IntechOpen

Biomedical Science, Engineering and Technology

Edited by Dhanjoo N. Ghista

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED
WEB OF SCIENCE™

BIOMEDICAL SCIENCE, ENGINEERING AND TECHNOLOGY

Edited by **Dhanjoo N. Ghista**

Biomedical Science, Engineering and Technology
http://dx.doi.org/10.5772/1020
Edited by Dhanjoo N. Ghista

Contributors

ANDREA MAHN, Ahmed Morsy, Magdalena Labieniec-Watala, Karolina Siewiera, Slawomir Gierszewski, Cezary Watala, Nathalie De Geyter, Rino Morent, Eduard A. Kurmazenko, Evgeniy Demin, Boris Morukov, Kamaletdinova Guzel, Nicolay Khabarovskiy, Kingsley Osagie Enohumah, Oluwafemi Omoniyi Oguntibeju, A. J. Esterhuyse, E. J. Truter, Vimal Selvaraj, Ambrose Jon Williams, William Bosl, John cason, Richard Leask, Alexander Emmott, Scott Cooper, Paul Jonak, Melissa L. Dick, Leonie Rouleau, Ying-Fang Yang, Michael Walsh, Barry O'Connell, Dragomir Stamenković, Dušan Kojić, Nikola Jagodić, Božica Antun Bojović, Djuro Koruga, Catauro, Anetor, Gloria Anetor, Segun Adeola, Ijeoma Esiaba, Cao, Lasse Dahl Jensen, Ziquan Cao, Eva-Maria Hedlund, Pegah Rouhi, Sharon Lim, Jennifer Honek, Ji Hong, Junwei Zhang, Kayoko Hosaka, Jason Ross, Caixia Yang, Hidenori Koyama, Tetsuya Yamamoto, Marta Wasilewska-Radwanska, Ewa Augustyniak, Ryszard Tadeusiewicz, Piotr Augustyniak, Deyin Lu, Bahram Khoobehi, Husheng Yan, Miao Guo, Keliang Liu, Jaroslav Turanek, Josef Mašek, Milan Raška, Miroslav Ledvina, Shuang-Qing Zhang, John Hancock, Mario Letelier, Juan Stockle, Porta, Ibrahim Banat, Letizia Fracchia, Massimo Cavallo, Maria Giovanna Martinotti, Guang Yang, Dhanjoo N. Ghista

© The Editor(s) and the Author(s) 2012
The moral rights of the and the author(s) have been asserted.
All rights to the book as a whole are reserved by INTECH. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECH's written permission.
Enquiries concerning the use of the book should be directed to INTECH rights and permissions department (permissions@intechopen.com).
Violations are liable to prosecution under the governing Copyright Law.

Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 3.0 Unported License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be foundat http://www.intechopen.com/copyright-policy.html.

Notice

Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in Croatia, 2012 by INTECH d.o.o.
eBook (PDF) Published by IN TECH d.o.o.
Place and year of publication of eBook (PDF): Rijeka, 2019.
IntechOpen is the global imprint of IN TECH d.o.o.
Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

Biomedical Science, Engineering and Technology
Edited by Dhanjoo N. Ghista

p. cm.

ISBN 978-953-307-471-9

eBook (PDF) ISBN 978-953-51-4376-5

We are IntechOpen,
the first native scientific
publisher of Open Access books

3,250+
Open access books available

106,000+
International authors and editors

112M+
Downloads

151
Countries delivered to

Our authors are among the
Top 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

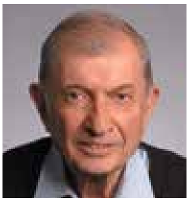
WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Meet the editor



The weaving of these diverse topics into an integrated book has been made possible by Professor Dhanjoo Ghista. His academic background and activities span engineering science, medical and health sciences, hospital administration and healthcare delivery system. A pioneer and authority in biomedical engineering (BME), he has a vision for professional development of BME, by incorporation of BME into the medical education curriculum, for offering of MD-PhD (bme) and MD-MBA (hospital administration) degrees, and setting up of BME departments in tertiary-care hospitals. He has also developed a pre-primary community healthcare system. Together, this will help to bring about (I) a technological evolution of novel medical procedures and devices, towards a higher order of translational medicine in patient care, as well as (II) more cost-effective hospital management and healthcare delivery system.

Contents

Preface XIII

Chapter 1 **Biomedical Engineering Professional Trail from Anatomy and Physiology to Medicine and Into Hospital Administration: Towards Higher-Order of Translational Medicine and Patient Care 1**
Dhanjoo N. Ghista

Part 1 Biomedical Science: Disease Pathways, Models and Treatment Mechanisms 49

Chapter 2 **Cell Signalling and Pathways Explained in Relation to Music and Musicians 51**
John T. Hancock

Chapter 3 **Chemical Carcinogenesis: Risk Factors, Early Detection and Biomedical Engineering 69**
John I. Anetor, Gloria O. Anetor, Segun Adeola and Ijeoma Esiaba

Chapter 4 **AGE/RAGE as a Mediator of Insulin Resistance or Metabolic Syndrome: Another Aspect of Metabolic Memory? 91**
Hidenori Koyama and Tetsuya Yamamoto

Chapter 5 **Mitochondria Function in Diabetes – From Health to Pathology – New Perspectives for Treatment of Diabetes-Driven Disorders 123**
Magdalena Labieniec-Watala, Karolina Siewiera, Slawomir Gierszewski and Cezary Watala

Chapter 6 **Red Palm Oil and Its Antioxidant Potential in Reducing Oxidative Stress in HIV/AIDS and TB Patients 151**
O. O. Oguntibeju, A. J. Esterhuyse and E. J. Truter

Chapter 7 **Medical Plant and Human Health 165**
Ahmed Morsy Ahmed

Chapter 8 ***In Vitro* Leukocyte Adhesion in Endothelial Tissue Culture Models Under Flow 191**
Scott Cooper, Melissa Dick, Alexander Emmott, Paul Jonak, Léonie Rouleau and Richard L. Leask

Chapter 9 **Pain in Osteoarthritis: Emerging Techniques and Technologies for Its Treatment 209**
Kingsley Enohumah

Part 2 Biomaterials and Implants 223

Chapter 10 **Non-Thermal Plasma Surface Modification of Biodegradable Polymers 225**
N. De Geyter and R. Morent

Chapter 11 **Poly(Lactic Acid)-Based Biomaterials: Synthesis, Modification and Applications 247**
Lin Xiao, Bo Wang, Guang Yang and Mario Gauthier

Chapter 12 **Multifunctional Magnetic Hybrid Nanoparticles as a Nanomedical Platform for Cancer-Targeted Imaging and Therapy 283**
Husheng Yan, Miao Guo and Keliang Liu

Chapter 13 **Arterial Mass Transport Behaviour of Drugs from Drug Eluting Stents 301**
Barry M. O'Connell and Michael T. Walsh

Chapter 14 **Biosurfactants and Bioemulsifiers Biomedical and Related Applications – Present Status and Future Potentials 325**
Letizia Fracchia, Massimo Cavallo, Maria Giovanna Martinotti and Ibrahim M. Banat

Chapter 15 **Contact Lenses Characterization by AFM MFM, and OMF 371**
Dušan Kojić, Božica Bojović, Dragomir Stamenković, Nikola Jagodić and Đuro Koruga

Chapter 16 **Synthesis and Characterization of Amorphous and Hybrid Materials Obtained by Sol-Gel Processing for Biomedical Applications 389**
Catauro Michelina and Bollino Flavia

Part 3 Biomedical Engineering 417

Chapter 17 **Diabetes Mechanisms, Detection and Complications Monitoring 419**
Dhanjoo N. Ghista, U. Rajendra Acharya, Kamlakar D. Desai, Sarma Dittakavi, Adejuwon A. Adeneye and Loh Kah Meng

Chapter 18 **Domain-Specific Software Engineering Design for Diabetes Mellitus Study Through Gene and Retinopathy Analysis 447**
Hua Cao, Deyin Lu and Bahram Khoobehi

Chapter 19 **A Shape-Factor Method for Modeling Parallel and Axially-Varying Flow in Tubes and Channels of Complex Cross-Section Shapes 469**
Mario F. Letelier and Juan S. Stockle

Chapter 20 **CSA – Clinical Stress Assessment 487**
Sepp Porta, Gertrud W. Desch, Harald Gell, Karl Pichlkastner, Reinhard Slanic, Josef Porta, Gerd Korisek, Martin Ecker and Klaus Kisters

Chapter 21 **Neurotechnology and Psychiatric Biomarkers 511**
William J. Bosl

Chapter 22 **Life Support System Virtual Simulators for Mars-500 Ground-Based Experiment 535**
Eduard Kurmazenko, Nikolay Khabarovskiy, Guzel Kamaletdinova, Evgeniy Demin and Boris Morukov

Chapter 23 **Educational Opportunities in BME Specialization - Tradition, Culture and Perspectives 559**
Wasilewska-Radwanska Marta, Augustyniak Ewa, Tadeusiewicz Ryszard and Augustyniak Piotr

Part 4 Biotechnology 585

Chapter 24 **Poly (L-glutamic acid)-Paclitaxel Conjugates for Cancer Treatment 587**
Shuang-Qing Zhang

Chapter 25 **Hydrophobic Interaction Chromatography: Fundamentals and Applications in Biomedical Engineering 603**
Andrea Mahn

Chapter 26 **Development and Engineering of CSαβ Motif for Biomedical Application 629**
Ying-Fang Yang

Chapter 27 **Application of Liposomes for Construction of Vaccines 653**
Jaroslav Turánek, Josef Mašek, Milan Raška and Miroslav Ledvina

Chapter 28 **iPS Cells: Born-Again Stem Cells for Biomedical Applications 679**
Ambrose Jon Williams and Vimal Selvaraj

Chapter 29 **Genetic Modification of Domestic Animals for Agriculture and Biomedical Applications 697**
Cai-Xia Yang and Jason W. Ross

Chapter 30 **Animal Models of Angiogenesis and Lymphangiogenesis 727**
L. D. Jensen, J. Honek, K. Hosaka, P. Rouhi, S. Lim, H. Ji, Z. Cao, E. M. Hedlund, J. Zhang and Y. Cao

Chapter 31 **Ethical and Legal Considerations in Human Biobanking: Experience of the Infectious Diseases BioBank at King's College London, UK 761**
Zisis Kozlakidis, Robert J. S. Cason, Christine Mant and John Cason

Part 5 Physiological Systems Engineering in Medical Assessment 779

Chapter 32 **Cardiac Myocardial Disease States Cause Left Ventricular Remodeling with Decreased Contractility and Lead to Heart Failure; Interventions by Coronary Arterial Bypass Grafting and Surgical Ventricular Restoration Can Reverse LV Remodeling with Improved Contractility 781**
Dhanjoo N. Ghista, Liang Zhong, Leok Poh Chua, Ghassan S. Kassab, Yi Su and Ru San Tan

Chapter 33 **Renal Physiological Engineering – Optimization Aspects 815**
David Chee-Eng Ng and Dhanjoo N. Ghista

Chapter 34 **Lung Ventilation Modeling for Assessment of Lung Status: Detection of Lung Disease and Indication for Extubation of Mechanically-Ventilated COPD Patients 831**
Dhanjoo N. Ghista, Kah Meng Koh, Rohit Pasam and Yi Su

Chapter 35 **Physiological Nondimensional Indices in Medical Assessment: For Quantifying Physiological Systems and Analysing Medical Tests' Data 851**
Dhanjoo N. Ghista

Preface

Biomedical Sciences (from anatomy, physiology and molecular biology to pathology) provide the information and knowledge base for biomedical engineering and technology. Formulation of biological and physiological mechanisms and correlates of organ functions, disorders and disease states in biomedical engineering terms makes them more clearly defined in terms of equations, formulas and indices. From bio-physiological disease mechanisms, we can proceed to engineering analysis and formulations of functions of physiological systems, and define normal and pathological ranges of physiological systems operations. This in turn leads to analysis of physiological systems' functional tests data or medical tests data, for carrying out medical diagnosis and prescribing medical treatments.

In this book, we start with chapter 1 on the biomedical engineering (BME) professional trail. Then, in **Section 1,** we deal with the biomedical sciences of disease pathways and mechanisms of action of treatments.

For biomedical engineering (BME) to be a professional discipline, we have addressed (in chapter 1) the professional needs of anatomy and physiology, medicine and surgery, hospital performance and management. The role of BME in Anatomy is to demonstrate how anatomical structures are intrinsically designed as optimal structures. In Physiology, the BME formulation of physiological systems functions can enable us to characterize and differentiate normal systems from dysfunctional and diseased systems. For BME in Medicine, we formulate the engineering systems analyses of physiological and organ system functions and medical tests, in the form of differential equations (Deqs), expressing the response of the organ system in terms of monitored data. The parameters of the Deq are selected to be the organ system's functional performance features. The normal and dysfunctional ranges of these parameters can enable reliable medical diagnosis, such as diagnosis of lung disease states or diagnosis of persons at risk of being diabetic. In Surgery, we can develop the criteria for candidacy for surgery, carry out pre-surgical analysis of optimal surgical approaches, and design surgical technology and implants. In Hospital management, we can develop measures of cost-effectiveness of hospital departments, budget development and allocation, such that no hospital department has a cost-effective index below a certain specific value. This chapter provides the basis of how biomedical engineering can be employed (i) to provide a new approach to the study of anatomy,

(ii) in the formulation of physiological systems' functional indices and their applications in medicine, and (iii) in combination with operations research methods in hospital management. All of this can be carried out by introducing biomedical engineering courses in the MD-PhD (BME) curriculum and biomedical engineering departments in tertiary care medical centers.

Section 1 is on Disease Pathways, Models and Treatment Mechanisms. We start with cell signalling (in chapter 2), which is an extremely important aspect of modern biology, involving control of cellular events in response to extracellular factors. In this chapter, it is suggested that music has many parallels with the principles of cell signalling. This chapter discusses (i) signalling between organisms and the production of signals, (ii) signalling systems, receptors and degeneracy, and (iii) threshold signalling levels, with timings and phrasing.

Chemical Carcinogenes is an important concern for us. In chapter 3, we discuss: cell regulatory mechanisms and their disruptions in cancer cells caused by carcinogens; mechanism of oxidative stress and DNA damage due to micronutrient deficiency; biomarkers usage in measurement of external dose, and determination of altered structure and function of cells as a marker of chemical carcinogenesis; the bioengineering technologies associated with these processes and measurements.

We next deal, in chapter 4, with the concept of Metabolic memory, of (i) early metabolic control on longer cardiovascular outcomes, and (ii) the underlying pathophysiology of metabolic syndrome and insulin resistance. The potential mechanisms for propagating this "memory" are the non-enzymatic glycation of cellular and tissue proteins, which are conceptualized as advanced glycation end-products (AGEs), the generation of which is implicated to be associated with increased oxidative stress and hyperglycemia. AGEs, with their receptors potentially mediate molecular and cellular pathways leading to metabolic memory. Interaction of the RAGE with AGEs leads to crucial biomedical pathway generating intracellular oxidative stress and inflammatory mediators, which could result in further amplification of the pathway involved in AGE generation. By utilizing genetically engineered mouse models, emerging evidence suggests that AGE/RAGE axis is also found to be profoundly associated with non-diabetic pathophysiological conditions, including 1) atherogenesis, 2) angiogenic response, 3) vascular injury, and 4) inflammatory response, many of which are now implicated in metabolic syndrome.

Next, in chapter 5, we present Mitochondria Function in Diabetes, on (i) various mechanisms present in mitochondria that lead to the development of diabetes, (ii) modulation of the "vicious circle" established between mitochondria, oxidative stress and hyperglycemia, and (iii) application of some agents possessing anti-glycation properties to reduce glycation phenomenon and to increase the antioxidant defense system by targeting mitochondria.

Infection by HIV and/ or TB is known to cause persistent chronic inflammation. There is evidence that patients infected with HIV and/ or TB are under chronic oxidative stress with a resultant decrease in endogenous and nutritional antioxidants as well as other micronutrients. Oxidative stress due to overproduction of free radicals and antioxidant deficiency, causes damage to vital biological macromolecules and organs and further contributes to disease complications, disease progression and morbidity. In chapter 6, we discuss the role of red palm oil from the African palm (Elaeis guinensis) in reducing oxidative stress. It is proposed that red palm oil supplementation could effectively scavenge free radicals and increase total antioxidant capacity, with the potential to (i) reduce disease progression and its complications, (ii) increase survival and (iii) improve the general wellbeing of people living with TB and HIV/AIDS.

Recent researches show that medical plants have ecological functions that have potential medicinal effects for humans. Diabetes mellitus is the major endocrine disorder responsible for renal failure, blindness or diabetic cataract, poor metabolic control, increased risk of cardiovascular disease including atherosclerosis and AGE (advanced glycation end) products. Antioxidants play an important role to protect against damage by reactive oxygen species, and their role in diabetes has been evaluated. Many plant extracts and products are shown to possess significant antioxidant activity. Accordingly, in chapter 7, we discuss some fundamental aspects of phytomedicinal plants with an overview of those plants that have received considerable use and attention in diabetes treatment.

Atherosclerosis, causing thrombosis (atherothrombosis), is the underlying pathology of the vast majority of cardiovascular diseases. It is responsible for up to 80% of all deaths in diabetic patients. Atherothrombosis is clinically manifested as coronary artery disease (heart attacks), stroke, transient ischaemic attack, and peripheral arterial disease. The atherosclerotic process starts early in life and, in almost one-third of all people, can progress to a complicated atheromatus plaque that generates thrombosis and blockage of blood supply. These plaques preferentially develop in regions of complex blood flow, such as bifurcations and regions of curvature. Local variations in hemodynamic forces, in particular wall shear stress (WSS), have been hypothesized to cause focal endothelial cell (EDC) dysfunction leading to a pro-inflammatory environment prone to atherosclerotic lesion development. These WSS profiles can manifest morphological and phenotypical changes in EDCs through a complex pathway of mechanotransduction. In chapter 8, we provide an understanding of endothelial-leukocyte interactions in atherogenesis and plaque stability, based on 3-d culture in vitro modeling.

Now, we come to chapter 9. Osteoarthritis (OSA) is a heterogenous condition that involves not only the articular cartilage but also an adaptive response of the bone and the synovium to a variety of environmental, genetic and biomechanical stresses. This chapter deals with pain in osteoarthritis: (i) mechanisms involving activation of nociceptors (naked nerve endings close to small blood vessels and mast cells) and nociceptive stimuli causing tissue damage; (ii) pathophysiology of gradual proteolytic degradation of the joint cartilage matrix, catalysed by metalloproteinases; (iii) receptors involved in the mechanisms of action for acute pain: a-amino-3-hydroxy- 5-methyl-isoxazole-4-propionic acid (AMPA) receptors; (iv) receptors of importance in the sensation of chronic pain: N-methyl- D-aspartate (NMDA) receptors; the activation of NMDA receptors causes the release of peptide neurotransmitter SP, which amplifies the pain by causing the spinal neurons transmitting the pain to be easily stimulated; (v) modes of treatment for OA for decreasing pain and improving function through analgesics, non-steroidal anti-inflammatory drugs and joint injections, and surgery involving joint replacement with plastic, metal or ceramic implants.

In Section 2, we deal with Biomaterials and Implants. Among biomaterials, we have included herein: (i) non-thermal plasma surface modification of biodegrable polymers employed in sutures and biodegradable scaffolds, (ii) synthesis and surface modification of polylactic acid (PLA) based biomaterials employed in tissue engineering scaffolds and drug delivery systems, and (iii) multifunctional magnetic nanoparticles as contrast agents for magnetic resonance imaging (MRI) and as carriers for drug delivery.

During the past two decades, there has been a considerable interest in the development and production of biodegradable polymers. Besides their use as packaging materials, biodegradable polymers play a major role in biomedicine as sutures, temporary prostheses and drug delivery vehicles. Biodegradable polymers have also been studied as three-dimensional porous structures (scaffolds) in the tissue engineering domain. The ultimate goal of this technology is to generate completely biocompatible tissues that can be used to replace damaged or diseased tissues in reconstructive surgery. Ideally, the scaffold material should be able to support initial cell growth and further proliferation, and should have the ability to biologically degrade over time while leaving behind a reproduced functional tissue. The success of polymeric biodegradable scaffolds is however determined by the response it elicits from the surrounding biological environment and this response is largely governed by the surface characteristics of the scaffold. In order to obtain the desired surface properties, the use of non-thermal plasmas for selective surface modification has been a rapidly growing field. Chapter 10 presents recent advances in plasma-assisted surface modification of biodegradable polymers.

Poly(lactic acid) (PLA) has gained increasing attention as a polyester material. Chapter 11 deals with (i) synthesis of PLA, (ii) modification of PLA to improve its properties, and (iii) biomedical application of PLA. For PLA synthesis, different synthetic methods are described, especially direct polycondensation and ring-opening polymerization, which are presently the main synthetic methods used to obtain PLA. In order to be suitable for specific biomedical applications, PLA has been modified mainly concerning its bulk properties and surface chemistry. To achieve this, both chemical modification and physical modification have been tried, involving the incorporation of functional monomers with different molecular architectures and compositions, the tuning of crystallinity and processibility via blending and plasticization. PLA has been employed to manufacture tissue engineering scaffolds, drug delivery system materials, and bioabsorbable medical implants, due to its bioresorbability and biocompatibility in the human body.

Multifunctional magnetic nanoparticles (MFMNPs) possess unique magnetic properties and the ability to function at the cellular and molecular level of biological interactions, making them an attractive platform as contrast agents for magnetic resonance imaging (MRI) and as carriers for drug delivery. Nanomedical platforms, based on superparamagnetic iron oxide nanoparticles, have useful applications, for magnetic targeting, contrast enhancement in magnetic resonance imaging, and hyperthermia in response to an external alternating magnetic field. For biomedical applications, superparamagnetic iron oxide nanoparticles are usually composed of a single domain magnetic core (less than 20 nm in diameter) and a hydrophilic coating that enables the nanoparticles to be biocompatible and dispersible in water. Chapter 12 deals with: (i) a multifunctional nanoplatform of a superparamagnetic Fe_3O_4 core and a block copolymer (methoxy poly(ethylene glycol)-*b*-poly(methacrylic acid-*co*-*n*-butyl methacrylate)-*b*-poly(glycerol monomethacrylate), denoted MPEG-*b*-P(MAA-*co*-*n*BMA)-*b*-PGMA) shell; (ii) the loading of anticancer agent adriamycin (ADR) into the nanocarrier, release of loaded drug molecules, and enhancement of delivery efficiency and cancer specificity by anchoring folic acid (FA) onto the nanoparticles for recognition by folate receptors on surface of cancer cells; (iii) fabrication of a nanoplatform with a magnetite core, for the targeted delivery of carboxyl group-containing drugs using anticancer agent chlorambucil; (iv) loading of chlorambucil into the nanocarrier by a combination of ionic and hydrophobic interactions, with the release rate of loaded chlorambucil at pH 7.4, and increasing significantly at acidic pH.

In chapter 13, on drug eluting stents (DES) deployed in blocked arteries, we have discussed how the drug coating suppresses the process of smooth muscle cell migration from the medial layer of artery to the lumen to thereby mitigate vascular restenosis. This chapter (i) addresses the mechanisms and biological implications of mass transport of drugs from the stents into the arterial wall, , and (ii) provides a validated numerical model to simulate arterial drug concentrations after stent implantation and the transport of therapeutic levels of drugs within the artery wall.

In Chapter 14, we discuss how biosurfactants application on medical insertion devices (such as urethral catheters) serve as anti-adhesive coating agents against pathogens for prevention of microbial biofilm formation on these devices. The antimicrobial activity property of biosurfactants disrupts membranes, leading to cell lysis against bacterial pathogens, fungi and viruses. Biosurfactants also serve as anti-inflammatory, anti-tumour, immunosuppressive and immunomodulating agents. They can be employed: (i) in self-assembly, human cells stimulation and differentiation, interaction with stratum corneum lipids, cell-to-cell signalling, and hemolytic activity; (ii) in biotechnology and nanotechnology, as means of introducing foreign genes into target cells due to their high transfection efficiency, low toxicity, ease of preparation and targeted application; (iii) in the enhancement of the gene transfection efficiency of cationic liposomes, in gene therapy and drug delivery.

Contact Lens (COL) production is one of the fastest growing sectors in medical device industry. Supporting this high development trend requires non-destructive surface analysis methods on the nanometer scale, to further enhance production quality as well as therapy efficiency. The magnetic property of contact lenses (COL), as optical material, has influence on electrical and magnetic light signals properties. This multimodal research comprises measurement of intermolecular interactions on the basis of optical, mechanical, morphological and magnetic properties of contact lens material. As discussed in Chapter 15, the approach to COL structure and function analysis on the molecular level requires the usage of high precision technologies, such as atomic force microscopy (AFM) and magnetic force microscopy (MFM), in order to describe and quantitatively measure the influence of processing parameters on the final surface quality.

The introduction of an implant in a living body causes inflammation phenomena and also frequently triggers infection processes. Those problems can be overcome by using local drug delivery methods to confine pharmaceuticals, as antibiotics, anti-inflammatory, and anti-carcinogens. In this context, the sol-gel process has been widely used in the preparation of organic-inorganic hybrid materials, non-linear optical materials, and mesomoporous materials. This family of organic-inorganic hybrid materials has interesting properties, such as molecular homogeneity, transparency, flexibility and durability. Such hybrids are promising materials for applications as biomaterials and contact lenses. Chapter 16 deals with synthesis and characterisation methods of organic-inorganic hybrid biomaterials to be used for controlled drug delivery applications, with a focus on the science of sol-gel processing, involving areas of physics (e.g. fractal geometry and percolation theory) and chemistry (mechanisms of hydrolysis and polycondensation) and ceramics (sintering and structural relaxation).

Section 3 is on Biomedical Engineering. Chapter 17 in this section is on Diabetes mechanisms, detection and monitoring. Diabetes mellitus (DIM), defined as a state chronic hyperglycaemia resulting from absolute or relative impaired insulin synthesis/secretion and/or insulin action, remains the most common endocrine disorder of carbohydrate and lipid metabolism, worldwide. This chapter develops an enquiry into diabetes from many angles: (i) the cellular and molecular mechanisms of development of diabetes and its complications; (ii) bioengineering of the glucose-insulin regulatory system, and its employment in the modeling of the oral glucose tolerance test data, to detect diabetes as well as persons at risk of being diabetic; (iii) analysis of heart rate variability signals to depict diabetes; (iv) analysis of retinal and plantar images to characterize diabetes complications; (v) diagnosis of diabetic autonomic neuropathy complication by means of an integrated index composed of indices based on heartrate variability power spectrum plots of normal subjects, diabetic patients and ischemic heart disease patients; (v) application of the glucose-insulin regulatory system to formulate an insulin delivery system for controlling blood sugar.

Software engineering designs and practices differ widely among various application domains. Chapter 18 is on high performance software engineering design for bioinformatics and more specifically for diabetes mellitus study through gene and retinopathy analysis. Complex gene interaction study offers an effective control of blood glucose, blood pressure and lipids. Early detection of retinopathy is effective in minimizing the risk of irreversible vision loss and other long-term consequence associated with diabetes mellitus.

The main objective of Chapter 19 is to present a method for modeling an ample variety of flows in tubes and channels, considering steady, non-steady, Newtonian and non-Newtonian flows. The method is based upon a specific shape factor that is imposed in the solution for the velocity field, thus making it possible to impose boundary conditions that determine tube or channel contour shapes. In this way, flows in tubes and channels of non-circular geometry or axially-varying cross-sections can be analyzed by means of the velocity, pressure and shear-stress fields. Knowledge of these flows is useful in the study of surgical interventions in pathological arteries and veins, and in microfluidics applications. In particular, zones of low velocity and low shear stress can be determined, which are considered risk zones related to the development of stenosis and other artery diseases. Specific applications included are (1) flow in straight tubes of constant non-circular cross-section: Newtonian unsteady, and steady plastic flows, (2) axially-varying flows in conduits: Newtonian flow in round tubes of arbitrarily axially- varying cross-section, and steady plastic flow in undulating channels.

Adrenaline and Noradrenaline changes incite changes in blood pH, buffer parameters like HCO3, lactate and blood glucose as well as electrolytes like K, Na, Ca and Mg. These parameters constitute interdependent stress-hormone effects. They can be put on organisms like a data-net, by especially designed online software, (i) to assess their workload, stress compatibility and stress duration, intensity and the kind of stress, (ii) by collecting 100 microliters of capillary blood within 3 minutes, using transportable intensive care equipment. In chapter 20 on Clinical Stress assessment, this approach is employed to: 1) determine the impact of sport training and military training units, fire fighters and others, to link changes of blood parameters not only with sportive success but also to predict success chances before competition; 2) determine mental stress as well as stress by combined psychical and physical workload; 3) determine idiosyncrasies of diabetic metabolism, namely importance of mineral deficiencies in type2 diabetics as well as new aspects of metabolic differences between hypertonic and normotonic diabetics; 4) mathematically develop “situation dependent values”, to assess responses to simulated stress, and predict ability to sustain stress; 5) quantify predictions of success chances in competing animals like horses or camels, and provide stress documentations for prevention of cruelty to animals.

Neuropsychiatric disorders account for over 30% of all years lived with disability (YLD), globally. The combination of relatively easy-to-administer psychiatric assessments and emerging health information technology can aid in the treatment of psychiatric disorders. Neurotechnology, that enables psychiatric conditions to be estimated from physiological measurements and more frequent feedback on the course of therapy, would be useful for treating neuropsychiatric disorders. Also, the development of neurotechnology, that can effectively measure changes in brain function due to administration of drugs, can be very useful during the long and expensive drug testing process. If brain function and behavior are mirrors of each other, then biomarkers of mental disorders may be hidden in subtle and complex patterns of neurobiological data. A key challenge in clinical neuroscience is to discover the relationship between brain function and behavioral patterns that are indicative of mental disorders. The challenge for biomedical engineers is hence to design devices and algorithms that enable affordable measurements of brain function that can be used in clinical setting for assessing neuro-psychiatric disorders.

Chapter 21 reviews recent advances in neuroscience. The physics of complex systems and neurotechnology together may enable innovations in the diagnosis, classification and management of psychiatric disorders. Complex neurophysiological mechanisms underlying abnormal mental function cannot be understood by reduction to simple measures. Measurements of brain electrical activity with EEG has long been a valuable source of information for neuroscience research, yet underutilized for clinical and diagnostic applications. To fully exploit this data, methods for discovering nonlinear patterns and deeper understanding of the relationship between emergent complex signal features and the underlying neurophysiology are needed. Analysis of EEG signal complexity and transient synchronization may reveal information about local neural structure and long-range communication between brain regions. Research suggests that patterns in these EEG signal features may contain key biomarkers of abnormal information processing that is a central characteristic of many mental disorders. The development of novel EEG sensors, with improved resolution (together with new algorithms), promises continued improvement in the ability to measure subtle variations in brain function and yield a new window into the mind.

Mars manned mission requires resolution of problems on the ground with test subjects, related to crew life-support and psychological stability. In chapter 22, we deal with life support system virtual simulators for Mars-500 Ground-based experiment. In order to make interplanetary missions a reality, it is necessary to provide special crew's trainings. However use of full-scale systems at first phases of ground simulation of spaceflight to Mars is extremely complicated and economically unprofitable. A more rational approach is (i) the application of standard system virtual simulators interacting with simulation models for both environment and crew as a load component, and (ii) integrated in a single Hardware/Software Complex for Serving Operational Systems (HSCSOS) by crew, intended for system functioning in normal, off-normal, emergency situations in systems and deviation of environment controllable parameters from specified values. An additional biomedico-engineering system can be incorporated in the HSCSOS hardware architecture to perform psycho-physiological tests. This chapter provides analysis of all possible approaches to development of such complexes based on simulation of long-duration space missions. The results can be used in development of similar hardware/software complexes to analyze complicated human-machine interaction and specialist training for various-purpose Man-Made Ecosystems (MMES).The final chapter 23 in this section describes the traditions and the present status of medical physics and biomedical engineering education in Poland. A detailed history of the development of these specializations is provided with the example of the Multidisciplinary School of Engineering in Biomedicine founded in 2005 at the Akademia Gorniczo-Hutnicza (AGH) University of Science and Technology in Krakow. This program of studies incorporates a single 7-semester track leading to the First (Undergraduate) Degree (Bachelor's/Engineer's); five domain-oriented 4-semester tracks leading to the Second (Graduate) Degree (Master's), and a single 8-semester track leading to the Third Degree (Doctor's). The program provides special adaptation mechanisms to develop students' connection to prospective workplaces. Considerable emphasis is placed on specific characteristics of BME-related corporate culture that requires mutual understanding and good cooperation within multidisciplinary teams striving for technical excellence. The chapter also describes opportunities and perspectives of all BME-teaching institutions in Poland. The syllabi and curricula of the degree programs are included in the Appendix.

Now we start the next **Section 4 on Biotechnology.**

The preparation of polymer-anticancer drug conjugates is an effective way to improve the efficacy and decrease the toxicity of anticancer drugs. Chapter 24 deals with polymer-drug conjugates, which are made by combining a suitable polymeric carrier, a biodegradable linker and a bioactive anticancer agent, to form the basis of a new generation of anticancer agents. Poly (L-glutamic acid)-paclitaxel conjugate is a polymer-drug conjugate that links anticancer agent paclitaxel (PTX) to poly (L-glutamic acid) (PG). PG-PTX conjugate can improve the anticancer activity and the pharmacokinetic properties of PTX.

Hydrophobic interaction chromatography (HIC) is a powerful technique used for separating homologous proteins, receptors, antibodies, recombinant proteins and nucleic acids. Macromolecule retention in HIC is promoted by hydrophobic interactions between the HIC support and the macromolecule, and it is governed by an entropy change. The thermodynamics fundamentals of protein retention in HIC are discussed in this chapter 25. The strength of the interaction depends mainly on the properties of the HIC support and on the macromolecule hydrophobicity, which can be defined by different approaches. The hydrophobic interaction is weakened by a decrease in the ionic strength in the mobile phase, thus producing the elution of the macromolecule. The effect of the type and concentration of salt has been modeled through a thermodynamic model that considers macromolecule retention due to electrostatic and hydrophobic interactions. The outcome of a HIC process is a chromatogram, which can be described by the dimensionless retention time (DRT) of a macromolecule. HIC constitutes a purification tool suitable for biomedical applications, such as purification of vaccines, therapeutic proteins, plasmids and antibodies. In addition, the use of chromatography in high-throughput studies, such as proteomics and protein interactomics, is increasing.

Protein scaffolds have been employed as frameworks for innovative peptide drug development. New functions can be introduced to protein scaffolds through engineering processes. The antibody scaffold is one of the most extensively studied scaffolds. Although it is widespread in biomedical applications, the disadvantages of antibody stagnate its development in biomedical applications. In recent years, there is an urgent demand for new promising protein scaffolds in biomedical applications. The cysteine-knot scaffold demonstrates a rigid structure and ultra-stable characteristics. The proteins containing the scaffold usually serve as the defender in the innate immunity of their host. These proteins exhibit low sequence identity, but share a common three-dimensional structure. The structure is stabilized and sealed with two to four disulfide bridges. The scaffold has been reported to be engineered and to exhibit new functions. For its excellent properties, it is believed that the scaffold can fit the required criteria and serve as a fundamental building block for peptide drug development. Proteins with CSαb motif widely exist in crops and vegetables; they affect physiological regulations, and have been employed as remedies in traditional Chinese therapies. In chapter 26, we discuss the possible stratagem and the bottlenecks to engineer the CSαb motif for biomedical applications.

Liposomes represent ideal carrier/delivery systems for the components of synthetic vaccines, due to their biodegradability and ability to retain and incorporate a variety of essential vaccine components simultaneously. Different synthetic vaccine components can be encapsulated within the aqueous cavities of liposomes (if hydrophilic) or associated with liposome bilayers (if at least partially hydrophobic in character). Furthermore, essential components can be attached to either internal or external outer leaflet membrane by electrostatic, covalent or metallochelation interactions. The most diverse synthetic vaccine components are typically adjuvants needed to provoke innate immune reactions (e.g. monophosphoryl lipid A [MPL A], CpG oligonucleotides, muramyl dipeptide [MDP] and analogues). In addition, these can be combined with antigens needed to provoke specific immunity such as soluble or membrane proteins, synthetic peptides and oligosacharide antigens. Finally, liposomes may present ligands to assist functional delivery of antigens and adjuvants to antigen-presenting cells necessary to invoke immunostimulation. Chapter 27 discusses applications of Liposomes for construction of vaccines. Owing to biodegradability and safety, liposomes are compatible with various routes of application (intranasal, intramuscular, intradermal, peroral, sublingual, etc.). This is the main advantage of liposomes over other adjuvants. Many new synthetic components like cationic lipids, neoglycolipids, activated lipids and metallochelating lipids are now available for construction of liposomal carriers tailored for specific antigen. New synthetic adjuvants are being designed and tested, e.g. compounds based on muramyl or norAbu-muramyl peptides, CpG oligonucleotides and MPL-A.

The potential for the participation of liposome-based recombinant vaccines in the human and veterinary vaccine market is very promising.

Embryonic Stem Cells (ESCs), the topic of chapter 28, have been a focus of biomedical research in regenerative medicine and tissue engineering for more than ten years, because of their potential to give rise to cells of all three germ layers, a property termed pluripotency. However, progress to clinical translation in this field faces significant obstacles that include immune incompatibility and ethical concerns surrounding the use of human blastocyst embryos and therapeutic cloning, which have led to several high- profile legal challenges to continued funding. It has been recently discovered that adult somatic cells, including easily-obtained fibroblasts and lymphocytes, can be directly reprogrammed back to a primordial state of being functionally identical to ESCs. These Induced Pluripotent Stem Cells (iPSCs) not only circumvent ethical obstacles to clinical use of ESCs, but also are isogenic and negate concerns of immune complications in patients. Additional iPSCs also provide optimal substrate for gene-specific targeting to fix the genetic defects and subsequently treat these diseases using regenerative approaches. Induced pluripotency has therefore significantly improved the potential of cell and tissue engineering and is poised to take it closer to translational regenerative medicine.

Chapter 29 is on Genetic modification of Domestic animals for Agriculture and Biomedical applications The production of genetically modified animals greatly improves their utility in agriculture, as biomedical research models of human diseases, for the production of recombinant pharmaceutical proteins, and for making organs with greater potential for xenotransplantation. While numerous strategies have been used in the production of transgenic large animals, cell-based transgenesis followed by somatic cell nuclear transfer (SCNT) is currently the most widely applied method. Novel strategies for making specific modifications to somatic cells are rapidly being developed that allow targeted, conditional and tissue specific modifications to the mammalian genome. Continued utilization of cell-based transgenesis followed by SCNT will require improvements in efficiency, particularly in the areas of making targeted genetic modifications and in SCNT. This chapter discusses current and expanding applications for transgenic domestic species, emerging strategies to improve targeted genetic modification frequency of somatic cells, and methods to improve the efficiency of SCNT.

Angiogenesis and lymphangiogenesis are involved in regulation of tissue growth during development, regeneration, and in adults. Furthermore, deregulated angiogenesis/lymphangiogenesis may result in the onset and progression of cancer, cardiovascular disease, obesity, diabetes, ophthalmological diseases and chronic inflammation. Knowledge of the fundamental mechanisms of angiogenesis and lymphangiogenesis can therefore assist us in identifying new molecular targets for therapeutic intervention against such pathologies. In vivo animal models are essential for the study of angiogenesis and lymphangiogenesis, and are employed to study vascular formation, remodeling, permeability, maturation, and stability. Chapter 30 provides methodological tools and fundamental information about the most commonly used animal models of angiogenesis and lymphangiogenesis, employed in angiogenesis research.

Chapter 31 describes the legal and ethical issues which surround the practice of biobanking human clinical materials. The storing of human tissues has long stimulated public debate due to a series of recent and historical scandals which have stimulated new legislation to regulate the practice. Examples of important criminal cases which have resulted in new legal requirements or clarification of ethical principles are highlighted in this chapter. Particular issues covered include issues of informed consent, which in modern history were described in the Nuremberg code and more recently in the Helsinki Declaration. These ethical and legislative aspects of biobanking in the UK are addressed in theory and in practice. We also describe the working practice of the Infectious Diseases BioBank in London (UK), as a model system which has the aim of facilitating and expediting medical research into infectious agents whilst meeting and often exceeding current day requirements.

The last Section 5 is on Physiological systems engineering in Medical assessment. It deals with formulation and analysis of physiological systems, identification of parameters representing systems performance, and combining these parameters into a system index which can be employed in medical assessment.

In Chapter 32 , we study the course (i) of cardiomyopathy diseased LVs (with myocardial infarcts) progressing to heart failure (HF) through LV remodeling and decreased LV contractility, and (ii) their recovery through surgical therapeutic interventions of CABG and Surgical ventricular restoration (SVR), by restoration of myocardial ischemic segments, reversal of LV remodeling and improvement in LV contractility. For this purpose, we first provide the methodology for detecting myocardial infarcts. Then, we characterize LV remodeling of cardiomyopathy diseased LVs (with myocardial infarcts) in terms of reduced change in curvedness from end-diastole to end-systole. In these LVs, there is also reduced contractility; so we provide an index for cardiac contractility, in terms of maximal rate-of-change of normalized wall stress, $d\sigma^*/dtmax$, and its decrease in an infarcted LV progressing to heart failure. We provide clinical studies of remodelled cardiomyopathy diseased LVs, in terms of reduced values of their curvedness index and contractility index. By way of CABG surgical intervention, we have presented the hemodynamic flow simulation of the CABG, and pointed out certain factors and sites of wall shear stresses that cause intimal damage of vessels and hyperplasia, as potential causes for decreased graft patency. We have shown that surgical ventricular restoration (SVR), in conjunction with CABG, is seen to benefit the ischemic-infarcted heart, by (i) restoration of cardiac remodeling index of ‘end-diastolic to end-systolic curvedness change’, (ii) reduction of regional wall stresses, and (iii) augmentation of the cardiac contractility index value.

In Chapter 33 , we present how the renal system is intrinsically designed as a functionally optimal system for filtration and regulation of urine concentration as well as renal clearance of unwanted metabolic substrates such as creatinine. This chapter analyses how the kidney performs its urine concentration ability, through various mechanisms, focussing on the countercurrent multiplier mechanism operating in the loop of Henle and its medullary vicinity. This mechanism is physiologically engineered to increase and critically maintain at steady-state the hyperosmolality of the renal medullary interstitium to as high as 4 times normal blood osmolality, so as to produce a highly concentrated urine in the interest of conserving needed water. The linear coupled system model of the Loop of Henle is seen to account for the salient physiological features of this mechanism quantitatively. Analysis of the way the kidney optimally handles waste metabolites, specifically creatinine (one of its most important functions) is carried out by using a single-compartment kinetic model, with continuous input of metabolic substrate. The continuous input case is aimed at reproducing the in-vivo physiological conditions under which the kidney functions within the body. The analytical solutions for the continuous input case are obtained, and predict that the body waste metabolite creatinine level in the blood varies with renal clearance as an inverse rectangular hyperbolic function. The kinetics of the kidney's handling of the metabolic waste product creatinine, shown by convolution analysis on the single-compartment model, demonstrates how the blood creatinine is bounded, and stabilizes to an asymptotically steady-state concentration. The analysis predicts reasonable estimates for the actual serum creatinine levels in the body, based on empirical renal clearance and creatinine substrate input parameters.

Next we present, in chapter 34 , Lung ventilation modeling for assessment of lung status, for detection of lung diseases and for prescribing an index for weaning of COPD patients on mechanical ventilation. In pulmonary medicine, it is important to detect lung diseases, such as chronic obstructive pulmonary disease (COPD), emphysema, lung fibrosis and asthma. These diseases are characterized in terms of lung compliance and resistance-to-airflow parameters. Another important endeavour of pulmonary medicine is mechanical ventilation of COPD patients and determining when to wean off these patients from the mechanical ventilator. In both these medical domains, lung ventilation dynamics plays a key role. So in this chapter, we develop the lung ventilation dynamics model in terms of monitored lung volume (V) and driving pressure (P_N), in the form of a differential equation with parameters of lung compliance (C) and resistance-to-airflow (R). Now, $P_N = P_L - P_{el0}$ (elastic *recoil p*ressure @ end-expiration) = P_m (pressure at mouth) - P_p (pleural pressure) – P_{el0} (= P_L @ end-expiration). We obtain the solution of this equation in the forms of lung volume (V) function of P_N, C and R. For the monitored lung volume V and pressure P_N data, we can evaluate C and R by matching the model solution expression with the monitored lung volume V and driving pressure P_N data. So what we have done here is to develop the method for determining the average values of C and R during the ventilation cycle. A more convenient way for detecting lung disease is to combine R and C along with some ventilator data (such as tidal volume and breathing rate) into a non-dimensional lung ventilator index (LVI). Then, we can determine the ranges of LVI for normal and disease states, and thereby employ the patient's computed values of LVI to designate a specific lung disease for the patient.

Now, in this methodology, we need to monitor (i) lung volume, by means of a spirometer, and (ii) lung pressure (P_N) equal to Pm (pressure at mouth) minus pleural pressure (Pp). The pleural pressure measurement involves placing a balloon catheter transducer through the nose into the esophagus, whereby the esophageal tube pressure is assumed to be equal to the pressure in the pleural space surrounding it. This procedure cannot be carried out non-traumatically and routinely in patients. Hence, for routine and noninvasive assessment of lung ventilation for detection of lung disease states, it is necessary to have a method for determining R and C from only lung volume data. So, then, we have shown how we can compute R, C and lung pressure values non-invasively from just lung volume measurement. Finally, we have presented how the lung ventilation modeling can be applied to study the lung ventilation dynamics of COPD patients on mechanical ventilation. We have shown how a COPD patient's lung C and R can be evaluated in terms of the monitored lung volume and applied ventilatory pressure. We have also formulated a lung ventilator index to study and assess the lung status improvement of COPD patients on mechanical ventilation, and to decide when they can be weaned off mechanical ventilation.

Now we finally arrive at an epochal concept of nondimensional physiological indices or physiological numbers. In medicine, for making diagnosis, many tests are needed. It may so happen that some tests results may be in the normal range, while some test results may be abnormal. So how is the doctor going to precisely decide how "sick" is the patient: is s/he at risk, or marginal, or very sick?

Hence, in the last chapter 35, we have presented a new concept of a Nondimensional Physiological Index (NDPI). This NDPI is made up a number of parameters characterizing an organ function and dysfunction or a physiological system function and disorder or an anatomical structure's property and pathology, in the format of a medical assessment test; the NDPI combines these parameters into one non-dimensional number. Thus, the NDPI enables the doctor to integrate all the parameters' values from the medical test into one non-dimensional index value or number. Then, by examining a large number of patients, we can determine the statistical distribution of that particular NDPI into normal and abnormal categories. This makes it convenient for the doctor to make the medical assessment or diagnosis.

Now for an organ or physiological system assessment test (such as a Treadmill test or Glucose tolerance test) or for an anatomical structure's property and pathology determination (such as for determining mitral valve calcification and pathology), the method of formulating and evaluating the NDPI (from the medical test) entails developing its bioengineering model's differential equation incorporating the parameters characterising the organ state or physiological system function or the anatomical structural constitutive property. These parameters are adroitly combined into a NDPI, so that the NDPI unambiguously conveys the normal and abnormal state of the organ or physiological system or the anatomical structure.

This bioengineering model's governing equation or its solution (involving the model parameters) is then applied to fit or simulate the monitored Test data of the physiological system or the anatomical structure. The model parameters are then evaluated (from the simulated solution to the Test data), and their ranges are determined for normal and abnormal states of the organ or physiological system or anatomical structure. Then, the NDPI (composed of the parameters of the organ function or physiological system function or the anatomical structural constitutive property) is also evaluated for normal and abnormal states of the patient's organ or physiological system or anatomical structure. In this way, we can apply these NDPIs to reliably diagnose the patient's health state, from preferably noninvasive medical assessment tests. In this chapter, we have developed a number of noninvasive medical tests involving NDPIs, based on biomedical engineering formulations of organ function, physiological system functional performance and anatomical structural constitutive property, to provide the means for reliable medical assessment and diagnosis. These tests include (i) some conventional tests, such as Treadmill and Glucose tolerance tests, as well as (ii) some of our newly formulated tests, to detect arteriosclerosis, aortic pathology, mitral valve calcification, and osteoporosis. Indeed, the development of NDPls for physiological systems and their clinical employment can revolutionise medical diagnosis and assessment.

Prof. Dhanjoo N. Ghista
Consultant, Department of Graduate and Continuing Education
Framingham University
Massachusetts, USA

1

Biomedical Engineering Professional Trail from Anatomy and Physiology to Medicine and Into Hospital Administration: Towards Higher-Order of Translational Medicine and Patient Care

Dhanjoo N. Ghista
*Department of Graduate and Continuing Education,
Framingham State University, Framingham, Massachusetts,
USA*

1. Introduction

1.1 Theme

For Biomedical Engineering (BME) to be a professional discipline, we need to address the professional needs of anatomy and physiology, medicine and surgery, hospital performance and management.

The role of BME in Anatomy is to demonstrate how anatomical structures are intrinsically designed as optimal structures. In Physiology, the BME formulation of physiological systems functions can enable us to characterize and differentiate normal systems from dysfunctional and diseased systems. In order to address Medical needs, we need to cater to the functions and disorders of organ systems, such as the heart, lungs, kidneys, and the glucose regulatory system. In Surgery, we can develop the criteria for candidacy for surgery, carry out pre-surgical analysis of optimal surgical approaches, and design surgical technology and implants. In Hospital management, we can develop measures of cost-effectiveness of hospital departments, budget development and allocation.

For BME in Anatomy, we depict how the spinal disc is designed as an intrinsically optimal structure. For BME in Medicine, we formulate the engineering system analyses of Organ system functions and medical tests,

- in the form of differential equations (DEqs), expressing the response of the organ system in terms of monitored data,
- in which the parameters of the DEq are selected to be the organ system's intrinsic functional performance features.

The solution of the organ system's governing DEq is then derived, and made to simulate the monitored data, in order to:

- evaluate the system parameters,
- and obtain the normal and disease ranges of these parameters.

These parameters can then be combined into a Non-dimensional Physiological Index (NDPI),

- by which the system can be assessed in terms of just one "number",
- whose normal and disease ranges can enable effective medical assessment.

Herein, we demonstrate how this methodology [1, 2] is applied to:

- Treadmill test, for evaluating cardiac fitness;
- Lung ventilation modelling, for assessment of status of mechanically ventilated COPD patients;
- Derive a Cardiac Contractility index, which can be determined non-invasively (in terms of auscultatory pressures) and applied to assess left ventricular contractile capacity;
- Glucose Tolerance tests, to detect diabetic patients and borderline patients at risk of becoming diabetic;
- Non-invasive determination of Aortic Pressure profile, systemic resistance and aortic elastance (*Eao*, to characterize the LV systemic load).

Finally, we have also shown the application of this concept and methodology to hospital management. There is a considerable (and hitherto under developed) scope for application of Industrial Engineering discipline for effective hospital administration, in the form of how to determine and allocate hospital budget to optimise the functional performances of all the hospital departments. This leads us to what can be termed as the *Hospital Management System*.

Herein, we have shown how to formulate a performance index (PFI) for ICU. This index divided by the Resource index gives us the cost-effectiveness index (CEI). The Management strategy is to maintain certain acceptable values of both PFI and CEI for all hospital departments, by judicious allocation of staff to the departments. This enables the determination of the Optimal Resource index (RSI) and hospital budget (HOB) to maintain a balance between PFI and CEI for all the hospital departments. This can constitute the basis of Hospital Management.

2. Anatomy: Spine analysed as an intrinsically optimal structure

2.1 Spinal vertebral body (an intrinsically efficient load-bearer)

The spinal vertebral body (VB) geometry resembles a hyperboloid (HP) shell (fig 2) which is loaded by compressive and torsional loadings, portrayed in fig 1 as resolved into component forces along its generators.

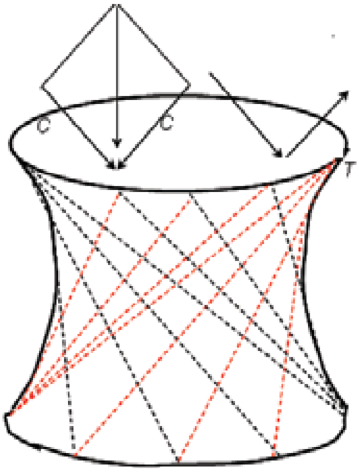
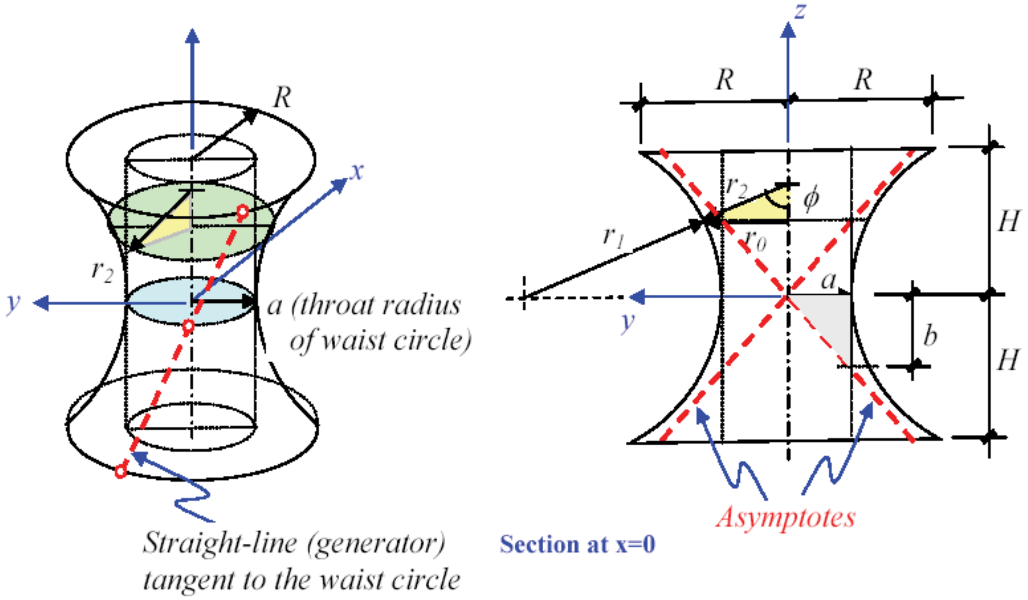


Fig. 1. Vertebral body, a hyperboloid (HP) shell formed of 2 sets of generators [3].

Stress analysis of the VB under Axial Compression [3]:

We now analyse for stresses in the HP shell (generators) due to a vertical compressive force P, as shown in figures 3 and 4.Assume that there are two sets of 'n' number of straight bars placed at equal spacing of ($2\pi a/n$) measured at the waist circle, to constitute the HP surface, as shown in figure 3 (right). Due to the axi-symmetric nature of the vertical load, no shear stresses are incurred in the shell, i.e. $\sigma_{\phi\theta}=0$, as in figure 3 (left). We then delineate a segment of the HP shell, and consider its force equilibrium (as illustrated in figure 4), to obtain the expressions for stresses N_ϕ and N_θ as depicted in figure 4.


R
x
r_2
y
a (throat radius of waist circle)
Straight-line (generator) tangent to the waist circle
Section at x=0
z
R
R
r_2
ϕ
r_1
r_0
a
y
b
H
H
Asymptotes

r_1: radius of curvature of meridian

r_2: slanted radius of horizontal section having radius r_o

Kinematic Relationship:

$$r_o = r_2\,(\sin\phi)$$

$$r_1 = -\left(\frac{b^2}{a^4}\right) r_2^3$$

Equation of HP curves:

$$\frac{x^2+y^2}{a^2} - \frac{z^2}{b^2} = 1$$

$$At\ x = 0,\quad \frac{y^2}{a^2} - \frac{z^2}{b^2} = 1$$

Equation of asymptotes:

$$z = \pm\left(\frac{b}{a}\right) y$$

Fig. 2. Geometry of a Hyperboloid (HP) shell**.** In the figure z = b, and y = a. We define tan β = a/b [3].

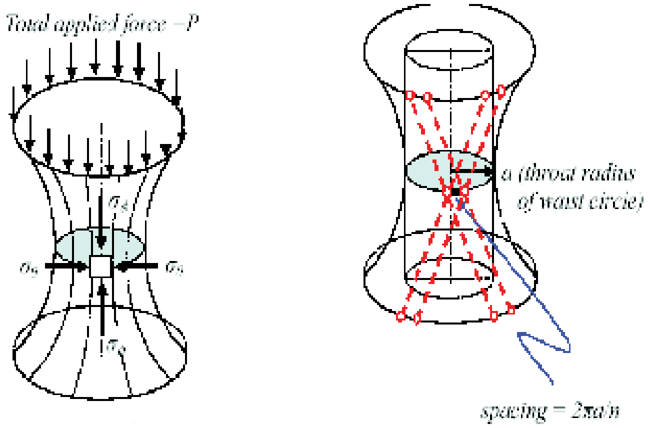

Total applied force – P
σ_ϕ
σ_θ
σ_θ
σ_ϕ
a (throat radius of waist circle)
spacing = 2πa/n

Fig. 3. Stress Analysis for Vertical Loading: Stresses at the waist section of the VB HP Shell: (a) stress components (b) equivalent straight bars aligned with the generators) placed at equal spacing to take up the stresses. In fig 3 (left) due to axi-symmetric vertical load, no shear stresses are incurred in the shell, i.e. $\sigma_{\phi\theta}=0$**.** In fig 3 (b), there are 2 sets of '*n*' number of straight bars placed at equal spacing of ($2\pi a/n$) measured at the waist circle, to constitute the generators of the HP surface [3].

Equilibrium of Forces on a Shell Segment under Vertical load P:

At the waist (r_o= a),

$$N_\phi = \frac{P}{2\pi a}, \text{ compressive}$$

Now since, $\frac{N_\phi}{r_1} - \frac{N_\theta}{r_2} = p_r$

Hence,

$$N_\theta = \left(\frac{a^2}{b^2}\right)\frac{P}{2\pi a} = \frac{P\tan^2\beta}{2\pi a}$$

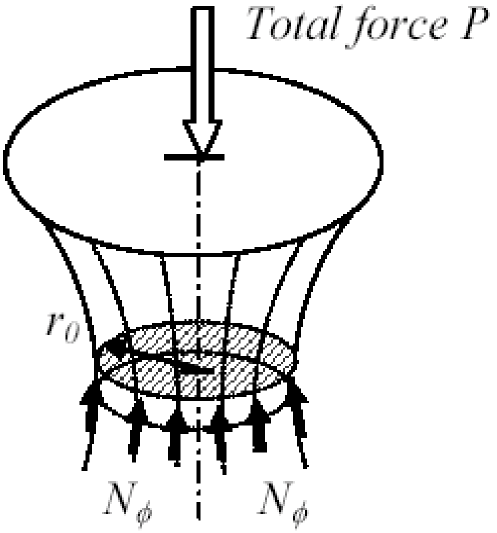
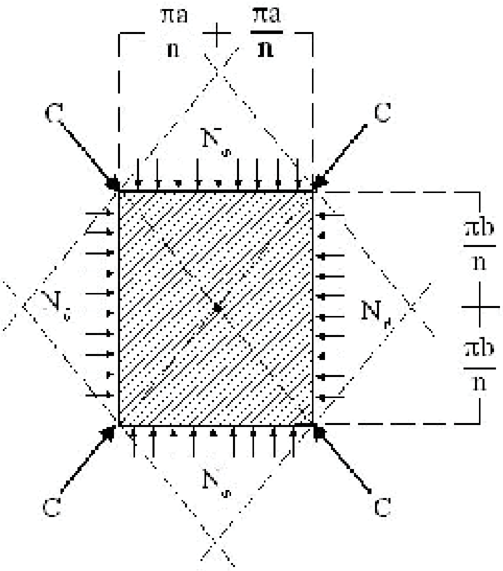

Total force P
r_0
N_ϕ
N_ϕ

Fig. 4. Equilibrium of Forces on a Shell Segment: Analysis for stresses N_ϕ *and* N_θ due to the vertical force *P* [3,2].

Then based on the analysis in Fig 5, we obtain the expression for the equivalent resultant compressive forces C in the fibre-generators of the VB HP shell. Thus it is seen that the total axial loading is transmitted into the HP-shell's straight generators as compressive forces. It is to be noted that the value of C is independent of dimensions R and a.



$N_{\phi} = P / 2\pi a$ and

$N_{\theta} = (a^2 / b^2)P / 2\pi a$

result in

$$C^2 = \left(N_{\phi}\frac{\pi a}{n}\right)^2 + \left(N_{\phi}\frac{\pi b}{n}\right)^2$$

$$C = \frac{P}{2nb}\sqrt{b^2 + a^2}$$

$$= \frac{P}{2n\cos\theta} = \frac{P(H^2 + a^2 - a^2)^{1/2}}{2nH}$$

Fig. 5. Equivalent compressive force C in the generators (corresponding to the stress-resultants acting on the shell element) equilibrating the applied axial loading [3,2].

It can be noted that the value of C is independent of dimensions R and a.

Stress analysis under Torsional loading [3]:

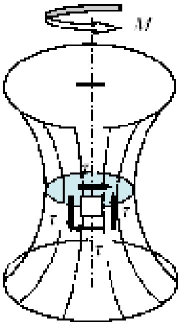
Next, we analyse the compressive and tensile forces in the HP shell generators when the VB is subjected to pure torsion (*T*).In this case (referring to **fig. 6)**, the normal stress resultants are zero, and we and only have the shear stress-resultant , as given by

$$N_{\phi} = N_{\theta} = 0 \qquad \text{and} \quad N_{\phi\theta} = \tau \text{t}$$

The equilibrium of a segment of the shell at a horizontal section at the waist circle (depicted in **figure** 6 **a**) gives the shear stress-resultant as follows:

$$[(\tau \cdot t)(2\pi a)]a = T \text{ , i.e., } N_{\phi\theta} = \frac{T}{2\pi a^2}$$

Stress Analysis for Torsional Loading M



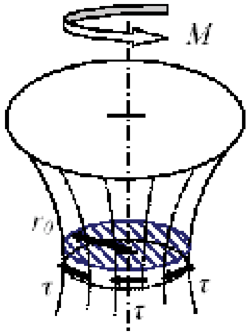
Equilibrium of a shell segment under torsion (M) and shear stresses (τ) (or shear stress-resultant $N_{\phi\theta}$)

For equilibrium,

$$(2\pi a)(\tau \cdot t)a = M$$

$$\tau = \frac{M}{2\pi a^2 t}$$

$$N_{\phi\theta} = \frac{M}{2\pi a^2}$$



Stresses in the HP shell element ($\sigma_\phi=\sigma_\theta=0$ and $\sigma_{\phi\theta}= \tau$) due to torsion M acting on the VB

Fig. 6. Stress analysis of the vertebral body under torsional loading [3].

Now, we consider an element at the waist circle as shown in **figure** 7. The equivalent compressive force (F_{cT}) and tensile force (F_{tT}), in the directions aligned to their respective set of shell generators, are given by

$$F_{cT}^2 = F_{tT}^2 = \left(N_{\phi\theta}\frac{\pi a}{n}\right)^2 + \left(N_{\phi\theta}\frac{\pi b}{n}\right)^2$$

or,

$$\left|F_{cT}\right| = \left|F_{tT}\right| = \frac{T}{2na\sin\beta}$$

wherein F_{cT} andF_{tT} are depicted as c and T respectively in figure 7.

$$= \frac{M(a^2 + a^2\cot^2\theta)^{1/2}}{2na^2}$$

$$= \frac{M}{2na\sin\theta}\left|T or C\right|$$

$$= \frac{M\cos ec\ \theta}{2na}$$

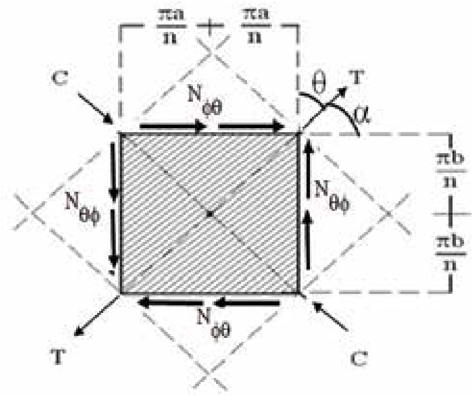


Fig. 7. Analysis of equilibrium of a shell element comprising of two intersecting generators: Expressions for tensile forces T and compressive forces C in the generators, indicate that torsion loading is also transmitted as axial compressive and tensile forces through the generators of the VB, which makes it a naturally optimum (high-strength and light-weight) structure [3].

Thus, a torsional loading on the VB HP shell is taken up by one set of generators being in compression and the other set of generators being in tension.

Equilibrium of a Shell Element Comprising of Two Intersecting Generators: The equivalent compressive forces (*C*) and tensile forces (*T*); in the shell generators (required to equilibrate the applied load), are given by:

$$|T| = |C| = \sqrt{\left(\tau \cdot t\left(\frac{\pi a}{n}\right)\right)^2 + \left(\tau \cdot t\left(\frac{\pi b}{n}\right)\right)^2} = \tau \cdot t\left(\frac{\pi}{n}\right)\sqrt{a^2+b^2} = \frac{M(a^2+b^2)^{1/2}}{2na^2}$$

2.2 Spinal disc optimal design (to bear loading with minimal deformation and maximal flexibility)

Fig 8 illustrates how the mechanism of how the spinal disc bears compression without bulging. It is seen that the nucleus pulposus plays the key role in this mechanism, as will be explained further in this section. Hence the absence of it in a denucleated disc causes the disc to bulge.

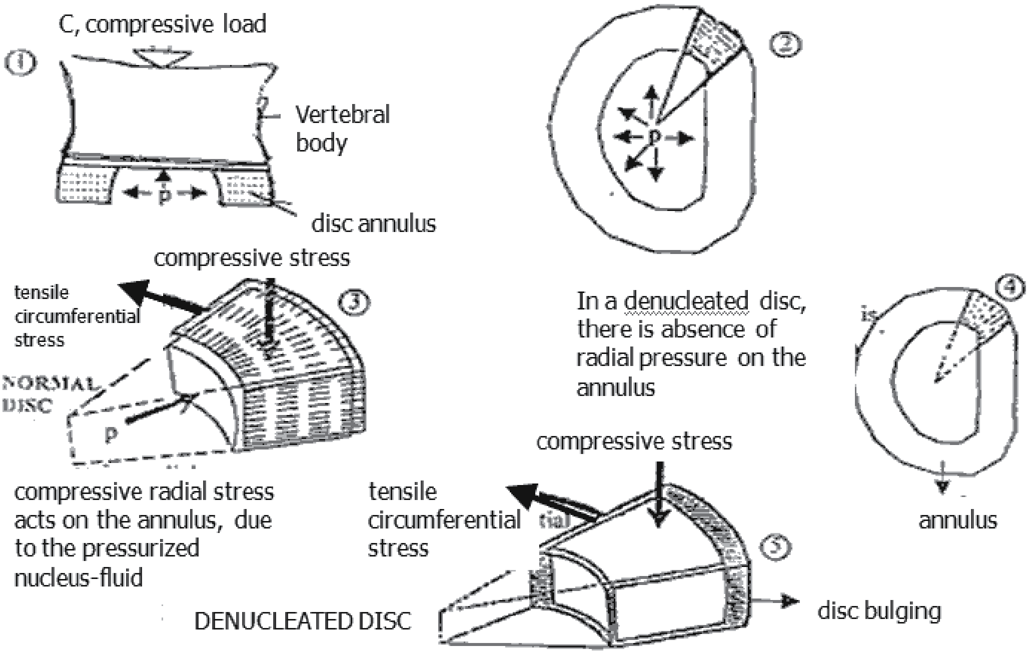

C, compressive load
①
Vertebral body
p
disc annulus
②
p
compressive stress
tensile circumferential stress
③
NORMAL DISC
p
In a denucleated disc, there is absence of radial pressure on the annulus
④
compressive stress
compressive radial stress acts on the annulus, due to the pressurized nucleus-fluid
tensile circumferential stress
annulus
⑤
DENUCLEATED DISC
disc bulging

Fig. 8. Mechanism of how the spinal disc bears compression without bulging.

Fig 9 displays the geometry and the deformation variables of the spinal disc. We now present the elasticity analysis of the disc to first obtain the radial, circumferential and axial stresses in terms of the disc deformations and annulus modulus *E*[4].

We next carry out the stress analysis of the disc under vertical loading *P* (fig 10), to obtain the expressions (equations 10) for the stresses in the disc annulus in terms of the load *P* , pressure p in the nucleus-pulposus, and the disc dimensions (a and b) [4].

We then derive the expressions for the disc axial and radial deformations δ_u and δ_h in terms of nucleus pulposus pressure *p*, the annulus modulus *E* and the disc dimensions, as given by equations (17) and (18). Now the annulus modulus E is a function of the stress in the annulus (k being the constitutive proportionality constant), and hence of the pressure p and the disc dimensions, as shown by equation (21). As a result of this relationship, we finally show that the disc deformations are only functions of k and the disc dimensions. This implies that irrespective of the increase in the value of load P, the disc deformations remain constant, and only depend on the constitutive property parameter k. This is the novelty of the intrinsic design of the spinal disc!

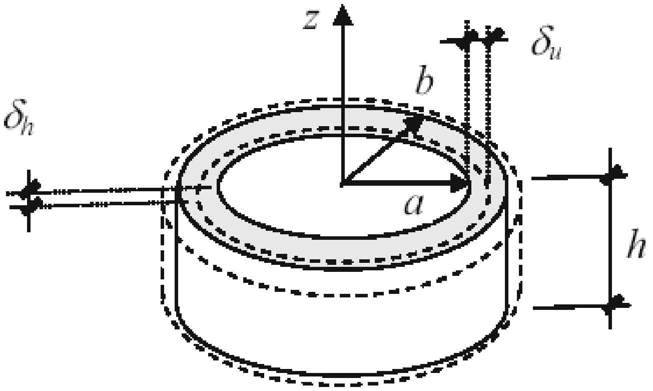


Fig. 9. Geometry and Deformation Variables of the Spinal Disc [4].

ANALYSIS

The equilibrium equation is:

$$\sigma_r - \sigma_\theta + r\frac{d\theta_r}{dr} = 0 \tag{1}$$

Strain-displacement relations: $$\in_r = \frac{du}{dr}, \in_\theta = \frac{u}{r}, \tag{2}$$

Constitutive relations:

$$\sigma_r = E\in_r = E\frac{du}{dr}\,,\; \sigma_\theta = E\in_\theta = E\frac{u}{r} \tag{3}$$

Substituting eqn (3) into eqn (1), we have:

$$\frac{d^2u}{dr^2} + \frac{1}{r}\frac{du}{dr} - \frac{u}{r^2} = 0\,,\; u = Ar + \frac{B}{r} \tag{4}$$

Because of the incompressible nucleus pulposus fluid inside:

$$\pi a^2 h = \pi(a+\delta_u)^2(h-\delta_h) \tag{5}$$

the (radial and axial) deformations δ_u and δ_h are related as:

$$\delta_u = \frac{\pi a^2 \delta_h}{2\pi a h} = \left(\frac{a}{2h}\right)\delta_h \tag{6}$$

Now, we designate:

$$u\big|_{r=a} = \delta_u, \sigma_r\big|_{r=b} = 0 \tag{7}$$

Substituting eqn (7) into eqns (3 & 4), we get

$$\begin{aligned} & Aa + \frac{B}{a} = \delta_u \\ & E\left[A - \frac{B}{b^2}\right] = 0 \end{aligned} \tag{8}$$

So that

$$A = \frac{a\delta_u}{a^2 + b^2},\ B = \frac{a\delta_u.b^2}{a^2 + b^2} \text{ in the 'u' function} \tag{9}$$

Then, substituting A and B into eqns (4 & 3), we obtain:

- for the radial stress, $\sigma r = \dfrac{Ea\left(r^2 - b^2\right)\delta u}{(a^2 + b^2)r^2} = \dfrac{a^2\left(r^2 - b^2\right)}{2h(a^2 + b^2)r^2}E\delta h$ (10-a)

- for the circumferential stress, $\sigma\theta = \dfrac{Ea\left(r^2 + b^2\right)\delta u}{(a^2 + b^2)r} = \dfrac{a^2\left(r^2 + b^2\right)}{2h(a^2 + b^2)r}E\delta h$ (10-b)

- for the axial stress $\sigma_z = E \in_z = E\dfrac{\delta_h}{h}$ $\delta u = \left(\dfrac{a}{2h}\right)\delta h$ (10-c)

Once σ_z is evaluated, δ_h will become known (from eqn. 10-c) and subsequently σ_r, σ_θ (from eqns 10-a & 10-b) and δ_u(from eqn 10-c).

Stress Analysis for Vertical Loading [4]:

For a vertically applied force *P*,

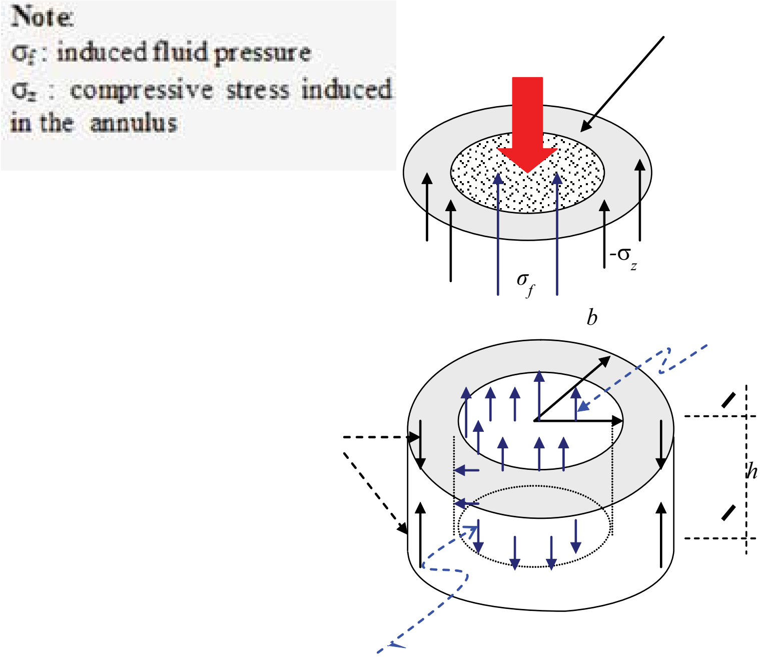
$$P = \left(\pi a^2\right)\sigma_f + \pi(b^2 - a^2)\sigma_z \tag{11}$$

where σ_f is the hydrostatic pressure in the NP fluid, and σ_z the axial stresses in the annulus. Because the disc height (*h*) is small, $\sigma_f \approx$ *constant, and hence*

$$\sigma_f = -\sigma_r\big|_{r=a} = p \tag{12}$$

Then, based on eqns (12, 10-a & 10-c), we obtain

$$\sigma_f = -\frac{\left(a^2 - b^2\right)}{2h\left(a^2 + b^2\right)}E\delta_h = \frac{\left(b^2 - a^2\right)}{2h\left(a^2 + b^2\right)}E\left(\frac{\delta h}{h}\right), \text{ or } \sigma_f = \frac{(b^2 - a^2)}{2(a^2 + b^2)}\sigma_z = p \tag{13}$$


Note:
σf : induced fluid pressure
σz : compressive stress induced in the annulus
-σz
σf
b
h

Normal stresses σ_f & σ_z equilibrating the applied force P

Fig. 10. Induced Stresses in the disc annulus and Pressure *p* in the nucleus-pulposus in response to the load *P* [4].

Substituting

$$\sigma_f = \frac{\left(b^2 - a^2\right)}{2\left(a^2 + b^2\right)}, \quad \sigma_z = p\,,\ \sigma_z = E\frac{\delta h}{h} = 2E\frac{\delta u}{a}$$

Into

$$P = (\pi a^2)\sigma_f + \pi(b^2 - a^2)\sigma_z$$

We get:

$$\delta_h = \frac{h\sigma_z}{E} = \frac{h}{E}\ \frac{P}{\pi(b^2 - a^2)}\left[\frac{2(a^2 + b^2)}{3a^2 + 2b^2}\right] \tag{14}$$

$$\delta_u = \frac{a}{2h}\delta_h = \frac{a}{2E}\ \frac{P}{\pi(b^2 - a^2)}\left[\frac{2(a^2 + b^2)}{3a^2 + 2b^2}\right] \tag{15}$$

Nucleus-pulposus fluid pressure,

$$p = \frac{P}{\pi(3a^2 + 2b^2)} \tag{16}$$

$$\sigma_z = \frac{2p(a^2 + b^2)}{(b^2 - a^2)} = \frac{2P(a^2 + b^2)}{\pi(b^2 - a^2)(3a^2 + 2b^2)} \tag{16-a}$$

$$\delta_h = \frac{h\sigma_z}{E} = \frac{h}{E} \quad \frac{P}{\pi(b^2 - a^2)} \left[\frac{2(a^2 + b^2)}{3a^2 + 2b^2} \right]$$

Stress in the Annulus

From,
$$\sigma\theta = \frac{Ea\left(r^2 + b^2\right)\delta u}{(a^2 + b^2)r} = \frac{a^2\left(r^2 + b^2\right)}{2h(a^2 + b^2)r} E\delta h$$

We get :
$$\frac{a^2\left(r^2 + b^2\right)P}{\pi r^2(b^2 - a^2)(3a^2 + 2b^2)} = \frac{pa^2\left(r^2 + b^2\right)}{r^2(b^2 - a^2)}, \tag{17}$$

$$\sigma_\theta\left(\mathrm{r} = \mathrm{a}\right) = \frac{p\left(a^2 + b^2\right)}{(b^2 - a^2)}, \text{ in the annulus} \tag{18}$$

wherein, $p = \frac{P}{\pi\left(3a^2 + 2b^2\right)}$, pressure in nucleus-pulposus fluid

The disc deformations have been obtained as:

Axial deformation,
$$\delta_h = \frac{2ph(a^2 + b^2)}{E_c(b^2 - a^2)} \tag{19}$$

Radial deformation
$$\delta_u = \frac{pa(a^2 + b^2)}{E_c(b^2 - a^2)} \tag{20}$$

wherein $E_c = E - E_o = k\sigma$.

$$\therefore\ E_c = k\sigma_\theta(r = a) = kp\frac{(a^2 + b^2)}{(b^2 - a^2)} \tag{21}$$

Now, as the magnitude of the load *P* increases, the pressure *p* in nucleus-pulposus fluid also increases. Then, as p increases, so does the modulus E_c according to eqn (21)

$$\therefore\ \frac{p}{E_c} = \frac{\left(b^2 - a^2\right)}{k\left(a^2 + b^2\right)} = \text{constant} \tag{23}$$

$$\delta_h = \frac{2ph(a^2 + b^2)}{E_c(b^2 - a^2)} = \frac{2h}{k}\text{, a constant, and } \delta_u = \frac{pa(a^2 + b^2)}{E_c(b^2 - a^2)} = \frac{2a}{k}\text{, a constant}$$

This means that, irrespective of increase in the value of load P, the disc deformations δ_u and δ_h remain constant, and only depend on the constitutive property parameter k. This is the novelty of the intrinsic design of the spinal disc!

3. Physiology: Mechanism of left ventricle twisting and pressure increase during isovolumic contraction (due to the contraction of the myocardial fibres)

Introduction and objective: The left ventricular (LV) myocardial wall is made up of helically oriented fibers. As the bioelectrical wave propagates along these fibers, it causes concomitant contraction wave propagation. Our LV cylindrical model is illustrated in figure 11. The contraction of the helical oriented myocardial fibers causes active twisting and compression of the LV (as illustrated in fig 11), thereby compressing the blood fluid contained in it. Then due to the very high bulk modulus of blood, this fluid compression results in substantial pressure increase in the LV cavity.

Herein we simulate this phenomenon of LV isovolumic contraction, which causes the intra-LV pressure to rise so rapidly during 0.02-0.06 seconds of isovolumic contraction. Our objective is to determine how the pressure generated during isovolumic contraction, due to by active torsion (with LV twisting) and compression (with LV shortening) caused by the contractile stress in the helically wound myocardial fibers [5].

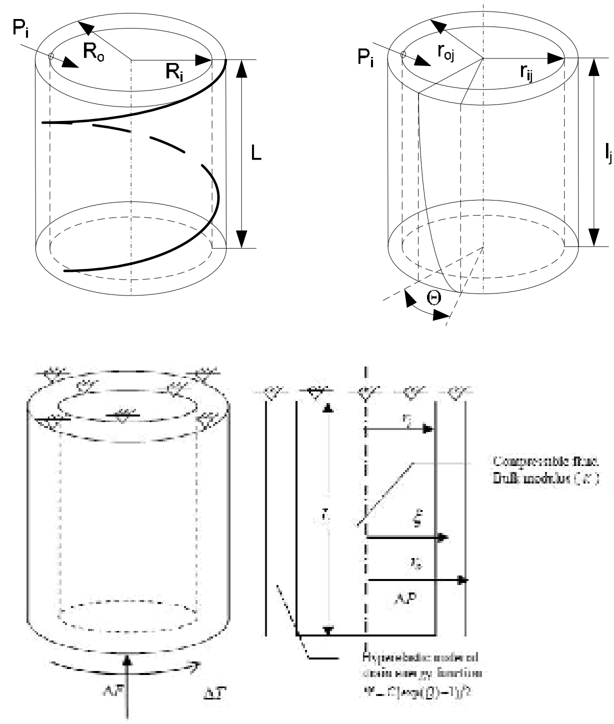


Fig. 11. Top: Fiber orientation and twisting model of the left ventricle (LV). Bottom: The fluid-filled LV cylindrical shell model: (i) geometry (ii) material property, and (iii) equivalent compression Δ*F* and Δ*T* associated with its internal stress state due to internal pressure rise within it.

Concept: In order to simulate the left ventricle deformation during isovolumic contraction, we have modeled it as a pressurized fluid-filled thick-walled cylindrical shell supported by the aorta along its upper edge. The LV cylindrical model consists of an incompressible hyperelastic material with an exponential form of the strain energy function ψ.

The contraction of the myocardial fibers causes active twisting and compression of the left ventricle, thereby compressing the blood fluid contained in it. Then, due to the very high bulk modulus of blood, this compression results in pressure increase in the ventricular cavity. Hence, we simulate this phenomenon by applying and determining equivalent active torque and compression to the LV-cylindrical model incrementally (as ΔF and ΔT), so as to raise the LV pressure by the monitored amounts.

Modeling approach:

We monitor LV pressure (p), LV volume (V), myocardial volume (V_M), and wall thickness (h) at time intervals during isovolumic contraction. From the monitored V, V_M, and h, we determine the LV model radius R and length L and the wall thickness h. We also monitor LV twist angle ϕ.

We then invoke blood compressibility to determine ΔV at subsequent instants , as $\Delta V = (\Delta p/K)\ V$, where Δp is the monitored incremental pressure and K is the bulk modulus of blood. From the volume strain $\Delta V/V$, we then determine the model length and radius strains $\Delta I/L$ and $\Delta r/R$, and hence the LV dimensions with respect to LV dimensions at the start of isovolumic contraction. This enables us to determine the stretches (strains) (λ_z, λ_r, λ_θ and γ), and thereafter the Lagrange strain tensor components (E_{rr}, $E_{\theta\theta}$, E_{zz}, $E_{\theta z}$) in terms of these stretches and the hydrostatic pressure.

Then, we express the LV wall stresses in terms of the strain energy density function Ψ, of the Lagrange strain tensor components in cylindrical coordinates and the material constitutive parameters (b_i), in which (i) the stretches (λ_z, λ_r, λ_θ and γ) have been calculated and are known, (ii) the hydrostatic pressure and the constitutive parameters b_i ($i = 1, 2, \ldots, 9$) are the unknowns.

So now, we substitute the stress expressions σ_{rr} and $\sigma_{\theta\theta}$ into the boundary conditions equations (of equilibrium between the internal pressure and the wall stresses σ_{rr} and $\sigma_{\theta\theta}$, and between the internal pressure and wall stress σ_{zz}), and determine the best values of the constitutive parameters (b_i) and the hydrostatic pressure to satisfy these equations.

We then go back, and determine the stress expressions. We utilize the stress expressions for σ_{zz} and $\sigma_{\theta z}$, to determine the generated values of torsion (ΔT) and axial compression (ΔF),due to the contraction of the helically wound myocardial fibres.

Finally, we determine the principal stresses and principal angle along the radial coordinate of the LV wall thickness, from which we can interpret the fibre orientations, which can be related to the LV contractility index.

This procedure is carried out at three instants of time from the start of isovolumic contraction, and at 0.02 s, 0.04 s, 0.06 s into the isovolumic contraction phase. Hence, from the monitored LV Δp and computed ΔV at these three instants (with respect to the pressure and volume at $t = 0$ at the start of isovolumic contraction phase, we determine (i) the time variation of the internally generated torque and axial compression during the isovolumic contraction phase (fig 12), as well as (ii) the time variations of the principal (tensile) stress and the principal angle (taken to be equivalent to the fiber angle) during the isovolumic contraction phase (fig 13).

Model Kinematics:

We model the LV as an incompressible thick-walled cylindrical shell subject to active torsion torque and compression as illustrated in fig 11. The upper end of the LV model is constrained in the long-axial direction to represent the suspension of the left ventricle by the aorta at the base. Now, considering the LV at end-diastole to be in the unloaded reference configuration, the cylindrical model in its undeformed state is represented geometrically in terms of cylindrical coordinates (R, Θ, Z) by

$$R_i \le R \le R_o,\ \ 0 \le \Theta \le 2\pi,\ \ 0 \le Z \le L \tag{1}$$

where R_i, R_o and L denote the inner and outer radii, and the length of the undeformed cylinder, respectively.

In terms of cylindrical polar coordinates (r, θ, z), the geometry of the deformed LV configuration (with respect to its undeformed state at the previous instant) is given by:

$$r_i \le r \le r_o,\ \ 0 \le \theta \le 2\pi,\ \ 0 \le z \le l \tag{2}$$

where r_i, r_o and l denote the inner and outer radii, and the length of the deformed cylinder, respectively.

We further consider the incompressible LV model in its reference state to be subjected to twisting, radial and axial deformations in the radial and long-axis directions during isovolumic contraction, such that (also based on incompressibility criterion), the deformations of incompressible LV cylindrical shell can be expressed as

$$r = \sqrt{\frac{R^2 - R_i^2}{\lambda_z} + r_i^2},\ \theta = \Theta + Z\frac{\phi}{L},\ z = \lambda_z Z \tag{3}$$

where λ_z is the constant axial stretch, r_i is the inner radius in the deformed configuration and ϕ is the measured angle of twist at the apex of the LV (relative to the base). It can be seen that the twist angle (θ) and the axial deformation (z) are zero at the upper end of the LV.

Model Dimensions:

At any instant (t), the geometrical parameters (or dimensions) of the LV cylindrical model (instantaneous radius R and length L, as defined in fig 11) can be determined in terms of the monitored LV volume (V), myocardial volume (V_M) and wall thickness (h), as follows:

$$R_i = \frac{2Vh/V_M + \sqrt{(2Vh/V_M)^2 + 4Vh^2/V_M}}{2} \tag{4-a}$$

$$L = V/\pi R_i^2 \tag{4-b}$$

Then, $$R_o = R_i + h. \tag{4-c}$$

These equations will be employed to determine the LV dimensions at the start of isovolumic contraction phase (t = 0). The determination of the dimensions of the deformed LV (due to contraction of the myocardial fibers) at the subsequent instants of the isovolumic contraction phase is indicated in the next subsection. We also utilize the information on the LV twist angle (ϕ) during the isovolumic phase, from MRI myocardial tagging. From this information, we can determine the stretches (λ_z, λ_r, λ_θ and γ).

Theoretical Analysis:

The strain energy density function suitable for the myocardium material, is given by:

$$\psi = C(\exp(Q) - 1)/2 \tag{5}$$

wherein Q is a quadratic function of the 3 principal strain-components (to describe 3-d transverse isotropy) in the cylindrical coordinate system, given by:

$$\begin{aligned} Q = {} & b_1 E_{\theta\theta}^2 + b_2 E_{ZZ}^2 + b_3 E_{RR}^2 + 2b_4 E_{\theta\theta} E_{ZZ} + 2b_5 E_{RR} E_{ZZ} \\ & + 2b_6 E_{\theta\theta} E_{RR} + 2b_7 E_{\theta Z}^2 + 2b_8 E_{RZ}^2 + 2b_9 E_{\theta R}^2 \end{aligned} \tag{6}$$

wherein b_i are non-dimensional material parameters, and E_{ij} are the components of the modified Green-Lagrange strain tensor in cylindrical coordinates (R, Θ, Z). In order to reduce the mathematical complexity of the problem, we assume negligible shear during isovolumic contraction. Thus E_{RZ} and $E_{\theta R}$ in equation (6) and their corresponding stress components (σ_{RZ} and $\sigma_{\theta R}$) are neglected.

The stress equilibrium equation (in the cylindrical coordinate system) is given by the following equation:

$$\frac{d\sigma_{rr}}{dr} + \frac{(\sigma_{rr} - \sigma_{\theta\theta})}{r} = 0 \tag{7}$$

The boundary conditions on the outer and inner surfaces of the LV cylindrical model are given by

$$\sigma_{rr}(r = r_o) = 0 \text{ , } \sigma_{rr}(r = r_i) = -p \tag{8}$$

where p is the LV pressure acting on the inner surface of the LV model; we employ incremental pressure Δp with respect to the LV pressure at $t = 0$, the start of isovolumic contraction.

By integrating eq (7), the Cauchy radial stress σ_{rr} is given by:

$$\sigma_{rr}(\xi) = \int_{\xi}^{r_o} (\sigma_{rr} - \sigma_{\theta\theta}) \frac{dr}{r}, \qquad r_i \le \xi \le r_o \tag{9}$$

There from, the boundary condition eq (8) of the internal pressure $p = -\sigma_{rr}\ (r=r_i)$ is obtained (by substituting equation 9 into the boundary condition 8), in the form:

$$p_i = -\int_{r_i}^{r_o} (\sigma_{rr} - \sigma_{\theta\theta}) \frac{dr}{r}, \tag{10}$$

Since the valves are closed during isovolumic contraction, we impose another set of boundary conditions (at both the top and bottom of the internal LV surface), giving:

$$\sigma_{zz}\pi\left(r_o{}^2 - r_i{}^2\right) = p\left(\pi r_i{}^2\right) \tag{11}$$

where σ_{ZZ} denotes the axial component of the Cauchy stresses. In the analysis, we will employ Δp with respect to the pressure at $t = 0$, at the start of isovolumic contraction.

The blood in the left cavity is assumed compressible, and the change in cavity volume (ΔV) due to the monitored incremental pressure (Δp), is given by

$$\Delta V = \Delta p / K; \qquad K = 2.0 \times 10^9 pa \tag{12}$$

where K is the bulk modulus of blood.

Analysis and computational procedure:

The following analysis is carried out at the three time instants t (or j) = 0.02 secs, 0.04 secs and 0.06 secs (from the start $t = 0$ of the isovolumic contraction phase, from the monitored Δp and computed ΔV (eq 12) at the three time instants with respect to p and V at t = 0 (the start of isovolumic contraction phase).

1. We obtain the increments in LV pressure (Δp) from the monitored pressure data, during isovolumic contraction. By taking $K=\Delta p/\Delta V$, we compute the corresponding changes in LV volume ΔV.
2. Left ventricular deformation: For this change ΔV in the LV volume, by assuming the ratios of cylindrical-model length and radius strains $\Delta l/L$ and $\Delta r/R$ to be equal during isovolumic contraction, we obtain their expressions as:

$$\Delta l_j = \left(1 - \sqrt[3]{1 - \Delta p/K}\right)L, \;\; \Delta r_{ij} = \left(1 - \sqrt[3]{1 - \Delta p/K}\right)R \tag{13}$$

From equation (13), the incremental quantities Δl_j and Δr_{ij} can be calculated, and hence:

$$l_j = L - \Delta l_j, \text{ and } r_{ij} = R - \Delta r_{ij} \tag{14}$$

where l_j and r_{ij} are the deformed model length and radius at time t (or j)
So the wall-thickness h can be obtained from:

$$h_j = \sqrt{\frac{V_M / l_j + \pi r_{ij}^{\,2}}{\pi}} - r_{ij} \tag{15}$$

Let $\Delta\phi$ denote the relative angle of twist measured at the apex, at each of the 3 stages of isovolumic contraction phase, obtained by magnetic resonance imaging (MRI).

3. Next we determine the stretches in the 3 directions (due to deformed dimensions l and r with respect to the undeformed dimensions L and R) as follows

$$\lambda_z(r) = \frac{1}{L}, \;\; \lambda_r(r) = \frac{\partial r}{\partial R} = \frac{R}{r\lambda_z}, \;\; \lambda_\theta(r) = \frac{r\partial\theta}{R\partial\Theta} = \frac{r}{R} \tag{16}$$

We define the twist stretch due to torsion, as

$$\gamma(r) = \frac{r\partial\theta}{\partial z} = \frac{r\phi}{l} \tag{17}$$

wherein ϕ is zero at the top surface of the LV held by the aorta

4. Next we express the components of the Lagrange Green's strain tensor in terms of the stretches and deformations obtained from equations (8-10), as:

$$\mathrm{E}_{rr}=\frac{1}{2}\left(\lambda_r^2\text{-}1\right),\ \mathrm{E}_{\theta\theta}=\frac{1}{2}\left(\lambda_\theta^2\text{-}1\right),\ \mathrm{E}_{zz}=\frac{1}{2}\left(\lambda_z^2\left(1+\gamma^2\right)\text{-}1\right),\ \mathrm{E}_{\theta z}=\frac{\gamma\lambda_z\lambda_\theta}{2} \tag{18}$$

5. Then, by using the strain energy function (eqs 5 and 6), we obtain the expressions of Cauchy stress tensor in terms of the parameters b_i:

$$\sigma_{\theta\theta}=\lambda_\theta^2\frac{\partial\Psi}{\partial\mathrm{E}_{\theta\theta}}+2\gamma\lambda_z\lambda_\theta\frac{\partial\Psi}{\partial\mathrm{E}_{\theta z}}+\gamma^2\lambda_z^2\frac{\partial\Psi}{\partial\mathrm{E}_{zz}}-\overline{p}\qquad\sigma_{rr}=\lambda_r^2\frac{\partial\Psi}{\partial\mathrm{E}_{rr}}-\overline{p}$$

$$\sigma_{zz}=\lambda_z^2\frac{\partial\Psi}{\partial\mathrm{E}_{zz}}-\overline{p}\qquad\sigma_{\theta z}=\lambda_z\lambda_\theta\frac{\partial\Psi}{\partial\mathrm{E}_{\theta z}}+\gamma\lambda_z^2\frac{\partial\Psi}{\partial\mathrm{E}_{zz}} \tag{19}$$

wherein ψ is given by equations (5 & 6), and the to-be-determined unknown parameters are the hydrostatic pressure $\overline{p}$ and the constitutive parameters b_i (in equation 6).

6. We now employ the stress expressions into the boundary conditions (10 & 11), and determine the hydrostatic pressure $\overline{p}$ and the parameters b_i (i=1,2,…7) and, using a nonlinear least squares algorithm.
7. We next determine (i) the stretches from eqs (16 & 17), and (ii) therefrom, the strains $E_{i,j}$ from eq (18). Then by substituting the computed strains components $E_{i,j}$ and the strain energy density function ψ (eqs 5 & 6) into eq (19), we can determine the LV wall stress components for the LV cylindrical model.
8. From the stress components $\sigma_{\theta z}$, $\sigma_{\theta\theta}$, and σ_{zz}, we can computes the principal stresses σ_1, σ_2 the principal angle ϕ.

$$\sigma_{1,2}=\frac{\sigma_{zz}+\sigma_{\theta\theta}}{2}\pm\sqrt{\left(\frac{\sigma_{zz}-\sigma_{\theta\theta}}{2}\right)^2+\sigma_{\theta z}{}^2} \tag{20}$$

$$\tan 2\phi=\frac{2\sigma_{\theta z}}{\sigma_{\theta\theta}-\sigma_{zz}} \tag{21}$$

9. Then, we compute the equivalent active axial force ΔF and the torsional couple ΔT as follows:

$$\Delta F=2\pi\int_{r_i}^{r_o}\sigma_{zz}r\mathrm{d}r,\quad M_T=2\pi\int_{r_i}^{r_o}\sigma_{\theta z}r^2\mathrm{d}r, \tag{22}$$

Results and comments:

1. The computational model results for a sample subject are provided in Tables 1 -5.
2. The variations of the equivalent active torque T and axial compressive force F during the isovolumic phase are calculated and shown in Fig. 12.

We hence demonstrate that the big increment of internal pressure in LV cavity is caused by the compression of the blood (due to its high bulk modulus) by the internally generated torque and axial force during the isovolumic phase, due to the contraction of the myocardial fibers.

3. The radial variations of the principal stress and principal angle are found to be almost uniform across the wall thickness [5]. The time variation of the principal stress and angle during isovolumic phase are shown in **Fig. 13**. At the end of the isovolumic phase, the magnitude of the tensile principal stress is around 1.75x10^{5}pa, which is in good agreement with the isometric tension value of 1.40x10^{5}pa achieved under maximal activation.

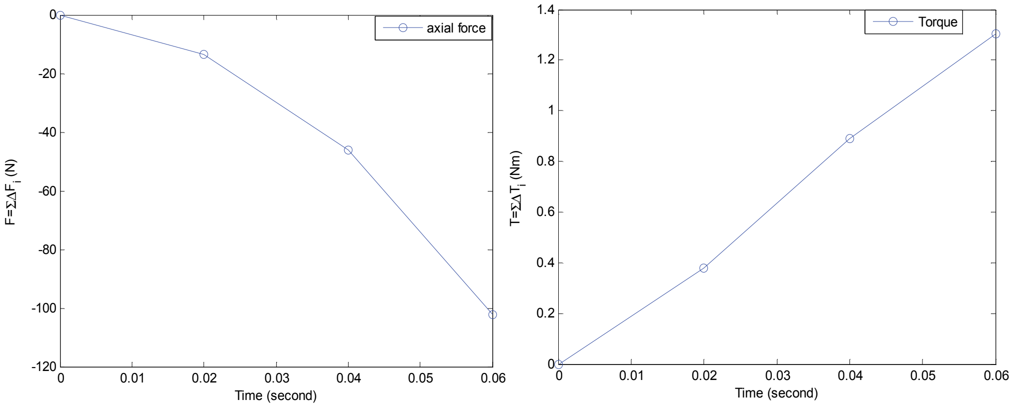


Fig. 12. Variations of axial force and torque as function of time during isovolumic phase [5].

The notable result from **Fig. 13** is that both the principle stresses and their orientation angle keep changing during the isovolumic phase. It is seen from Figure 13, that the myocardial fiber was oriented 38^0 at the start and became 33^0 at the end of the isovolumic phase, as the monitored internal pressure increased during isovolumic phase (due to active torsion and compression, corresponding to active contraction of the helically woven fibers from 38^0 to 33^0) from 24 to 44 to 62 mmHg during the isovolumic contraction phase. Observing the values of the principle stresses, it is seen that the LV is acted upon by internally generated (i) active torsion (T) causing its twisting, and (ii) compression (F) to cause compression of the blood and develop the LV pressure rise.

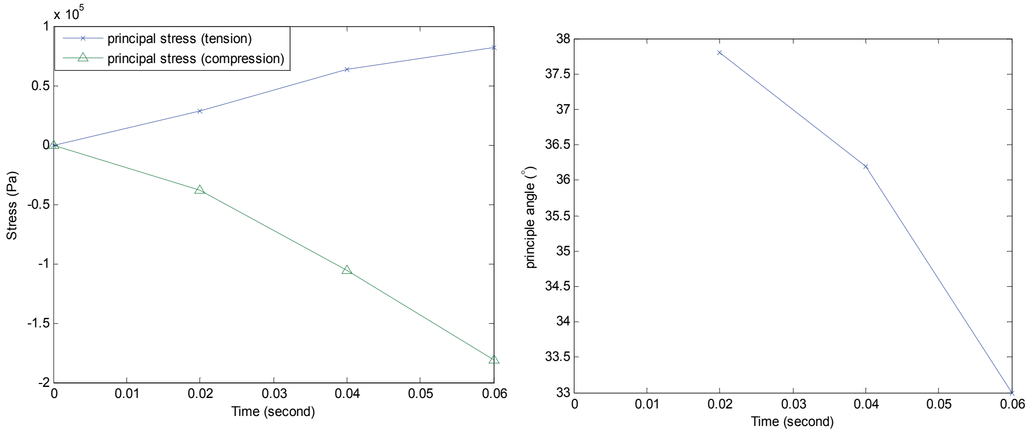


Fig. 13.Variation of principal stresses as functions of time, during isovolumic phase [5].

In conclusion: we have indirectly shown that the contraction of the myocardial fibers (i) develops active stresses in the LV wall (and corresponding principal stresses), and (ii) causes LV twist and shortening, resulting in the development of substantial pressure increase during isovolumic contraction.

The tensile principal stress corresponds to the active contractile stress generated in the myocardial fibers, while the angle of the tensile principal stress corresponds to the fiber helical angle, which is in agreement with the experiment data on the fiber angle.

The computational results are shown in Tables 1 to 5, for a sample subject.

t (second)	P (mmHg)	ΔP (mmHg)	V (ml)	ΔV (ml)	r_{ij} (cm)	Δr_{ij} (cm)	l_j (cm)	Δl_j (cm)	h_j (cm)	r_{oj} (cm)	ΔΦ (°)
0	18		1.36700000E+02		2.03208400E+00 (=R_i)		1.053745000E+01 (=L)		1.085247000E+00	3.117331000E+00 (R_o)	0
0.02	43	25	1.36699773E+02	2.27263750E-04	2.032083992E+00	8.46701669E-09	1.053744996E+01	4.39060418E-08	1.085246679E+00	3.117330670E+00	0.667
0.04	63	45	1.36699591E+02	4.09074750E-04	2.032083985E+00	1.52406301E-08	1.053744992E+01	7.90308757E-08	1.085246684E+00	3.117330669E+00	1.333
0.06	81	63	1.36699427E+02	5.72704650E-04	2.032083979E+00	2.13368822E-08	1.053744989E+01	1.10643226E-07	1.085246689E+00	3.117330667E+00	2.00

Table 1. Pressure-volume and model parameters for a sample subject with V_M=185ml.

parameter	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
value	5946.2278	15690.58158	422.514993	16157.10454	16360.53744	33299.28998	680.7385218	0	0

Table 2. The parameters of the strain energy function for the sample case shown in Table 1.

t	λ_z	λ_r	λ_θ	Υ
0				
0.02	0.9999999958	1.000000008E+00	9.99999996E-01	1.28626948E-01
0.04	0.9999999925	1.000000015E+00	9.99999993E-01	2.56482520E-01
0.06	0.9999999895	1.000000021E+00	9.99999990E-01	3.85688000E-01

Table 3. The stretches calculated from equations (16 & 17) for the sample case shown in Table 1.

t	Endocardium--Epicardium									
(second)	0	1.99E-02	1.02E-01	2.37E-01	4.08E-01	0.591715	0.762765	0.89833	0.98014	1
0.02	-24.0809	-23.4615	-17.2801	-12.9839	-8.73378	-4.9888	-2.11726	-0.41173	-0.2144	0
0.04	-44.2917	-43.1341	-31.6543	-23.7482	-15.9694	-9.13029	-3.88098	-0.75569	-0.3542	0
0.06	-6.21E+01	-6.05E+01	-5.43E+01	-4.48E+01	-3.39E+01	-2.31E+01	-1.34E+01	-5.75E+00	-1.13E+00	0

Table 4. Radial stresses distributions along LV wall from endocardium to epicardium.

t (second)	Results								
	Circumferential stress (Pa)	Radial stress (Pa)	Axial force (N)	Torque (Nm)	Axis stress (Pa)	Shear stress (Pa)	Principal stress-tension (Pa)	Principal stress-compression (Pa)	Principal direction (°)
0	0.00	0.00	0.00	0.00	0.00E+00	0.00E+00	0.00	0.00	
0.02	3358.33	-2844.53	-13.51	0.38	-1.35E +04	3.29E+04	28221.68	-38395.38	37.82
0.04	5318.61	-5956.38	-46.06	0.89	-4.70E +04	8.21E+04	63836.28	-105547.96	36.18
0.06	4481.83	-8530.94	-102.38	1.30	-1.04E +05	1.22E+05	81673.18	-180897.35	33.02

Table 5. Results for the sample subject at different instants

4. Clinical evaluation of Physiological systems in terms of Non-dimensional physiological Indices

Non-dimensional numbers (made up of several phenomenon related terms) are employed to characterise – disturbance phenomena. For example, in a cardiovascular fluid-flow regime, the Reynold'snumber

$$N_{re} = \rho VD / \mu \tag{1}$$

(*V :* flow velocity, *D:* diameter, μ: fluid viscosity, ρ: fluid density) characterizes turbulent flow, which can occur in the ascending aorta when the aortic valve is stenotic (giving rise to murmurs) and accentuated in the case of anaemia (decreased blood viscosity).

Integration of a number of isolated but related events into one non-dimensional physiological index (NDPI) can help to characterise an abnormal state of a particular physiological system [1].

For utilization of an NDPI diagnostically, we evaluate it in a large patient population and develop its distribution into normal and dysfunctional ranges, as illustrated in figure 14. Then, upon evaluation of the NDPI of a particular subject, we can see if that value falls in the normal or dysfunctional range, and accordingly make the medical diagnosis.

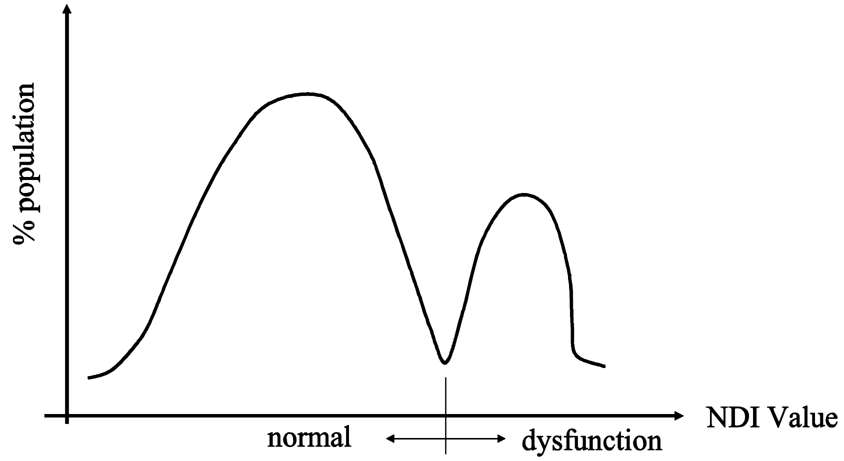


Fig. 14. Illustration of the Distribution of an NDI, showing its normal and dysfunctional ranges which can be employed for diagnostic purpose.

5. Medical test: Cardiac fitness index based on treadmill HR variation

In this procedure, the cardiac fitness model consists of a first-order differential-equation system model, describing the heart rate (HR) response (y) to exertion (exercise, jogging etc) monitored in terms of the work-load, where y is defined as follows:

$$y = \frac{HR(t) - HR(rest)}{HR(rest)} \tag{1}$$

The subject is exercised on the treadmill for a period of time *te* (minutes) at a constant work load (W), while the HR(t) (and hence y) is monitored, as displayed in fig 15. Now we develop a model to simulate (i) the y(t) response during exercise, i.e. during $t \le te$, and (ii) thereafter for y(t) decay, after the termination of exercise. In a way, *te* represents the exercise endurance of the subject [6].

For a person, the model equation for HR response is represented by:

$$\frac{dy}{dt} + k_1 y = C_0 W \tag{2}$$

For $t \le t_e$, the solution is given by:

$$y = \frac{y_e(1-e^{-k_1 t})}{(1-e^{-k_1 t_e})} \tag{3}$$

For the recovery period *(t ≥ te),* the solution of eqn. (2) is :

$$y = y_e e^{-k_2(t-t_e)} \tag{4}$$

where k_1 and k_2, are the model parameters, which can serve as cardiac-fitness parameters (in min^{-1}).

Non-dimensional Cardiac Fitness Index:

A typical y(t) response is illustrated in Fig 15.

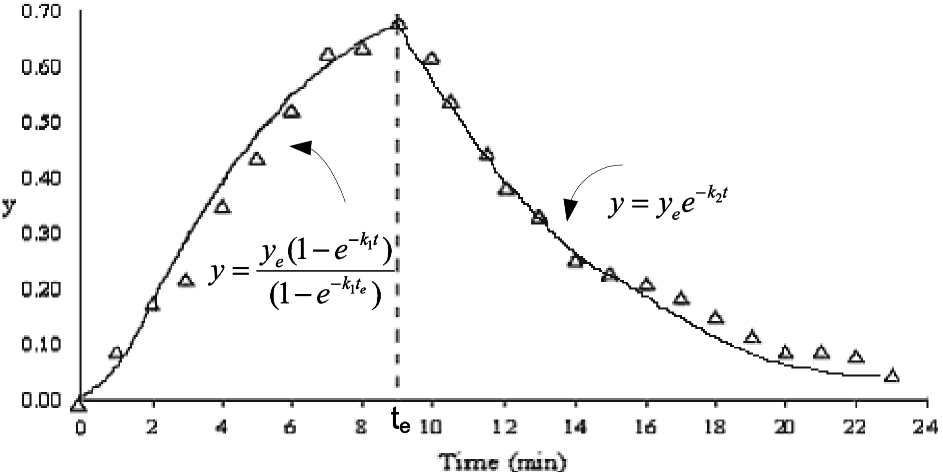


Fig. 15. Graph of y (the computed HR response) vs. t during treadmill exercise *t=te* and during the recover of period *t=te.*

The parameters k_1 and k_2 can be continued into a single nondimensional cardiac-fitness index (CFI):

$$CFI = k_1 k_2 t_e^{\,2} \tag{5}$$

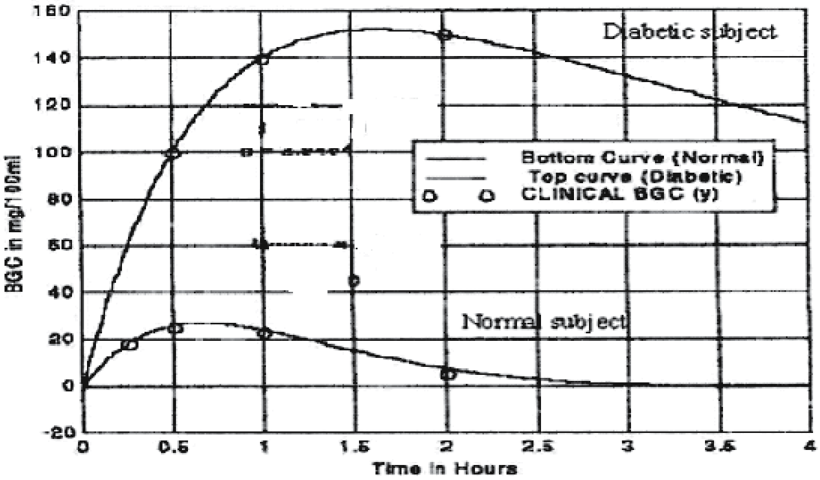
According to this formulation of CFI, a healthier subject has (i) greater k_1 (i.e., slower rate-of-increase of HR during exercise (ii) greater k_2 (i.e., faster rate-of-decrease of HR after exercise) (iii) greater t_e (i.e., exercise endurance), and hence (iv) higher value of CFI.

Subject	Classification	CFI (relative values)
1	Athletic	271
2	Occasionally runs	90
3	Rarely exercise	65
4	Sedentary	19
5	Exercises regularly	155

Table 1. CFI values for athletic, fit and unfit subjects.

Now, we need to evaluate CFI for a big spectrum of patients, and then compute its distribution curve, to determine the efficiency of this index, in order to yield distinct separation of CFI ranges for healthy subjects and unfit patients. This CFI can also be employed to assess improvement in cardiac fitness following cardiac rehabilitation regime. This CFI is non-dimensional, and it can be useful to clinicians as they are able to predict the heart condition or fitness performance of a person by referring to the value of a single index value.

6. Medical physiology: A non-dimensional diabetes index with respect to Oral-Glucose-Tolerance testing



$$y(t) = \frac{G}{\omega} e^{-At} \sinh \omega t$$
$$(AT_d = 4.9)$$
$$\lambda = 0.27 hr^2$$
$$T_d = 6.08 hr$$
$$G = 2.84 gmlit^{-1} hr^{-1}$$
$$A = 0.808 hr^{-1}$$

$$y(t) = \frac{G}{\omega} e^{-At} \sin \omega t$$
$$(AT_d = 1.4)$$
$$\lambda = 2.6 hr^2$$
$$T_d = 1.08 hr$$
$$G = 1.04 gmlit^{-1} hr^{-1}$$
$$A = 1.3 hr^{-1}$$

Fig. 16. OGTT response-Curve: A=1.3hr^{-1} (i.e., higher damping.coefficient value) for the normal subject for the diabetic patient A=0.808 hr^{-1}, i.e., the damping coefficient is smaller [7, 8].

Oral Glucose Tolerance Test (OGTT) is a standard procedure for diagnosing diabetes and risk of becoming diabetic. However, the test data is assessed empirically. So, in order to make this procedure more reliable, we have carried out a biomedical engineering analysis of the OGTT data, to show how to distinguish diabetes subjects and those at risk of becoming diabetic.
For Oral-glucose Tolerance Test simulation (entailing digestive & blood-pool chambers), the differential equation is as follows [7,8]:

$$\begin{aligned} & y'' + 2Ay' + \omega_n^2 y = G\delta(t); \quad \text{y in gms/liter, G in gm/liter/hr} \\ & or, \\ & y'' + \lambda T_d y' + \lambda y = G\delta(t) \end{aligned} \tag{1}$$

wherein $\left(\omega_n = \lambda^{1/2}\right)$ is the natural oscillation-frequency of the system, A is the attenuation or damping constant of the system, ω = $(\omega_n^2 - A^2)^{1/2}$ is the (angular) frequency of damped oscillation of the system, λ= $2A/T_d$=, vω_n^2 with (λy) representing the proportional-control term of blood-glucose concentration (y).

($\lambda T_d y'$)is the derivative feedback control term with derivative-time of T_d $G\delta(t)$ represents the injected glucose bolus.

The input to this system is taken to be the impulse function due to the orally ingested glucose bolus [G], while the output of the model is the blood-glucose concentration response y(t). For an impulse glucose-input, a normal patient's blood-glucose concentration data is depicted in **Figure 16** by open circles. Based on the nature of this data, we can simulate it by means of the solution of the Oral-glucose regulatory (second-order system) model, as an under-damped glucose-concentration response curve, given by:

$$y(t) = (G / \omega)e^{-At}\sin\omega t, \tag{2}$$

$$\omega = \left(\omega_n^2 - A^2\right)^{1/2}$$

wherein A is the attenuation constant, is the damped frequency of the system, thenatural frequency of the system = ω_n, and $\lambda = 2A/T_d$.

The model parameters λ and T_d are obtained by matching eqn.(1) to the monitored glucose concentration y(t) data (represented by the open circles). The computed values of parameters are: $\lambda = 2.6$ hr^{-2}, $T_d = 1.08$ hr. This computed response is represented in Figure 1 by the bottom curve, fitting the open-circles clinical data.

ParametricIdentification (sample calculation for Normal Test Subject No.5)

$$\begin{aligned}
&y\ (1/2) = (G/\omega)e^{-A/2}\sin\omega/2 = 0.34\ g/L \\
&y\ (1) = (G/\omega)e^{-A}\sin\omega = 0.24\ g/L \\
&y\ (2) = (G/\omega)e^{-2A}\sin 2\omega = -0.09\ g/L
\end{aligned}$$

Using trignometry relations, we get

$$\begin{aligned}
&A = 0.8287\ \text{hr}^{-1} \\
&\lambda = \omega_n{}^2 = A^2 + \omega^2 = (0.82875)^2 + (2.0146)^2 = 4.7455\ \text{hr}^{-2}, \\
&T_d = 2A/\lambda = 0.3492\ \text{hr}
\end{aligned}$$

Upon substituting the above values of λ and Td, the value of the third parameter,

$$G = 1.2262\ g\ (l)^{-1}\ \text{hr}^{-1}$$

For a diabetic subject, the blood-glucose concentration data is depicted by closed circles in Fig 16. For the model to simulate this data, we adopt the solution of model eqn(17), as an over-damped response function:

$$y(t) = (G/\omega)\ e^{-At}\sinh\omega t \tag{3}$$

The solution ($y = (G/\ \omega)\ e^{-At}\sinh\omega t$) is made to match the clinical data depicted by closed circles, and the values of λ and T_d are computed to be 0.27 hr^{-2} & 6.08 hr, respectively. The top curve in **Figure 16** represents the blood-glucose response curve for this potentially diabetic subject. The values of T_d, λ and A for both normal and diabetic patients are indicated in the figure, to provide a measure of difference in the parameter values.

It was found from these calculations that not all of the normal test subjects' clinical data could be simulated as under-damped response. Similarly, not all the diabetic test subjects' clinical data corresponded to over-damped response.

However it was found that the clinical data of these test subjects (both normal and diabetic) could indeed be fitted by means of a critically-damped glucose-response solution of the governing equation.

$$y(t) = G\,t\,e^{-At} \tag{4}$$

for which, $\omega = 0$, $\omega_n^2 = A^2 = \lambda$, and $T_d = 2A/\lambda = 2$

Clinically-based Diagnosis:

The blood glucose 'normal' values, used for the clinical studies, were: Fasting: 70 to 115 mg/dl, At 30th min.: less than 200 mg/dl,

At 1st hour: less than 200 mg/dl, At 2nd hour: less than 140 mg/dl,

Modeling-based Diagnosis:

The test subjects have been classified into four categories:

Normal-test subjects based on under-damped model-response;

Normal test-subjects based on critically-damped model-response, at risk of becoming diabetic;

Diabetic test-subjects based on critically-damped model-response, being border-line diabetic;

Diabetic test-subjects based on over-damped model response;

Non-Dimensional Number for Diagnosis of diabetes:

We decided to develop a unique diabetes index number (DIN) to facilitate differential diagnosis of normal and diabetic states as well as diagnose supposedly normal but high (diabetic) risk patients and diabetic patients in early stages of the disorder [8].

$$\text{DIN}= \frac{y(\max)}{G}\times A\times\frac{T_d}{T(\max)} \tag{5}$$

wherein,

y(max) = maximum blood glucose value in gram/liter

G = glucose dose administered to the system in gram/liter hour

A= attenuation constant in 1/hour

T_d= derivative-time ($\alpha+\delta$) in hours

T(max) = the time at which y(max) is attained in hour

This non-dimensional number DIN consists of the model parameters (A & T_d) or (A & ω_n) or (λ & T_d). The DIN values for all four categories were computed from equation (5). A distribution plot of the DIN is plotted in fig 17, wherein the DIN is classified into sections with 0.2 increments (for all the four categories of subjects) and the number of subjects which fall into these sections (frequency) is determined.

In the distribution plot (shown in Fig 17), the DIN values 0-0.2 is designated as range 1, the DIN 0.2-0.4 is range 2, 0.4-0.6 is range 3, and so on up to DIN 2.2-2.4, which is range 12.

As can be seen from figure 17, normal (i.e., non-diabetic) subjects with no risk of becoming diabetic, will have DIN value less than 0.4, or be in the 1 – 2 range. Distinctly diabetic subjects will have DIN value greater than 1.2, or be in the 7 – 12 range categories.
Supposedly, clinically-identified normal subjects who have DIN values between 0.4 and 1.0, or are in the 3 – 5 range, are at risk of becoming diabetic.
On the other hand, clinically-identified diabetic subjects with DIN value between 0.4 – 1.2, or in the 3 – 6 range category are border-line diabetics, who can become normal (with diet control and treatment).

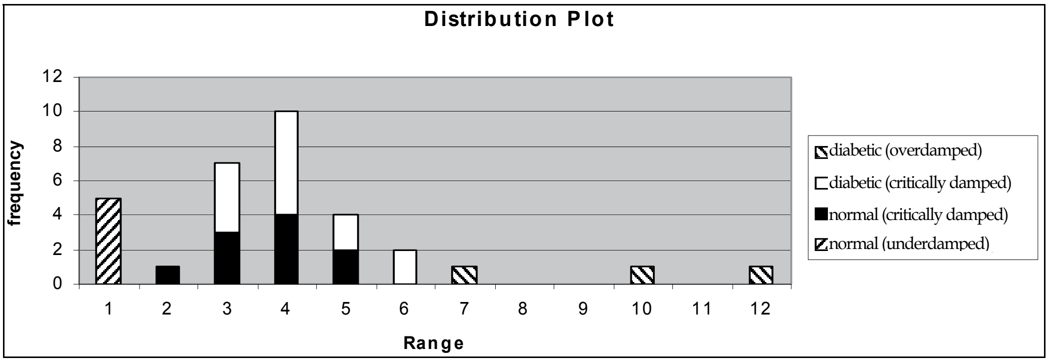


Fig. 17. DIN distribution plot of all the four categories subjects [8]. In the figure, diabetic (critically damped) category of subjects are designated as border-line diabetic; normal (critically damped) category of subjects are designated as normal subjects at risk of becoming diabetic.

7. Cardiology: LV contractility index based on normalized wall-stress

The traditional cardiac contractility index $(dP/dt)_{max}$ requires cardiac catheterization.

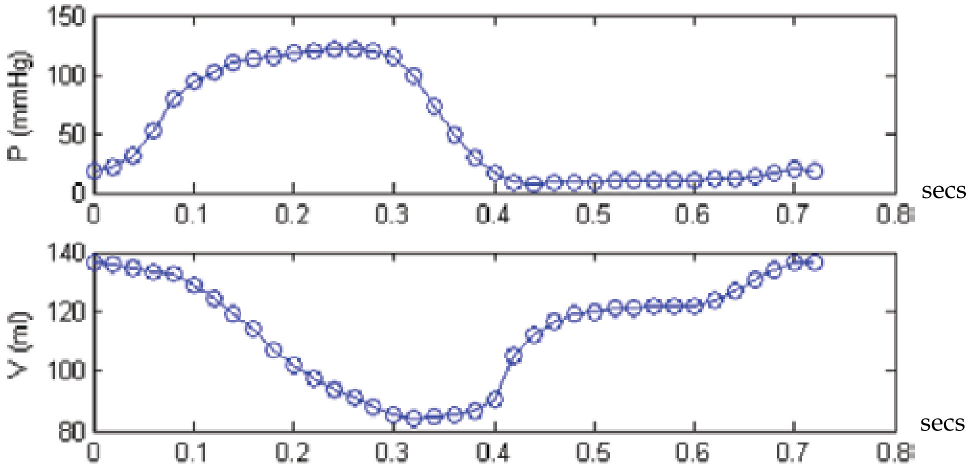


Fig. 18. Sample cyclic variation of LV pressure and volume.

Since LV Pressure is developed by LV wall stress σ_θ(based on sarcomere contraction), we have developed a contractility index based on σ_θ (normalized with respect to LV pressure) [9].
For a thick walled spherical model of LV, the circumferential wall stress:

$$\sigma_\theta(r_i) = P\frac{(r_i^3 / r_e^3 + 0.5)}{(1 - r_i^3 / r_e^3)} \tag{1}$$

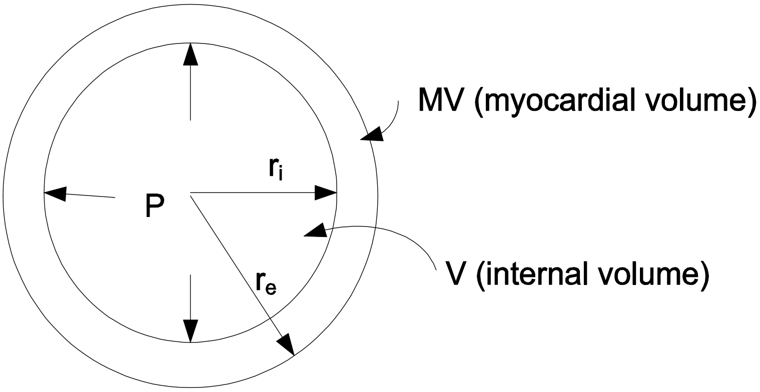


Fig. 19. Thick walled Spherical Model of the LV.

$$\text{The normalised Stress} = \frac{\sigma_\theta(r_i)}{P} = \sigma^* = \frac{3V}{2MV} + \frac{1}{2} \tag{2}$$

We define the contractility index as:

$$\textit{Contractility Index (CONT1)} = \left|\frac{d\sigma^*}{dt}\right|_{\max} = \frac{3}{2MV}\left(\frac{dV}{dt}\right)_{\max} \tag{3}$$

For the data shown in the figure 18, we have:

$$CONT1 = \frac{3}{200cc}(224cc \cdot s^{-1}) = 3.3s^{-1}$$

Now we formulate a non-dimensional cardiac contractility index,

$$CONT2 = \left|\frac{d\sigma^*}{dt}\right|_{max} \times ejection\ period(= 0.3second) \times 100$$
$$\approx 100$$

Our new contractility indices do not require measurement of LV pressure, and can hence be evaluated noninvasively. In fig 20, we can see how well our contractility index *CONT*1 $(d\sigma^*/dt)_{max}$ correlates with the traditional contractility index $(dP/dt)_{max}$. This provides a measure of confidence for clinical usage of this index.

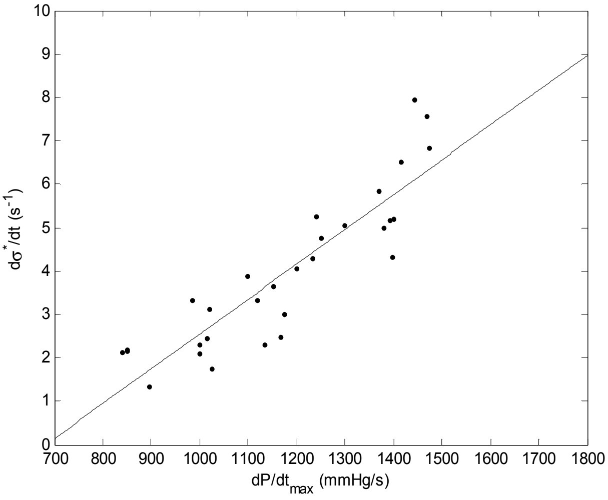


Fig. 20. Correlation of our new contractility index *(dσ*/dt)_{max}* with the traditional contractility index(dP/dt)_{max}

8. Diagnostics: LV contractility index based on LV shape-factor

Cardiologists have observed shape changes taking place in an impaired LV. We have investigated the effect of ventricular shape on contractility and ejection function. In this study, a new LV contractility index is developed in terms of the wall-stress (σ*, normalized with respect to LV pressure) of a LV ellipsoidal model (Fig. 21) [10, 11].

Using cine-ventriculography data of LV volume (V) and myocardial volume MV, the LV ellipsoidal model (LVEM) major (B) and minor axes (A) are derived for the entire cardiac cycle. Thereafter, our new contractility index $(d\sigma^*/dt)_{max}$ is derived in terms of the LV ellipsoidal shape factor *(s=B/A)*.

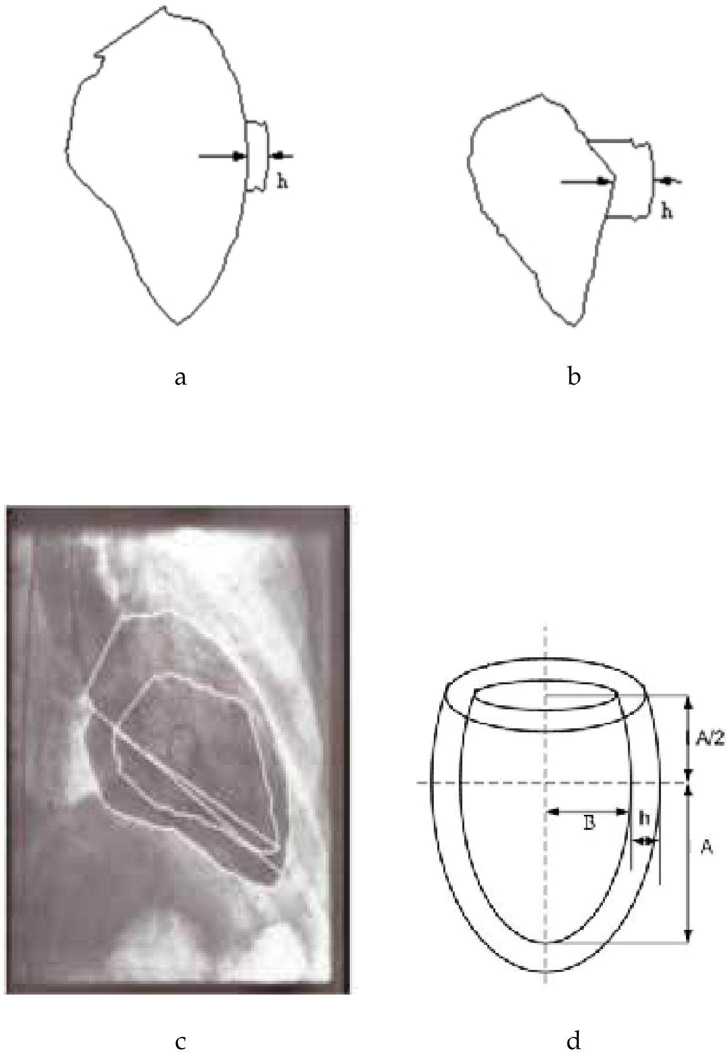
For the LV model (Fig 21) of a prolate spheroid truncated 50% of the distance from the equator to the base, we first put down the for σ* ($= \sigma_\theta / P$), and then determine dσ*/dt [2.10]. Thereby, we obtain the following expression for the contractility index:

$$\text{Contractility index} - 1\,(CONT1) = \left|(d\sigma^* / dt)\right|_{\max}$$

$$= \left| \frac{\dot V(2+s)+V\dot s}{MV} - \frac{V^2 s}{MV(4V+2Vs+MV)^2} \begin{pmatrix} s\dot V(8+4MV/V+(8+2MV/V)s+2s^2) \\ +V\dot s(16+4MV/V+(16+3MV/V)s+4s^2) \end{pmatrix} \right|_{\max} = F(s,\dot s,V,\dot V,MV)$$

where s is B/A, ś is first-time derivative of s; V and MV are LV volume and myocardial wall volume, Ṽ is the first-time derivative of V.

This index is analogous to the traditional employed index $(dP/dt)_{max}$, but does not involve determination of the intra-LV pressure by catheterization. For patient A (with myocardial infarct and double vessel disease) and B (with double vessel disease and hypertension), the values of CONT1 are obtained to be 3.84 and 6.90 s^{-1}, whereas the corresponding values of $(dP/dt)_{max}$ are obtained to be 985 and 1475 mmHg/s.



Comparison of the constructed LV ellipsoidal model geometry with the cineangiography derived geometries:

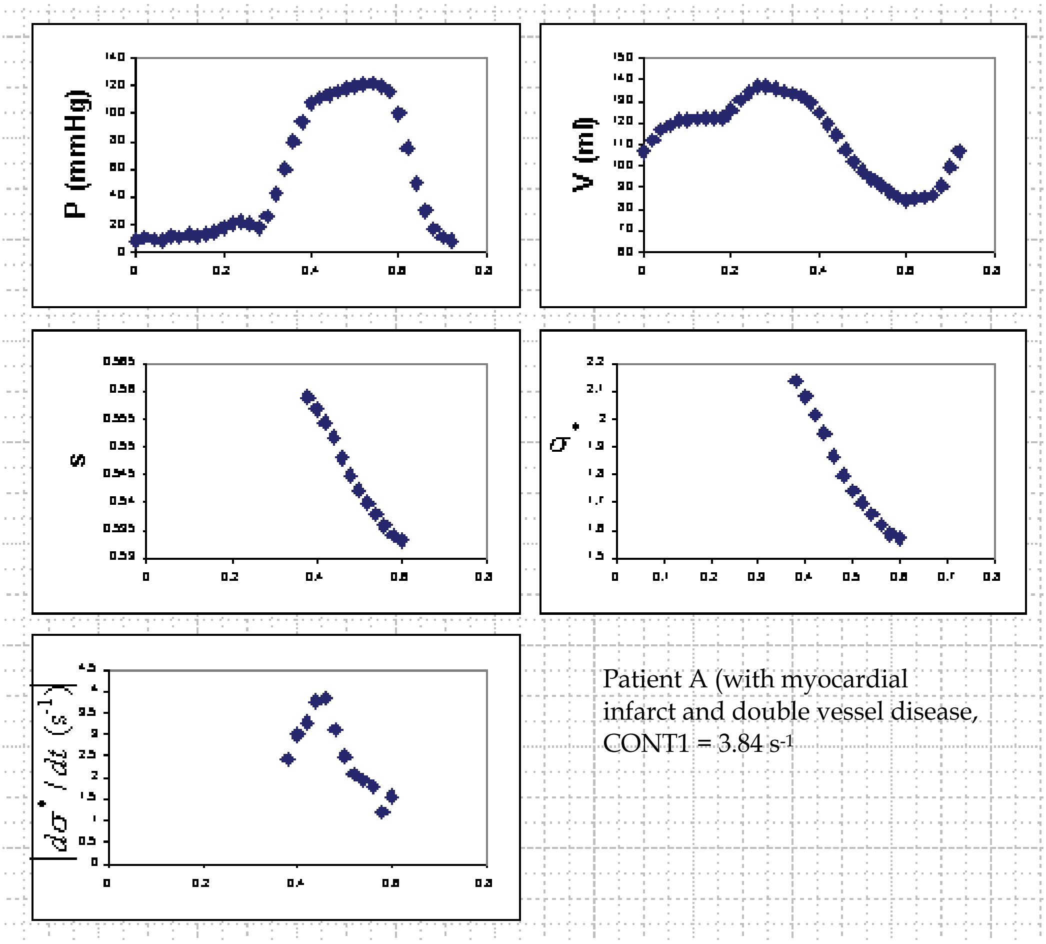
(a) at an instant during the diastolic filling phase,

(b) at an instant during the systolic ejection phase,

(c) cineangiography measured LV dimensions,

(d) LV ellipsoidal model geometry, showing the model geometrical dimensions.

Fig. 21. Cineangiography imaged LV geometry and the corresponding constructed LV ellipsoidal model: (a) at an instant during the diastolic filling stage, (b) at an instant during the systolic ejection stage, (c) measured LV dimensions, (d) LV ellipsoidal model, depicting its geometrical parameters.



Patient A (with myocardial infarct and double vessel disease, CONT1 = 3.84 s^{-1}

Fig. 22. Patient A: Cyclic variation of LV pressure-volume data, LV model shape factor s, computed σ*, computed index $(d\sigma^*/dt)_{max}$.

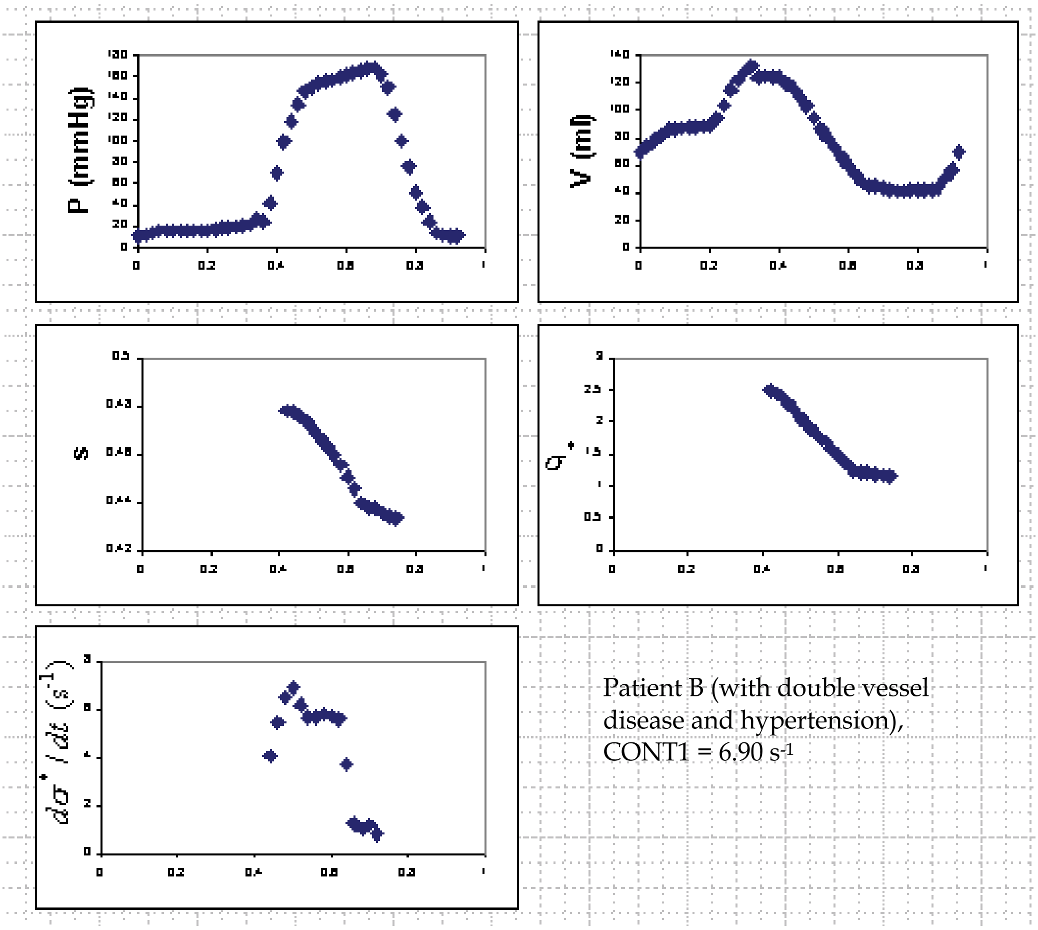
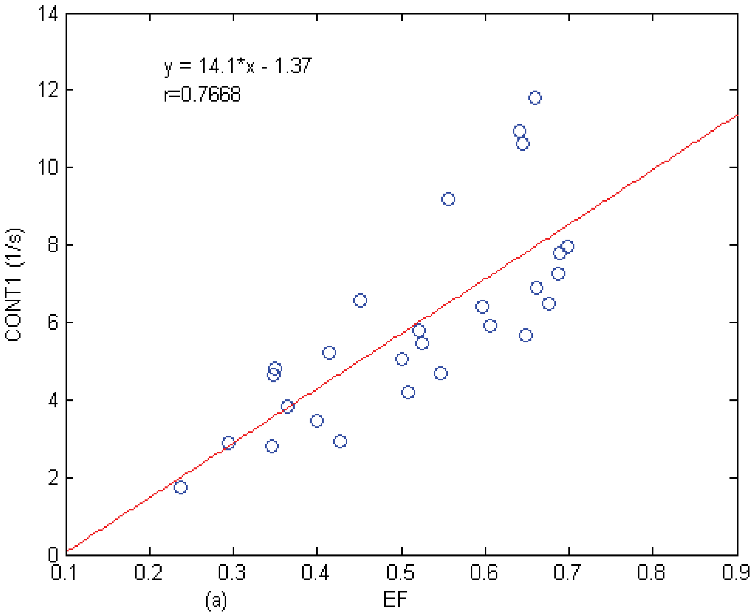


Fig. 23. Patient B: Cyclic variation of LV pressure-volume data, LV model shape factor s, computed σ*, computed index *(dσ*/dt)*$_{max}$.



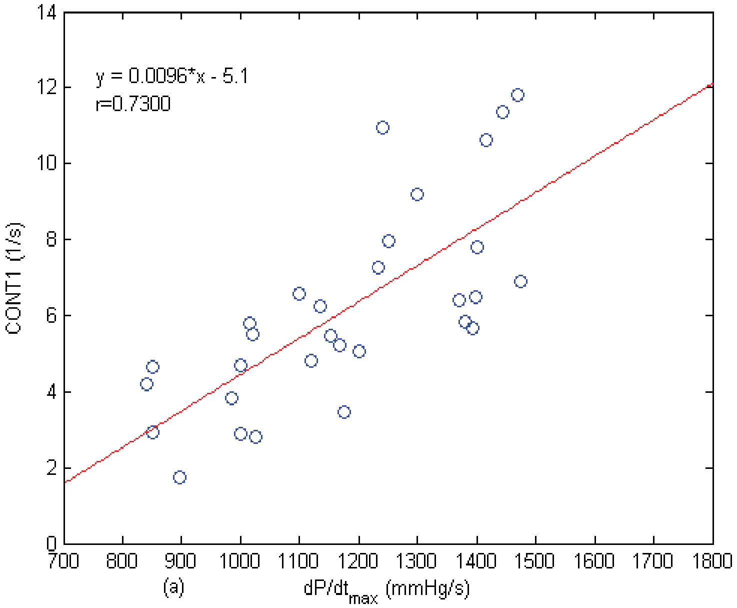


Fig. 24. Correlationships between (i) CONT1 and EF, and (ii) CONT1 and *(dP/dt)*$_{max}$.

For our patient group, we have computed and plotted CONT1 vs EF and $(dP/dt)_{max}$ in Fig. 24. We can note the good level of correlation with the traditional contractility index of $(dP/dt)_{max}$. Additionally, our new index can be determined noninvasively, and hence be more conducive for clinical use.

From these results, we can also infer that a non-optimal less-ellipsoidal shape (or a more spherical shape, having a greater value of $S = B/A$) is associated with decreased contractility (and poor systolic function) of the LV, associated with a failing heart. This has an important bearing on a quick assessment of a failing heart based on the values of S and $(d\sigma^*/dt)_{max}$

9. ICU Evalution: Indicator for lung-status in mechanically ventilated copd patients (using lung ventilation modelling and assessment)

In chronic-obstructive-pulmonary-disease (COPD), elevated airway resistance and decreased lung compliance (i.e. stiffer lung) make breathing difficult. After these patients are mechanically ventilated, there is a need for accurate predictive indicators of lung-status improvement, for ventilator discontinuation through stepwise reduction in mechanical support, as and when patients are increasingly able to support their own breathing, followed by trials of unassisted breathing preceding extubation, and ending with extubation.

So, in this section, we have provided a biomedical engineering analysis of the lung ventilator volume response to mechanical ventilation of COPD patients, and developed an index to assess the lung status as well as the basis of weaning the patient from ventilator support.

Figure 25 depicts the model for the lung volume (V) response to the net driving pressure P_N=P_L– P_e(*end-expiratory pressure)*, in which (i) the *driving pressure*P_L= P_m (*pressure at the mouth*) minus P_p (*the pleural pressure*), and (ii) P_p is determined by intubating the patient, and assuming that the pressure in the relaxed esophageal tube equals the pressure in the pleural space surrounding it.

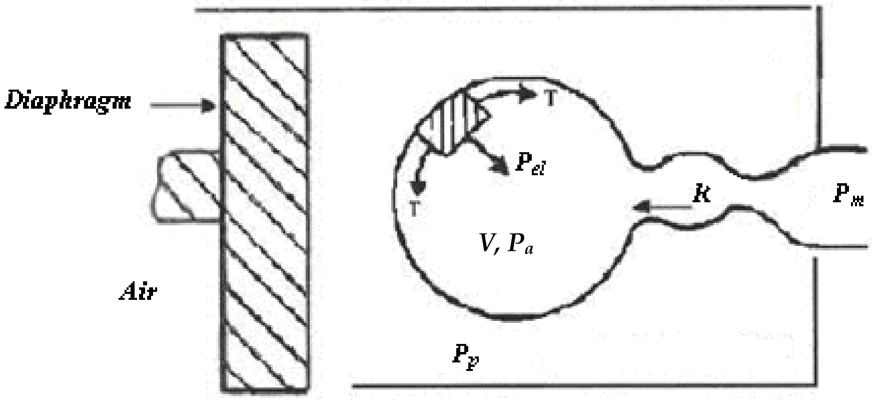


Fig. 25. Model of the Lung, depicting P_m ,P_p, P_{el} (lung elastic pressure recoil) = P_a (alveolar pressure) – P_p (pleural pressure) = 2T/(radius of alveolar chamber), and
R (resistance to airflow rate) = $(P_m - P_a) / (dV/dt)$.

The equation representing the lung model response to the net driving pressure in terms of the model parameters lung compliance *(C)* and airflow-resistance *(R)*, is given by:

$$R\overset{o}{V}+\frac{V}{C}=P_L(t)-P_e=P_N(t) \tag{1}$$

wherein:

1. the driving pressure = P_L ; Pe = the end-expiratory pressure; net pressure $P_N = P_L - Pe$
2. the parameters of the governing equation (1) are lung compliance (C) and airflow-
3. resistance(R), with both R & C being instantaneous values
4. V= V(t) – Ve (wherein Ve is the end expiratory lung volume)

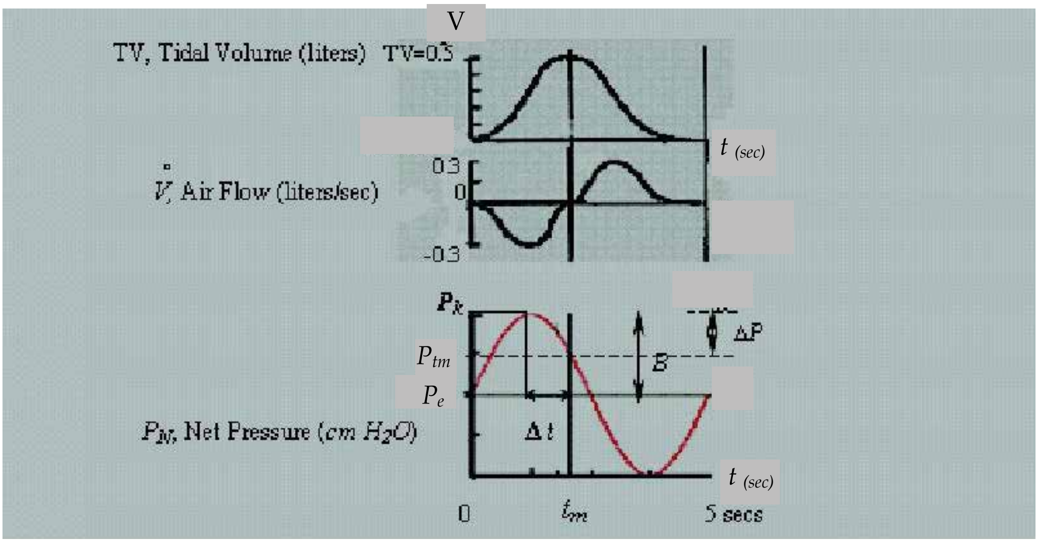


Fig. 26. Lung ventilatory model data shows air-flow ($\overset{o}{V}$) and volume (*V*) and net pressure (P_N). Pause pressure (P_{tm}) occurs at t_m, at which the volume is maximum (TV = tidal volume). Δ*t* is the phase difference between the time of maximum volume and peak pressure (P_k). It also the time lag between the peak and pause pressures. *B* is the amplitude of the net pressure waveform P_N applied by the ventilator. This P_N oscillates about P_e with amplitude of *B*. The difference between peak pressure P_k and pause pressureP_{tm} is Δ*p* [12].

We measure Peak pressure (P_k), Pause pressure (P_{tm}), t_m & ω (or t_m ω).
Then,

$$P_k = P_L(t = \pi / 2\omega) = P_e + B$$

$$P_{tm} = P_L[t = t_m = (\pi - \theta) / \omega] = P_e + B\sin \omega\, t_m = P_e + B \sin [\omega (\pi\text{-}\theta)/\omega]$$

$$= P_e + B\sin(\pi - \theta) = P_e + B\sin\theta$$

$$\therefore\ B = (P_k - P_{tm})/(1 - \sin\theta) = \Delta P/(1\text{-}\sin\theta)$$

B is the amplitude of the net pressure wave form applied by the ventilator. Let C_a be the average dynamic lung compliance, R_a the average dynamic resistance to airflow, the driving pressure $P_L = P_e + B \sin(\omega t)$, and the net pressure $P_N = B \sin(\omega t)$. The governing equation (1) then becomes [12]:

$$R_a \overset{o}{V} + \frac{V}{C_a} = P_N = B\sin(\omega t) \tag{2}$$

The volume response to P_N, the solution to eqn (2), is given by:

$$V(t) = \frac{BC_a\{\sin(\omega t) - \omega k_a \cos(\omega t)\}}{1+\omega^2 k_a^2} + He^{-t/k_a} \tag{3}$$

wherein:

i. ka (=RaCa)is the average time constant,
ii. the integration constant H is determined from the initial conditions,
iii. the model parameters are Ca and ka(i.e. Ca and Ra), and
iv. ω is the frequency of the oscillating pressure profile applied by the ventilator

An essential condition is that the flow-rate is zero at the beginning of inspiration and end of expiration. Hence, applying this initial condition of *dV/dt* =0 at t=0 to our differential eqn (3), the constant *H* is obtained as:

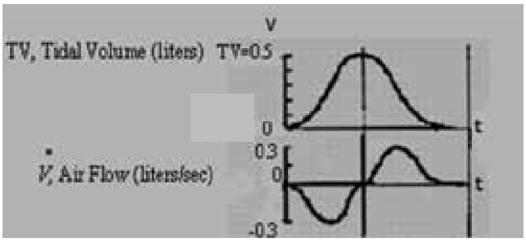
$$H = \frac{BC_a \omega k_a}{1+\omega^2 k_a^2} \tag{4}$$

Then from eqn (3) & (4), we obtain:

$$V(t) = \frac{BC_a\{\sin(\omega t) - \omega k_a \cos(\omega t)\}}{1+\omega^2 k_a^2} + \frac{BC_a \omega k_a}{1+\omega^2 k_a^2} e^{-t/k_a} \tag{5}$$

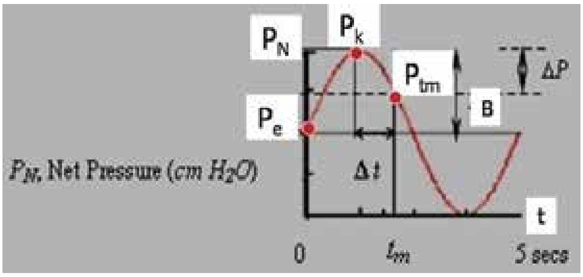
Evaluating parameters R_a & C_a [12,2]:

For evaluating the parameter $k_a(R_aC_a)$, we will determine the time at which *V (t)* is maximum and equal to the tidal volume (TV), Hence putting *dV/dt* =0 in eqn (5), we obtain:



$$\cos(\omega t) + \omega k_a \sin(\omega t) = e^{\left(\frac{-t}{k_a}\right)}, \text{ at t = tm} \tag{6}$$

From equation (6), we obtain (by neglecting $e^{-t/ka}$), the following expression for k_a:



$$k_a = -1/\omega \tan(\omega t_m) \tag{7a}$$

$$or, \quad \tan^{-1}(1/\omega k_a) = \pi - (\omega t_m) = \theta \tag{7b}$$

From eqn (5 & 6):

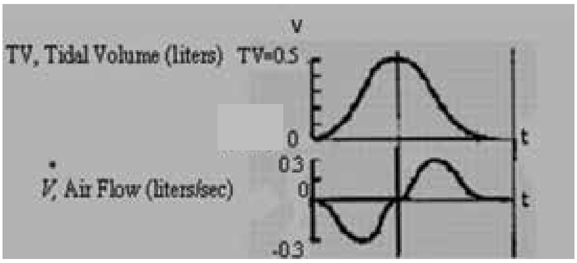
$$V(t = t_m) = \mathrm{TV} = \frac{BC_a\{\sin(\omega t_m) - \omega k_a \cos(\omega t_m)\}}{1+\omega^2 k_a^{\;2}} + \frac{BC_a \omega k_a}{1+\omega^2 k_a^2} e^{-tm/ka} \tag{8}$$

At $t = t_m$, the second term,

$$H = \frac{BC_a \omega k_a}{1+\omega^2 k_a^2} e^{-tm/ka} \approx 0 \tag{9}$$

Hence, eqn (8) becomes:

$$V(t = t_m) = \mathrm{TV} = \frac{BC_a\{\sin(\omega t_m) - \omega k_a \cos(\omega t_m)\}}{1+\omega^2 k_a^2} \tag{10}$$



In eqn (10), if we put:

$$N = \sin(\omega t_m) - \omega k_a \cos(\omega t_m) \tag{11}$$

Then, based on equations (6 and 7), we get:

$$N = \frac{1}{\sqrt{1+\omega^2 k_a^2}} + \frac{\omega^2 k_a^2}{\sqrt{1+\omega^2 k_a^2}} = \sqrt{1+\omega^2 k_a^2} \tag{12}$$

Now then, based on equation (12), equation (10) becomes:

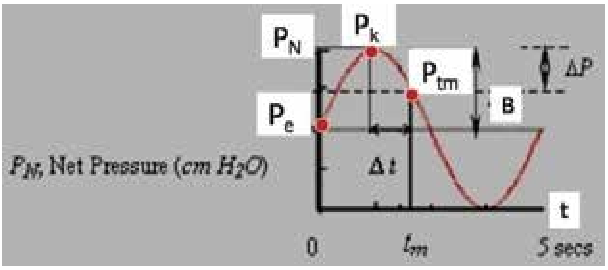
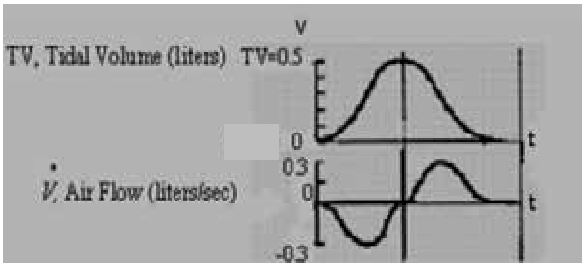


Fig. 26. (reproduced)

$$V(\ t = t_m) = \mathrm{TV} = \frac{BC_a}{\sqrt{1+\omega^2 k_a^2}} \tag{13}$$

Determining Lung-Compliance (Ca) and Air-Flow Resistance (Ra):



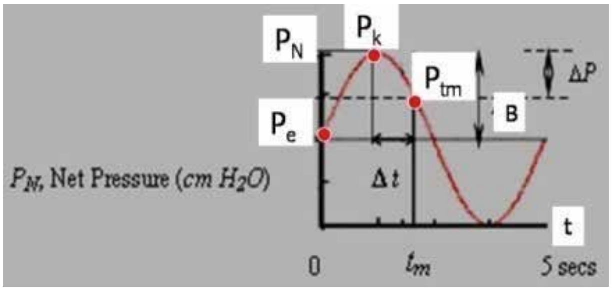
From equations (13 and 7), we get:

$$C_a = \frac{\mathrm{TV}\sqrt{1+\omega^2 k_a^2}}{B} = \frac{\mathrm{TV}}{B\sin\theta} = \frac{\mathrm{TV}(1-\sin\theta)}{\Delta P\sin\theta} \tag{14}$$

Hence, from eqns (7 & 13), the average value of airflow-resistance (R_a) is:

$$R_a = k_a / C_a = \frac{\Delta P\sin\theta(1/\omega\tan\theta)}{\mathrm{TV}(1\text{-}\sin\theta)} = \frac{\Delta P\cos\theta}{\mathrm{TV}\omega(1\text{-}\sin\theta)} \tag{15}$$

For our patients, the computed ranges of the parameters are:



$$C_a = 0.020 - 0.080 \text{ L / cmH}_2 0 \tag{16}$$

$$R_a = 9 - 43 \text{ cmH}_2 0 \cdot \text{s / L}$$

Now, that we have determined the expressions for R_a and C_a, the next step is to develop an integrated index incorporating these parameters.

Formulating a Lung Ventilatory Index (LVI) incorporating Ra and Ca: We now formulate a Lung Ventilatory Index (LVI), incorporating Ra and Ca, as:

$$LVI = \frac{R_a(RF)P_k}{C_a(\text{TV})} \tag{17}$$

Let us now obtain order-of-magnitude values of *LVI,* for a mechanically ventilated COPD patient in acute respiratory failure:

$$LVI(Intubated\ \text{COPD}) = \frac{[15cmH_2Os / L][0.33s^{-1}][20cmH_2O]}{[0.035L / cmH_2O][0.5L]}$$
$$= 5654\ (cmH_2O / L)^3$$

wherein

$$R_a = 15\,cmH_2Os / L \qquad C_a = 0.035L / cmH_2O \qquad RF = 0.33s^{-1}$$
$$\text{TV} = 0.5L \qquad P_k = 20cmH_2O$$

Now, let us obtain an order-of-magnitude of *LVI* (by using representative computed values of R_a, C_a, *RF,* TV, andP_k) as abovefor a COPD patient with improving lung-status just before successful discontinuation.

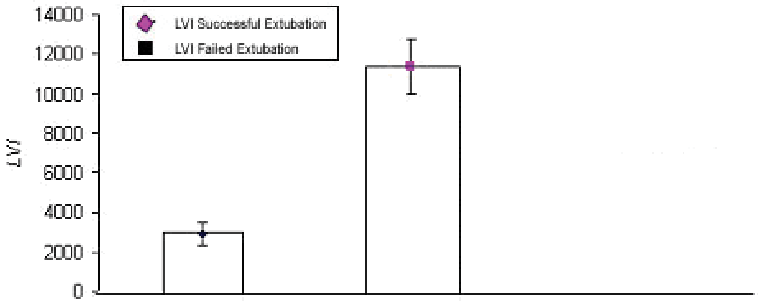
$$LVI(Outpatient\ \text{COPD}) = \frac{[10cmH_2Os / L][0.33s^{-1}][12cmH_2O]}{[0.05L / cmH_2O][0.35L]}$$
$$= 2263(cmH_2O / L)^3$$

wherein

$$R_a = 10\,cmH_2Os / L \qquad C_a = 0.050L / cmH_2O \qquad RF = 0.33s^{-1}$$
$$\text{TV} = 0.35L \qquad P_k = 12cmH_2O$$

Hence, for *LVI* to reflect lung-status improvement in a mechanically ventilated COPD patient in acute respiratory failure, it has to decrease to the range of *LVI* for an outpatient COPD patient at the time of discontinuation.

In fig. 27, it is shown that for the 6 successfully-discontinued cases, the *LVI* was (2900) ± (567) $(cmH_2O/L)^3$; for the 7 failed-discontinuation cases the *LVI* was (11400) ± (1433) $(cmH_2O/L)^3$. It can be also observed that *LVI* enables clear separation between failed and successful discontinuation, which again points to the efficacy of *LVI.*



(6) Successful Extubation Case (7) Failed Extubation Case

Fig. 27. Distribution of *LVI* at discontinuation for patients with failed and successful discontinuation. For the 6 successfully-discontinued cases, the *LVI* was (2900) ± (567) $(cmH_2O/L)^3$; for the 7 failed-discontinuation cases the *LVI* was (11400) ± (1433) $(cmH_2O/L)^3$. It is observed that *LVI* enables clear separation between failed and successful discontinuation [12].

10. Monitoring: Noninvasive determination of aortic pressure, aortic modulus (stiffness) and peripheral resistance)

The aortic blood pressure waveform contains a lot of information on how the LV contraction couples with the aortic compliance and peripheral resistance. Since accurate measurement of aortic blood pressure waveform requires catheterization of the aorta, we have developed a noninvasive method to determine the aortic pressure profile along with the aortic volume-elasticity and peripheral resistance. Fig 28 displays such a constructed aortic pressure profile.

The input to the model consists of auscultatory cuff diastolic and systolic pressures, along with the MRI (or echocardiographically) measured ejection volume-time profile (or volume input into the aorta). The governing differential equation for pressure response to LV outflow rate $I(t)$ into the aorta is given by [13]:

$$\frac{dP}{dt} + \lambda P = m_a \left[I(t) + T_a \frac{dI}{dt} \right] \tag{1}$$

where (i) m_a is the aortic volume elasticity (dP/dV), (ii) R_p is the resistance to aortic flow $(=P/Q)$, (iii) $\lambda = m_a / R_p$, (iv) $I(t)$ is the monitored inflow rate into the aorta, and (v) T_a is the flow-acceleration period.

This governing equation is solved for measured $I(t)$ and dI/dt during the systolic phase from time T_1 to T_3 (Fig. 28). For the diastolic-phase solution from time T_3 to T_4, the right-hand side is zero. The solutions for diastolic and systolic phases are given below by equations (2) and (3) respectively.

Solutionequations:

- During diastolic phase,

$$P_d(t) = P_1 e^{\lambda(T-t)} \tag{2}$$

At $T = T_4$, $P_d(T)$=auscultatory $P_{ad} = P_1$

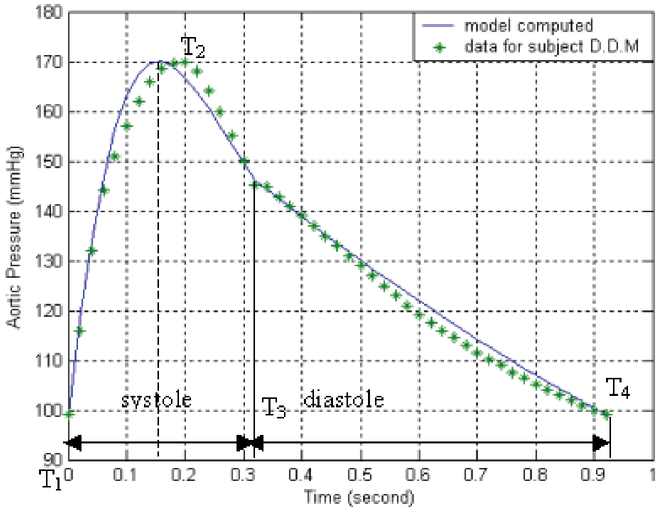


Fig. 28. Model computed cyclic Aortic pressure profile, compared with measured pressure values. The systolic phase from *T1* to *T3* is when blood is ejected into the aorta. The diastolic phase is from *T3* to *T4* [13].

- During systolic phase,

$$P_s(t) = (P_1 - A_1)e^{-\lambda t} + e^{-bt}\left[A_1\cos(\omega t) + A_2\sin(\omega t)\right] \tag{3}$$

where

$$A_1 = \frac{m_a a\omega(T_a\lambda - 1)}{(b-\lambda)^2 + \omega^2}$$

$$A_2 = \frac{m_a a[(\omega^2 + b^2)T_a + \lambda - b - T_a b\lambda]}{(b-\lambda)^2 + \omega^2}$$

Also, as noted in Fig 28, the boundary values are:
At $t = T_2$, $dP_s/dt = 0$; at $t = T_2$, $P_s(T_2)$ = auscultatory $P_{as} = P_2$; at $t = T_3$, $P_s(t = T_3) = P_d(t = T_3)$
Hence, based on these boundary values, the following equations are to be solved.

$$\frac{dP_s(t)}{dt}(t = T_2) = -C_1\lambda e^{-\lambda T_2} + e^{-bT_2}\left[A_3\cos(\omega T_2) - A_4\sin(\omega T_2)\right] = 0 \tag{4}$$

- At $t = T_2$, $P_s = P_2$. Hence, from equations (3 & 4), we get:

$$P_s(t = T_2) = P_2 = C_1 e^{-\lambda T_2} + e^{-bT_2}\left[A_1\cos(\omega T_2) + A_2\sin(\omega T_2)\right] \tag{5}$$

- At $t = T_3$, $P_s(t = T_3) = P_d(t = T_3)$. Hence, from equations (2 and 3), we obtain:

$$P_1 e^{\lambda(T - T_3)} = (P_1 - A_1)e^{-\lambda T_3} + e^{-bT_3}\left[A_1\cos(\omega T_3) + A_2\sin(\omega T_3)\right] \tag{6}$$

We now solve equations (4, 5 & 6), to determine the unknown parameters m_a, R_p (and T_2).

We have determined the expressions for the aortic pressure during the systolic and diastolic phases, by solving the governing equation (1), for the monitored LV outflow rate (or input into the aorta) $I(t)$, using (i) the monitored auscultatory diastolic pressures (P_{ad}), to serve as the boundary condition at the beginning of the systolic-phase solution (at time T_1) and at the end of the diastolic-phase solution (at time T_4), (ii) the monitored auscultatory systolic pressure (P_{as} $=P_2$) to represent the maximum value of the systolic-phase solution.

Because the pressure solution of equation (1) is a function of m_a and R_p, we first determine the values of these parameters by making the solution satisfy the above-mentioned boundary conditions, which in turn yields the pressure profile given by $P_d(t)$ and $P_s(t)$. This non-invasively determinable aortic pressure can provide hitherto unavailable information on vascular compliance and resistance status, as well as on the capacity of the LV to respond to it.

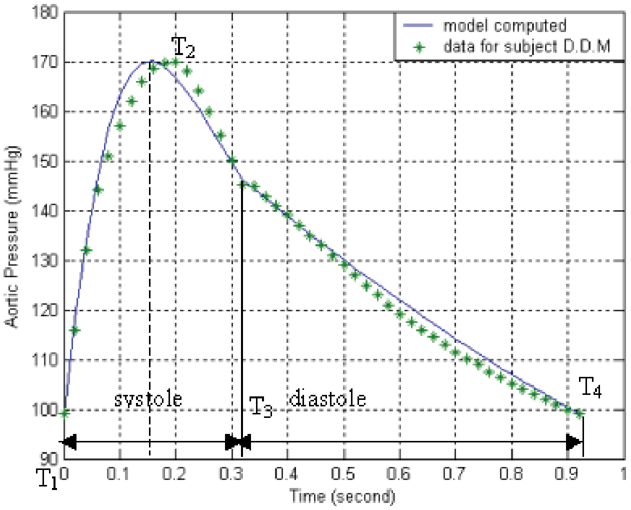


Fig. 29. Plot of aortic pressure (of patient B) during one cardiac cycle, t_e=0.32s. Herein, T_1-T_3 represents the systolic phase, and T_3-T_4(=0.92s) represents the diastolic phase. At T_2, the aortic systolic pressure profile has its maximal value ($=P_s$). The scatter points are the data measured from catheterization. The solid line is the model-computed profile; RMS=2.41 mmHg. Note the excellent correlation between the model-derived aortic pressure profile and the catheter-obtained aortic pressure profile.

In **Fig. 29**, the model-computed aortic pressure profile patient B (with double vessel disease and hypertension) is shown, along with the actual catheter pressure data. We can note how well the model-computed result matches the actual catheterization data, with RMS 2.41 mmHg. The aortic stiffness (m_a) and peripheral resistance (R_p) are obtained to be 1.03 *mmHg/ml* and 1.59 *mmHg ·s/ml*, respectively.

Let us consider yet another benefit of this analysis. We have determined aortic pressure profile, aortic stiffness (m_a) or aortic elastance (*Eao*) and peripheral resistance. From the instantaneous aortic pressure and aortic inflow rate, we can determine the instantaneous left ventricular (LV) systolic pressure, in terms of the instantaneous dimensions of the LV outflow tract. We hence determine the LV systolic pressure profile, from which we can evaluate the traditional contractility index (*dP/dt max*) as well as the LV systolic elastance (*Elv*).

We can then determine the ratio of *Elv/Eao*, to represent the LV-Aorta Matching Index (*VAMI*). In ischemic cardiomyopathy patients, this *VAMI* value is depressed. However, following surgical vascular restoration, this index value is partially restored.

11. Coronary Bypass surgery: Candidacy

- **Interventional candidacy based on computed intra-LV flow-velocity and pressure-gradient distributions**

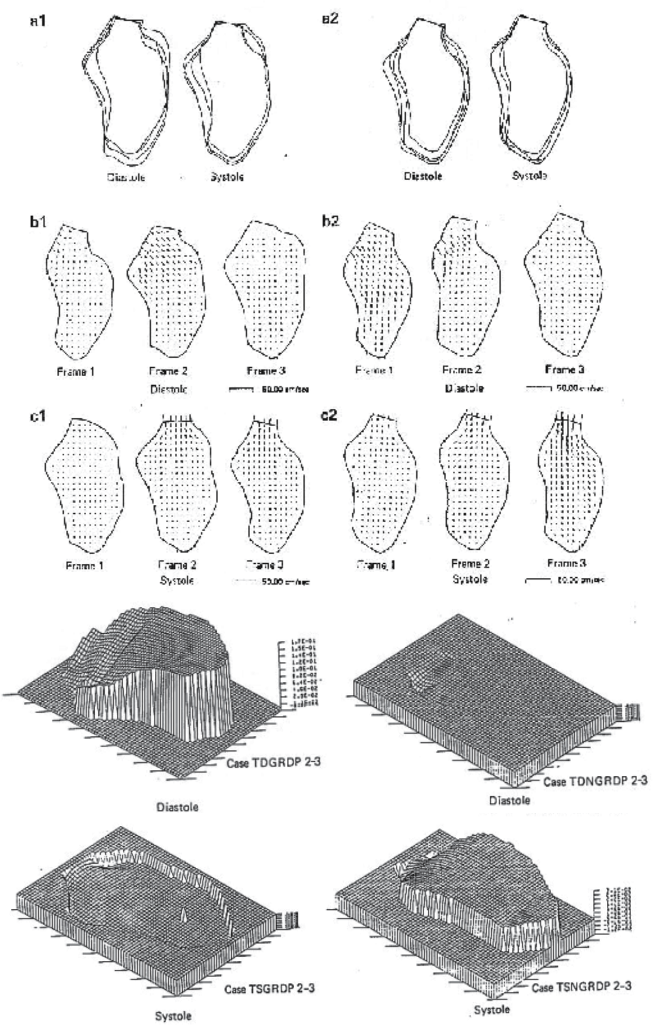

a1
a2
Diastole
Systole
Diastole
Systole
b1
b2
Frame 1
Frame 2
Frame 3
Diastole
50.00 cm/sec
Frame 1
Frame 2
Frame 3
Diastole
50.00 cm/sec
c1
c2
Frame 1
Frame 2
Frame 3
Systole
50.00 cm/sec
Frame 1
Frame 2
Frame 3
Systole
50.00 cm/sec
Case TDGRDP 2-3
Diastole
Case TDNGRDP 2-3
Diastole
Case TSGRDP 2-3
Systole
Case TSNGRDP 2-3
Systole

Fig. 30. Construction of Intra-LV Blood flow velocity and pressure-gradient distributions for a patient with myocardial infarct: (a) Superimposed sequential diastolic and systolic endocardial frames (whose aortic valves centers and the long axis are matched), before (1) and after (2) administration of nitroglycerin; (b) Instantaneous intra-LV distributions of velocity during diastole, before (1) and after (2) administration of nitroglycerin; (c) Instantaneous intra-LV distributions of velocity during ejection phase, before (1) and after (2) administration of nitroglycerin, (d) Instantaneous intra-LV distributions of pressure-differentials during diastole, before (1) and after (2) administration of nitroglycerin; (e) Instantaneous intra-LV distributions of pressure-differential during ejection phase, before (1) and after (2) administration of nitroglycerin. (Adapted from reference 14: Subbaraj K, Ghista DN, Fallen EL, *J Biomed Eng* 1987, 9:206-215.).

A left ventricle with ischemic and infarcted myocardial segments will have lowered ejection fraction and cardiac output, because it will not be able to generate adequate myocardial contraction to raise the intra-LV pressure above aortic pressure for a ling enough duration to generate adequate stroke volume. These patients need coronary bypass surgery, and a pre-surgical assessment of their candidacy for it, on how much they can from it. For this purpose, we need to determine the intra-LV blood flow velocity and pressure-gradient profiles before and after the administration of nitroglycerin.

So, we carry out a CFD analysis of intra-LV blood flow. The data required for the CFD analysis consists of: LV 2-D long-axis frames during LV diastolic and systolic phases; LV pressure vs. time associated with these LV frames; Computation of LV instantaneous wall velocities as well as instantaneous velocity of blood entering the LV during the filling phase and leaving the LV during the ejection phase.

From this CFD, we have determined the instantaneous distributions of intra-LV blood-flow velocity and differential-pressure during filling and ejection phases, to intrinsically characterize LV resistance-to-filling (LV-RFT) and contractility (LV-CONT) respectively. The results are summarized in the above Fig. 30**.**

By comparing intra-LV pressure-gradients before and after administration of nitroglycerin (a myocardial perfusing agent, and hence a quasi-simulator of coronary bypass surgery), we can infer how the myocardium is going to respond and how these LV functional indices will improve after coronary by bypass surgery.

12. Theory of hospital administration: Formulation of hospital units' performance index and cost-effective index

A Hospital has clinical services departments, medical supply and hospital-services departments, financial-management and administrative departments. Each of these five sets of departments has to function in a cost-effective fashion.

Let us, for example, consider the Intensive-Care Unit (ICU) department. The human resource to an ICU dept consists of physicians and nurses. Using activity-based costing, we can determine the human-resource strength, based on an assumed reasonable probability-of-occurrence of (for instance) two patients simultaneously (instead of just one patient) having life-threatening episodes.

Performance Index: We can formulate the ICU performance-indicator in terms of the amounts by which the physiological health-index (PHI or NDPI) values of patients were (i) enhanced in the ICU for those patients discharged into the ward from the ICU, and (ii) diminished in the ICU in the case of patients who died in the ICU.

Let us say that patients are admitted to the CCU if their Physiological-health-index (PHI) value falls below 50%. So if the PHI of a patient improves from 30 to 50, then the **Patient-HealthImprovement Index (PHII)** for that patient is given by [15],

$$\mathrm{PHII} = \left(\frac{50-30}{30}\right) 100 = 67 \ (or\ 67\%) \tag{1}$$

Thus the **patient health-improvement index (PHII)** value is higher if a more seriously-ill patient is discharged from the ICU, and lower if a not-so-seriously ill patient is discharged, i.e., if

$$\mathrm{PHII} = \left(\frac{50-40}{40}\right) 100 = 25 \ (or\ 25\%) \tag{2}$$

We can then formulate the **Performance-index (PFI)** for an ICU as follows:

$$ICU\text{ Performance index }(\text{PFI}) = \frac{\sum PHII\text{ of the patients}}{\#\text{ of those patients treated during a life time period}} \tag{3}$$

Hence, the higher the value of ICU performance index, the better is the performance of the ICU. If now a patient dies, as a result of the PHII becoming negative (i.e slipping from (say) 30 to 10), then

$$PHII = 100\left(\frac{10-30}{30}\right) = -67 \tag{4}$$

As a result, ΣPHII (in equation 3) will decrease, and the overall value of ICU performance index (namely PFI, as calculated by means of eq. 3), will fall.

Cost-Effective Index: Now, let us consider that (i) we have one physician and five nurses for a 10 bed CCU, based on the probability-of-occurrence of two patients having life-threatening events being say 0.2 (or 20%), and that (ii) for this human resource/staffing, the ICU performance index value (PFI) is (say) 40.

If we increase the staffing, the ICU performance index value could go upto 50 or so, at the expense of more salary cost. So now we can come up with another indicator namely, **Cost-effectiveness index (CEI)**

$$\text{CEI} = \frac{\text{Performance index}}{\text{Total salary index (in salary - units)}}$$

$$= \frac{\text{Performance index}}{\text{Resource index (in terms of salary - units)}} \tag{5}$$

wherein, say, a salary of 1000=0.1 unit, 10,000=1 unit, 20,000=2 units, and so on.

So if an ICU has one physician with a monthly salary of 20,000 (i.e. 2 salary-units) and four nurses each with a total monthly salary of 5,000 (i.e. total of 2.0 salary-units), then

$$\begin{aligned}\mathbf{CEI}(\text{ICU}) &= \frac{\text{Performance index (of 40)}}{\text{Salary-units Index or Resource Index,}R_i\,[=(2+2.0)]}\\ &= \frac{40}{4.0} = 10\end{aligned} \tag{6}$$

Now, let us assume that we raise the **PFI** (ICU) to (say) 60 by augmenting the nursing staff, so as to have six nurses (R_i = 3 units) and 1.5 full-time equivalent physician-on-duty (R_i = 3 units), then

$$CEI\ (ICU) = \frac{PFI}{R_i} = \frac{50}{(3+3)} = 8.3 \tag{7}$$

So while the PFI of ICU has gone up from 40 to 50, the CEI of ICU has gone down from 10 to 8.3.

Strategy of Operation: Our strategy would be to operate this Performance-Resource system, in such a way that we can determine the resource index R_i for which we can obtain acceptable values of both PFI and CEI. In a way, figure (31) could represent this balance between PFI & CEI, in order to determine optimal Resource Index or resource [15].

With reference to this **figure 31**, if we have a resource value of $R_i = R_1$, then the corresponding PFI ($=PFI_1$) will be unacceptable, as being too low; hence, we will want to increase the value of R_i. If we have a resource value of $R_i = R_2$, then our corresponding CEI ($= CEI_2$) will also be unacceptable, for being too low; hence we will want to decrease R_i. In doing so, we can arrive at the optimal value R_{io}, for which both CEI and PFI are acceptable. This procedure, for converging to $R_i = R_{i0}$, can be formulated computationally.

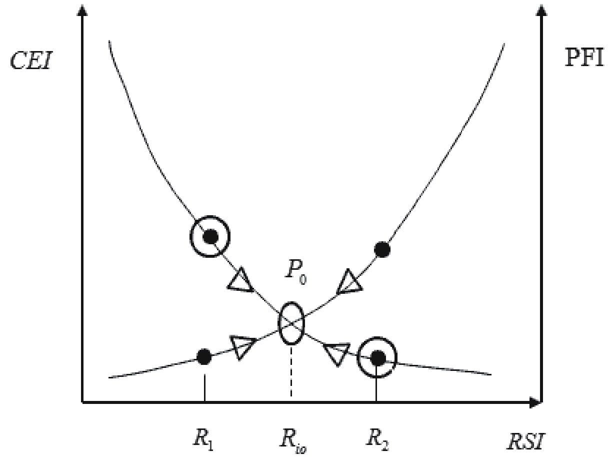


Fig. 31. Optimising the value of R_i so to obtain acceptable values of CEI and PFI [15].

Now then, let us formulate how a hospital budget can be optimally distributed. Let us say that a Hospital has 'n' number of departments and a prescribed budget (or budget index, BGI). We would want to distribute the budget among the departments, such that none of the 'n' departments has a PFI below the acceptable value of PFI_a and a CEI below the acceptable value of CEI_a.

So the Operational research problem is to be formulated as follows:

How to distribute or divide the given Budget (or Budget Index Value) into R_i (i= 1,……..,n), such that $PFI_i \geq PFI_a$, for all i; and $CEI_i \geq CEI_a$, for all i This then is the prime task of a Hospital administrator!

Summarizing:

The ICU Operational cost (OPC) is the cost of operating the ICU over this one month period. The ratio of this ICU PFI and the ICU OPC is the Cost-effectiveness index (CEI) of ICU.
The Management strategy is to maintain certain acceptable values of both PFI and CEI for all hospital departments, by judicious allocation of staff to the departments.
This enables the determination of optimal Resource index (RSI) and hospital budget (HOB) to maintain a balance between PFI and CEI for all the hospital departments. This can constitute the basis of Hospital Management.

13. *How to proceed*: For biomedical engineering to become a professional field

So in this chapter, we have talked about:

1. *Anatomy:* Spine as an optimal structure,
2. *Physiology:* Mechanism of LV pressure generation

3. Non-dimensional Physiological Index
4. *Medical Test:* Cardiac Fitness based on Treadmill Heart Rate Variation
5. *Medical Physiology:* A Non-dimensional Diabetes Index with respect to Oral-Glucose-tolerance Testing
6. *Cardiology:* LV Contractility based on its normalized wall stress
7. *Diagnostics:* LV Contractility based on its shape factor
8. *ICU Evaluation:* Indicator for lung-status in Mechanically ventilated Copd patients, using lung ventilation modeling and assessment
9. *Monitoring:* Noninvasive determination of Aortic pressure, aortic stiffness and peripheral resistance
10. *Coronary Bypass Surgery:* candidacy
11. Hospital Operations Management

We have shown how biomedical engineering (BME) can open up a new approach to the study of Anatomy, in terms of how anatomical structures are intrinsically designed as optimal structures for their function.

We have also seen how BME can be applied to the study of Physiology, by modelling of physiological systems (as bme models), which can enable us to assess their performance for clinical usage by means of our novel NDPIs comprised of the physiological systems' model parameters. The NDPIs can considerably facilitate medical assessment, and lead us to what can be effectively termed as Higher-order Translational Medicine (HOTM), entailing the incorporation of physical and engineering sciences into medical sciences and clinical sciences for more reliable and effective patient care.

Even Medical tests can be adroitly modelled as BME systems. These systems are formulated in the form of mathematical models, whose solutions are simulated to the medical tests data, to evaluate the model parameters. The model parameters can be combined into NDPIs, by means of which the test data can be analysed for making more reliable medical assessment and decisions.

Finally, we have seen how we can bring to bear the enormous scope of Industrial Engineering (and constrained optimisation theory) to hospital operational management, so as to develop (i) cost-effective performance indices of hospital departments, and (ii) more knowledgeable framework for budget development and allocation to the various department.

Putting all of this together is what will justify (i) the incorporation of BME into the MD curriculum, and (ii) its constituting an indispensable patient-care oriented department in tertiary- care hospitals.

This verily constitutes the biomedical engineering professional trail, from anatomy and physiology to medicine and into hospital administration. This is what the professional role of biomedical engineering needs to be, to promote a higher order of translational medicine and patient care.

14. References

[1] Nondimensional Physiological Indices for Medical assessment, by Dhanjoo N. Ghista, in *J of Mechanics in Medicine and Biology*, Vol 9, No 4, 2009.

[2] *Applied Biomedical Engineering Mechanics*, by Dhanjoo N. Ghista, CRC press (Taylor & Francis Group) Baton Rouge Florida 334872-2742, ISNBN 978-0-8247-5831-8,2008

[3] Human Lumbar Vertebral Body as an Intrinsic Functionally-optimal Structure, by D.N. Ghista, S.C. Fan, K.Ramakrishna, I.Sridhar, in *International Journal of Design and Nature*, 2006, 1(1): 34-47.

[4] The Optimal Structural design of the human Spinal Intervertebral disc", by D.N.Ghista, S.C.Fan, I.Sridhar, K.Ramakrishna, in *International Journal of Design and Nature*, 2006, 1(2).

[5] Mechanism of Left Ventricular Pressure increase during Isovolumic contraction, and determination of its Equivalent myocardial fibers orientation, by Ghista, DN, Liu Li, Chua LP, Zhong L, Tan RS, Tan YS; in *J Mech Med Biol*, 2009, 9 (2), 177–198

[6] Cardiac Fitness mathematical Model of Heart-rate response to V02 during and after Stress-Testing", Lim GeokHian, Dhanjoo N. Ghista, Koo TseYoong, John Tan Cher Chat, Philip EngTiew& Loo Chian Min; *International Journal of Computer Application in Technology(Biomedical Engineering & Computing Special Issue),* Vol 21, No 1/2, 2004.

[7] Glucose Tolerance Test Modeling & Patient-Simulation for Diagnosis, by Sarma Dittakavi & Dhanjoo N. Ghista, *Journal of Mechanics in Medicine & Biology,* Vol. 1, No.2, Oct.2001.

[8] Clinical Simulation of OGTT Glucose Response Model for Diagnosis of Diabetic Patient, by Dhanjoo N.Ghista, Patrick S.K. Chua, Andy UtamaAulia, Peter L.P. Yeo, in *Journal ofMechanics in Medicine & Biology* 2005 Vol 5, No. 1.

[9] Validation of a novel noninvasive characterization of cardiac index of left ventricle contractility in patients, by Zhong L, Tan RS, Ghista DN, Ng E. Y-K, Chua LP, Kassab GS, *Am J Physiol Heart CircPhysiol*2007, 292:H2764-2772.

[10] LV shape-based contractility indices, by Liang Zhong, Dhanjoo N. Ghista, Eddie Y-K. Ng, Lim SooTeik and Chua Siang Jin, CN Lee, in *Journal of Biomechanics*, 2006, 39: 2397-2409.

[11] Measures and Indices for Intrinsic Characterization of Cardiac Dysfunction during Filling and Systolic Ejection, by Liang Zhong, Dhanjoo N. Ghista, Eddie Y. Ng, Lim SooTeik, and Chua Siang Jin, in *Journal of Mechanics in Medicine and Biology*,Vol 5, No. 2, 2005.

[12] Indicator for Lung-status in a mechanically Ventilated (COPD) Patient using Lung Ventilation Modeling and Assessment, by D.N.Ghista, R. Pasam, S.B. Vasudev, P.Bandi, and R.V. kumar in *Human Respiration Anatomy and Physiology, Mathematical Modellings and Applications,* ed by Vladimir Kulish, WIT Press, 2006.

[13] Determination of Aortic Pressure-time Profile, Along with Aortic Stiffness and Peripheral Resistance, by Liang Zhong, Dhanjoo N. Ghista, Eddie Y-K. Ng, Lim SooTeik and Chua Siang Jin, in *Journal of Mechanics in Medicine & Biology* 2004, 4(4):499-509.

[14] Intrinsic Indices of the Left Ventricle as a Blood Pump in Normal and Infarcted Left Ventricle, by K. Subbaraj, D.N. Ghista, and E. L. Fallen, in *J of Biomedical Engineering*, Vol 9, July issue, 1987

[15] Physiological Systems' Numbers in Medical Diagnosis and Hosipital Cost-effective Operation, by Dhanjoo N. Ghista, in *Journal of Mechanics in Medicine & Biology* 2005, vol 4, No.4.

Part 1

Biomedical Science: Disease Pathways, Models and Treatment Mechanisms

2

Cell Signalling and Pathways Explained in Relation to Music and Musicians

John T. Hancock
University of the West of England, Bristol,
UK

1. Introduction

Cell signalling is arguably the most important area of modern biology. The subject encompasses the control of cellular events, especially in response to extracellular factors. It has been suggested that the human body is one of the most complex machines ever produced (Dawkins, 1989) and the regulation of the activities within it are also equally complex.

Interest in cell signalling does not simply stem from an academic viewpoint either. Certainly there is a vast resource of research which is focuses on the investigation of signalling pathways and the control which they bestow on a cell. However, there are tangible reasons to take an interest here too. The vast majority of new pharmaceutical compounds under development are aimed at the modulation of proteins involved in cell signalling events (Filmore, 2004). Such proteins may be G protein-coupled receptors (GPCRs) or perhaps kinases which are downstream of such receptors. Many anti-cancer studies are now focused on the development of compounds which modulate Mitogen Activated Protein Kinases (MAPKs) for example. Therefore, an understanding the working of the components of a signal transduction opens up avenues for the future modulation of such activities with the development of new therapies and pharmaceutical agents.

The study of cell signalling can seem very daunting. Vast diagrams full of acronyms can put off the most ardent reader, but there are many basic principles which underpin the subject. In cell signalling compounds are made and initiate a response, and this is true whether the molecule originates outside the cell or is created inside. The signal transduction pathway carries a “message”, with such a message originating in one place, either outside of the cell, or from another part of the same cell, but having a response elsewhere. The keys to cell signalling are that the message needs to be conveyed in a specific manner, so that it is not scrambled and misconstrued and that the cell must be able in most cases to revert back to a state or activity in which it was engaged before the message arrived, that is, the signal transduction pathway needs to be stopped when the message is no longer needed to be conveyed.

Even though the principles are simple, it is still hard to understand the complexities of cell signalling. Often signalling events are over-simplified, and components are aligned in neat rows. However, a more holistic view shows that signalling is extremely complicated and hard to understand. There are many books and chapters which explain cell signalling (Hancock, 2003; Hancock, 2010; Krauss, 2008) but these are all based on the description of the science, with the molecular details often putting off the reader who may be new to the field. Therefore, often an analogy to explain such a complex subject would be very useful, and may offer a more attractive way to teach the subject and to engage those who seek a better understanding of the area of study. In this chapter music is used as an analogy to try to shed light on some of the events in cell signalling. It has already been suggested that the use of such an analogy will be useful to those trying to get to grips with the subject (Hancock, 2005; 2009) and this chapter will expand and elaborate on those ideas. It is suggested that this can be used by those studying and teaching cell signalling.

2. Music and musical terminology

Listening to music and watching music being played are both events which rely on cell signalling. Sound waves are perceived in the ear, while photons are sensed in the eye and both lead to downstream series of signal transduction events (Hancock, 2010). In fact early work which led to the discovery of a major class of proteins, the G proteins, was due to work on the eye (Fung & Stryer, 1980). However, this is not of particular relevance to the discussion here. There is, however, growing, if controversial evidence that music can have effects on biological systems, including humans (Trappe, 2010). A term has been coined, "The Mozart Effect". This has come about from work where Mozart's music has been played and effects measured. Some tangible effects have been reported, perhaps more in the popular press than the scientific literature, but there are examples of serious reports looking into this (Jenkins, 2001).

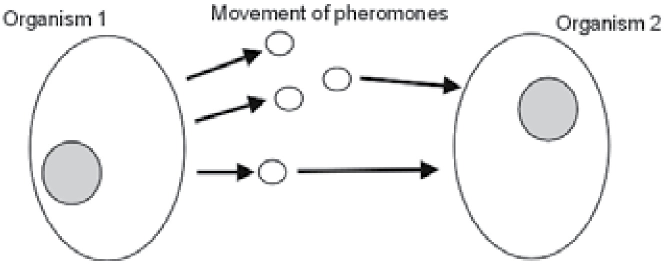


Fig. 1. Movement of pheromones transmits signals between individuals in a population. One organism releases a compound which is sensed by a second individual, in which a response is mounted. Pheromones could be thought of as music moving through the medium separating the two organisms.

Of course it is not only humans which are responsive to sound with studies reporting effects on plants for example (Qin *et al.*, 2003). Mechanical action on cells has been shown to affect cellular function (Wan *et al.,* 2004) and this includes exposure to music. For example, the activity or expression of some proteins has been shown to be changed if music is played (Chikahisa *et al.*, 2006).

Not only does music provoke cell signalling events in organisms, but music terminology is often used to describe such cellular activities. It is often said in research papers that a signalling molecule "orchestrates" or "conducts" events for example (Polo & de Fiore, 2008). In his most recent book Nick Lane uses music to explain his thinking on more than one occasion. On the theme of biological variation he discusses the musical variations of Bach and Beethoven. On the topic of protein structure he says "Yet the deeper music of the protein spheres is still there to be discerned by crystallography" and later when talking about the eye he writes “Like an orchestral conductor conjuring up the most beautiful music without sounding a note himself, the gene calls forth the structures of the eye by ushering in individual players, each with their own part to play” (Nick Lane, 2010). Therefore it is an extension of this idea which can enable music in its wider sense to act as an analogy for cell signalling. Previously the idea has been discussed (Hancock 2005; 2009) but here the ideas are expanded and enhanced.

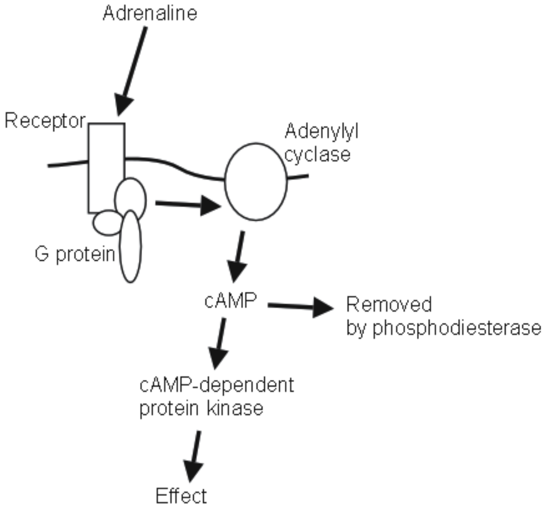


Fig. 2. cAMP is involved in the signalling invoked by adrenaline. Adrenaline is received by a plasma membrane bound receptor. The receptor is linked to a heterotrimeric G protein, which releases its alpha subunit on activation. The G protein subunit can activate adenylyl cyclase which resides in the plasma membrane. Adenylyl cyclase produces cAMP from ATP. cAMP can activate cAMP- dependent protein kinase and so lead to downstream responses. cAMP is removed by the action of phosphodiesterase.

3. Signals between organisms

Music is often produced by an individual or group of individuals, and listened to, or perceived by, another individual or group of individuals. There is an excellent example of such action in cell signalling and that is the generation and response to pheromones (Agosta, 1992; Kell *et al.*, 1995). Here small compounds are made and released into the medium outside of the organism (Figure 1). They are then carried in the flow of the liquid or gas, perhaps the atmosphere, which surrounds the organism to be sensed and responded to by another individual of the same species, but not the individual organism that released them. Pheromones are used for attracting a mate and sexual arousal for example. In human behaviour interestingly often music is used for the same purpose.

4. Production of signals

For a cell to use a molecule in a cell signalling pathway it needs to firstly be made. Many of the components are constitutively produced and are present to partake in the required activity when called upon to do so. Examples would be large proteins such as kinases.

However, there are many situations where a molecule needs to be present, or released, in a rapid manner. Cyclic AMP (cAMP) for example is needed in response to adrenaline (for an example of a signalling pathway in which cAMP is involved see Figure 2) while insulin is released in to the blood stream when required. One of the main underlying principles of signalling is that the system is able to convey the message when and where required, in a temporal manner appropriate to the required response. Therefore molecules need to be able to partake in such signalling when called upon to do so.

There are two main ways to make a signal. Either the molecule is produced when required, or it is made and stored, to be released when required (Figure 3). It a similar manner it could be argued that there are two main ways to listen to music. You either go to a concert and in the presence of the musical instruments you listen to the sounds being made, or you let the band record the music, store it until required and then play it. In this scenario the instruments are the enzymes, producing the message. At the concert instruments make the signal as needed, to the required amount, for the required time. This is just like an enzyme such as adenylyl cyclase which is turned on, generates cyclic AMP (cAMP) for a set period of time, and then turns off. Just like the person at the concert, the protein responding to cAMP can perceive its presence and when it is all over revert back to a quiescent state – concert over.

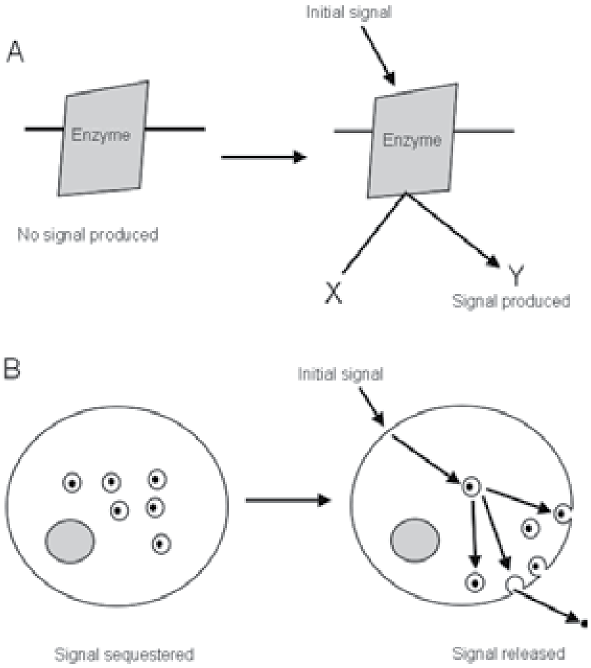


Fig. 3. Signals can be generated using different scenarios. In (A) there is no signal produced until an enzyme is activated. At that point a compound X and be converted to compound Y, so generating a signal. That is Y can now be recognised and a response or effect produced due to its presence. In (B) the signalling molecule can be pre-made but sequestered into vesicles. On arrival of the appropriate stimulus the vesicles will translocate to the plasma membrane for example. The signalling molecule will then be released and be able to move to its site of recognition and action. In the case of hormones the site of perception may be a different organ or tissue, with the signalling compound being carried by the vascular system of the organism.

The production of insulin on the other hand is more akin to a recording. Insulin is encoded for by a single gene, giving rise to a single protein, referred to as pre-proinsulin. This is heavily modified, primarily through cleave events, to produce the active insulin molecule which comprises of two polypeptides. This "ready to use" insulin is then stored in vesicles in the cell until required. This is like a musical recording, the music is created and then stored, sat in its CD case, or as an MP3 file, awaiting to be played. On demand insulin is released by the fusion of vesicles to the plasma membrane in the islets of Langerhans and is released to the outside.

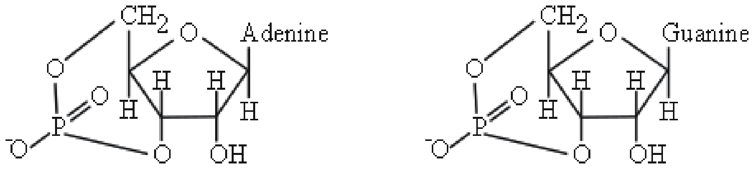
5. Uniqueness of signals

Signals used by cells have to be specific and often are unique so that their presence does not get confused with another. If a signalling molecule needs to provoke a particular response it is vital that the cell's machinery recognises the presence of that specific molecule. If there was doubt then the cell may mount a response in the presence of the wrong molecule. A good example here is role of the molecules cAMP and cyclic GMP (cGMP). As can be seen in Figure 4, at first glance both these compounds look very much the same. Both have a ribose ring, a cyclic phosphate group and an added base unit.

However, the base is different in each and therefore a cell can recognise them as different. Indeed, different enzymes make them, adenylyl cyclase produces cAMP from ATP and guanylyl cyclase makes cGMP from GTP. Downstream they are recognised as separate compounds too. cAMP controls protein kinase A, while cGMP controls amongst other things protein kinase G (Figure 5).

Musical instruments are the same. Take a quick glance at a viola, and then at a violin, and they look the same. They have the same basic shape and the same basic parts. But they are different. In an orchestra they will play different music, at different times perhaps, but what is important here is just like cAMP and cGMP, a violin and a viola have their own distinct roles and parts to play in the construction of the whole, despite the fact that they are outwardly so similar.

There are occasions when cAMP and cGMP can have similar activities, and in some cases both are removed by the same phosphodiesterase. Does this mean that our analogy breaks down? Perhaps not, as if pushed different string players can pick up alternate instruments and allow the orchestra to continue. If a violin player gets ill, a viola player can often step up to the breach to fill the gap.



cAMP cGMP

Fig. 4. Structures of signalling molecules can be very similar and yet unique. Here, the structures of cAMP and cGMP are given as examples, but they have very different signalling roles, controlling different proteins for example, that is, cAMP-dependent protein kinase and cGMP-dependent protein kinase respectively.

6. Domains and common features

In the discussion above it was argued that instruments may be similar but unique. However, there are often quite diverse instruments which share common mechanisms or structures. In signalling many proteins may also share common structures, with those structures having similar roles within the protein. A good example here is the EF hand (Lewitt-Bentley & Rety, 2000), which binds and causes a conformational change in a protein in response to changes in the levels of calcium ions in cells. EF hands can be found in a calcium controlled kinases which are able to phosphorylate downstream proteins, but EF hands are found in a wide range of other proteins too, for example the DUOX proteins involved in reactive oxygen species metabolism (Lambeth *et al.*, 2007). Although there might be subtle differences in the EF hands in different proteins the structural domain is still identifiable as being an EF hand, having the same basic function but being involved in a proteins which when taken as a whole have different functions in the cell.

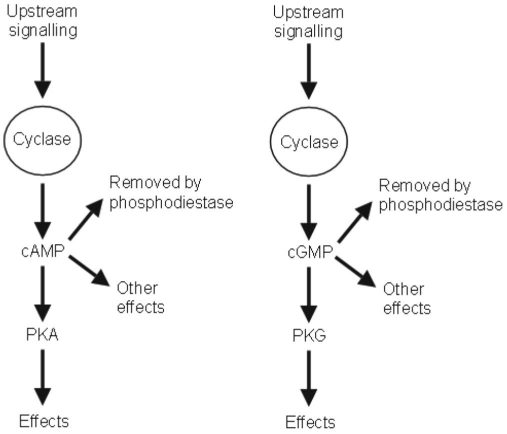


Fig. 5. cAMP and cGMP pathways are very similar. As well as cAMP and cGMP having very similar structures, if drawn in a simplistic view as shown here the signalling pathways in which they are involved is also very similar. Both are produced by cyclases, both are perceived by protein kinase, and in fact both are removed by phosphodiesterases. Both can have other effects and if viewed in more detail there are significant differences in their pathways. Importantly, cAMP and cGMP are involved in specific signalling, despite the similarities. PKA: cAMP-dependent protein kinase; PKG: cGMP-dependent protein kinase.

Therefore many signalling proteins have domains which are similar to each other and to continue our analogy musical instruments are often the same. An idea is repeated, but perhaps has a slightly different role. Consider keyboards on pianos and organs. The idea of having a set of keys which can be pressed can be found on a whole range of instruments, including different types of piano, electric keyboards, organs, harpsichords, accordions, melodicas and many others. Beyond the keyboard the mechanisms may be very different. A piano uses hammers to strike the strings, an organ opens valves to control air into pipes and so on. But the keyboard is a common feature or common structure, like the domain of a protein. And just like the EF hands discussed above, the keyboards in these different instruments may be subtly different, but they are still recognisable as being keyboards despite the overall instruments being quite different, both in shape and the type of music they play.

7. Making up a signalling system

Cells in an organism have the same genetic background, that is, they contain the same genome. The genome will encode for all the proteins which are possible to make in that organism, but different cells will have their own unique complement of proteins. Some cells will express genes for a particular signalling pathway, and others won't, perhaps having a different set of signalling proteins. Some cells will need receptors for a selection of hormones or cytokines, but other cells will have no requirement for the recognition of those extracellular signals and therefore will not express those receptors.

In music the genome is perhaps all the instruments that would be available to someone who wishes to form a group of musicians. There is a vast array of instruments available in, for example, the UK and this could be seen as the "music genome" for that country. The "music genome" for perhaps China or a Far East country may contain other instruments not commonly found in the West. So where ever a person is based to create a group of musicians they will have a "genome" to draw from, just as the cell has a genome containing the genes encoding cell signalling components to draw from. Different places could be viewed like different organisms, having different genomes, although often related. But no cell, and no music producer, would wish to enlist all the possible players. An orchestra conductor will take a wide range of instruments, from violins to kettle drums. A pop group producer may take a couple of guitars and a set of drums and little else. It is the vast array of possibilities that allow a cell to tailor its complement of proteins to allow the signalling that it needs, just as a music group will tailor its array of instruments to create the sound it requires. The cell will be able to respond to a particular group of extracellular signals while the leader of a group of musicians will be able to play a specific selection of music.

8. Receptors and their specificity

Cells are bombarded by signals all the time. It does not matter if it is a single-celled organism or a complex multicellular organism. Outside of the cell will be environmental factors such as salinity, pH, temperature, osmolarity, but on top of these will be the likely presence of compounds such as pheromones, hormones and cytokines. Therefore the cell has to be able to "decide" which it will recognise and mount a response to. The job of such a decision rests with the receptors that the cell has. If the cell synthesises the correct receptor and places it at the right place at the right time then the cell will be able to recognise and respond to the correct signal.

At the start of a practice session for an orchestra the conductor will arrive with all the scores for the piece of music which they intend to play. There will be a score for all the different instruments. The conductor generally has a score with all the parts, and s/he is like a cell which can recognise all the music. However, it would be a waste of time and effort to give such complex scores to all the musicians. Also it would make playing the music extremely unwieldy. With all the parts of the music on the score there is only a short section of notes on any page that can be seen and the musicians would be required to turn the page extremely often. This is fine for the conductor, but the musicians need to be playing, not turning pages. It would be a waste of time and effort for a cell to make all the receptors it is capable of making, that is, having the whole score. What would be the point of making a receptor for a protein hormone that it would never encounter? It would be a misuse of precious materials and space to synthesis such a receptor that will never be used. Therefore the conductor, just like an organism, ensures that each musician has the appropriate score, just with their own music to play.

This analogy can be taken further if the musicians themselves are thought of as receptors. What happens at the start of the practice session? The scores are handed out and each musician will look at the score and confirm that it is the right part, if only silently in their mind. A cello player will only wish to receive cello music: if they take violin music for example it will be in the wrong cleft and probably not able to be played. Violins not only have to get music for the violin but have the right one, being either in first or second violins. If a musician gets the wrong music they will not be able to use it and give the response that the conductor is hoping for. The musician will remain silent, even when they are supposed to be playing, like a cell receptor in the presence of the wrong ligand. It is not uncommon for musicians to return music and ask for the correct score like a receptor rejecting a ligand and leaving themselves free for the arrival of the correct ligand onto their music stand so cellular harmony can be obtained.

9. Single instruments playing a simple string of notes

Cell signalling is often presented as a neat array of components all in a line, as depicted in Figure 6A. In fact Krauss said: "The classical view of signalling pathways has been that of sequential transmission of signals in a linear signalling chain". This would be like listening to solo instruments, playing a linear line of notes on a page. Some representations of cell signalling show little else, making the systems look simple and easy to understand. But just as we don't often listen to single instruments, unless listening to a sonata perhaps, cell signalling is also a complex mix of many players, all adding to the harmony at the same time. As Krauss goes on to point out, signalling is far more complicated (Krauss, 2008).

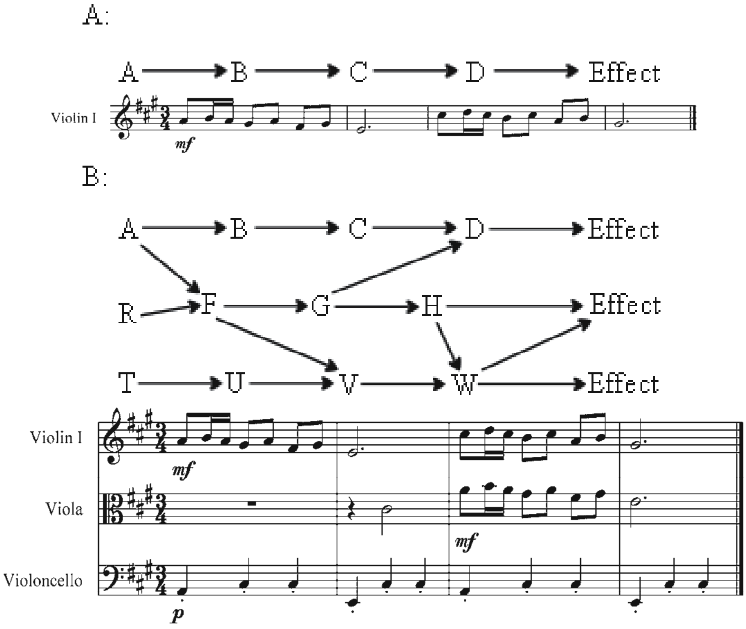


Fig. 6. Signalling pathways are often depicted as a single series of components in a line, akin to a single line of music (A). However, they are usually much more complex, more like several instruments all playing at the same time (B).

Therefore cell signalling should be thought of as a group of instruments and voices, all competing for attention at the same time. Perhaps it is more like depicted in Figure 6B. Just as in a musical group, not all these musicians need to be in action all the time, and in fact it could be rather boring if that was the case. But they will all be there on stage, awaiting their cue from the music, always ready for action, but only acting when needed. In a cell the situation is the same, many signalling components will be quiescent until they are drawn upon to play their role in the control of the cell.

10. Degeneracy

It is often a puzzle in molecular biology that proteins may be able to replace each other. Perhaps an inhibitor has been added which is supposed to remove the functioning of a cell signalling component, but the effect is far less than anticipated. In knock-out or knock-down studies, where the expression of a protein is completed ablated, or severely reduced, then the cell sometimes shows little effect (see Colucci-Guyon *et al.*, 1994 as an example). In a cell signalling response, again often far less than anticipated is seen than from the theory (see Zhang *et al.*, 1994 as an example). In such situations the most likely scenario is that proteins are replacing each other when needed, and protein function is said to be degenerate.

If proteins can replace each other in function can we again invoke a music analogy. The leader of the orchestra is a very important position, often helping and advising the conductor during practise sessions, but during the concert they are not redundant when it comes to control either. Like a signalling component in cells, the leader will be signalling to the rest of the violins when to start playing, so that the whole section plays together and sounds like one. However, what happens when the leader breaks a string. Violin strings are under a large amount of tension and break quite often. If the leader's violin breaks, does the orchestra have to stop. In reality, especially in a rehearsal, it probably would and a new string would be rapidly fitted. But if as in a cell, waiting for a response is not a pragmatic option and the orchestra has to continue then another violin player, probably the one in the second seat next to the leader, will take over and allow the orchestra to continue. One musician, just like the cell's protein, will take up the important role vacated by the other.

Other scenarios in music also can be envisaged. Often musicians can play more than one instrument, and especially in non-professional orchestras and groups it would not be uncommon for one player to take the place of an absence colleague. This would explain why the absence of what was thought to be an important player, or protein, can be seen to not have a devastating effect on the harmony of the music group, or cell.

11. Multiple functions of some proteins

It has been recognised that proteins often have more than one role and are said to be "moonlighting" (Jeffrey, 2009). Here, as well as proteins being able to cover the roles of each other as in degeneracy, some proteins have other very disparate roles.

Some proteins had roles discerned many years ago only to have new additional roles assigned to them more recently. A good example here is cytochrome *c* which was for many years assigned to a redox role in the mitochondria, only later to be found to be instrumental in controlling apoptosis (Figure 7. Reviewed by Jiang & Wang, 2004). Glyceraldehyde 3-phosphate dehydrogenase is central to glycolysis, but it too has now been assigned cell signalling roles, in particular as a protein which translocates to the nucleus to control gene expression (Tristan *et al.*, 2011).

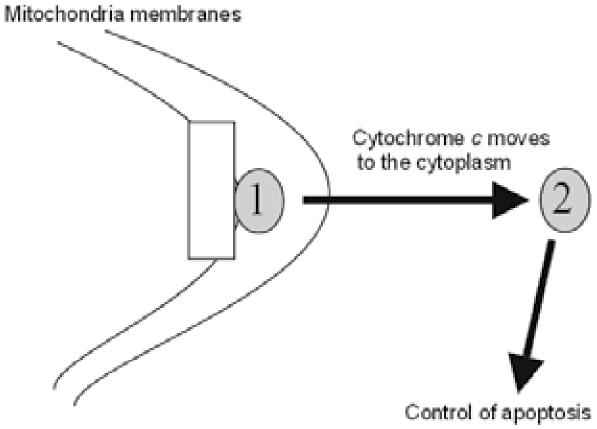


Fig. 7. Cytochrome *c* has more than one function. Cytochrome *c* is normally found in the mitochondria. It resides associated with the inner mitochondrial membrane, where it acts to shuttle electrons from Complex III to Complex IV. However, there is a signalling pathway which leads to cell suicide, or apoptosis, in which cytochrome *c* leaves the mitochondria and moves into the cytoplasm. Here it interacts with the caspase system which leads to ultimately to cell death. Therefore cytochrome *c* has two very distinct and disparate functions.

Musicians and their instruments are like such moonlighting proteins. They may play a violin, but in fact such instruments can be used in multiple music genres. During the day the violinist may be in an orchestra, but some evenings may be playing jazz in the local bar. As with the proteins, their roles are not fixed, and translocation from one venue to another will allow them to partake in a new role. Furthermore, many if not most musicians are experts at more than one instruments. Therefore if an orchestra is short of a viola player, then perhaps a violinist can take their place. Or the change of instrument can be more dramatic, with a flutist can take over on the kettle drums. Just like many proteins, temporal and spatial location of the player, and the interactions in which they may partake, may dictate the exact role they play at any moment in time.

12. Subtle changes make a big difference

In signalling systems there are many components all vying for attention, and to initiate a response. However, even though it is individual components which are studied in many cases, the overall response of the cell will be dictated by the sum of the signalling which is taking place, an idea modelled by Rachmilewitz & Lanzavecchia (2002). Even so, within this holistic approach it needs to be realised that subtle changes in the levels of some signalling components can give a large effect even if there is no change in other proteins and molecules. Certainly some signals have been described as dominant over others. An example here would be a paper by Reya and Clevers (2005) who write “Current evidence indicates that the Wnt cascade is the single most dominant force in controlling cell fate along the crypt–villus axis”. Therefore, a small change in a dominant pathway would initiate a significant response regardless of the activity of less dominant pathways.

Music is often like this, with the single notes in amongst many others having a profound effect. If a pianist plays a major chord, perhaps a C, E and G, to produce a chord of C major, the effect is recognisable, and being major will sound cheerful. However, keeping the C and G the same, but lowering the tone of the E by a semitone to E flat and the chord becomes minor. This is a significant effect. One note amongst the three has changed by a relatively small amount, but the chord is now recognisably different, and the resultant sound has gone from a happy major to a rather sad minor. A small musical change, with a large result. Beethoven uses the alteration of one note in amongst many to great effect in the first movement of the "*Moonlight Sonata*" (Piano Sonata no. 14 in C sharp minor - Op. 27 no 2) for example.

13. Background, volume and thresholds

Signalling in cells needs to take place in the presence a background level of "signalling noise". Cells are bombarded by extracellular signals all the time, whether from the environment of the organism or from other cells of the same organism. The demands on a cell will be constant and varied in many cases. Therefore, if a major response is needed, the signalling that is invoked needs to be "heard" above the noise of the rest of the activity of the cell.

Most signalling systems will in fact be in a state of equilibrium. Often levels of signalling components are measured, perhaps before and after a treatment. However, rarely do the levels of activity, levels of signalling molecules or levels of phosphorylation go from a base level of zero to a higher level. In the vast majority of cases the levels rise from a low level to a transiently higher level. Therefore researchers define threshold levels for signals, (for example Pereyra *et al.*, 2000).

How individual signals get heard may be liken to phrases in the orchestra when one instrument temporally is dominate and can be clearly heard, especially an instrument such as a kettle drum. In music there are often many instruments play all at the same time, and often it is hard to discern the exact contribution of any one instrument. The holistic effect may be pleasing but the parts played by individuals are assumed to be part of the whole. However, one instrument can dominate over the others and be heard above the rest. Perhaps a trumpet is playing a strident part. It will be heard above the other hundred instruments which make up the orchestra. And the audience will follow the trumpet, the tune from which will carry the music and the mood. Cells will have a similar system. Many signals are all contributing, but the arrival of a new hormone may need to dominate. A pathway may be activated, and the activity of the players in that pathway will reach a threshold allowing them to have their effect above that of the other signalling components, which will after all be carrying on doing what they were doing before. However, transiently, the pathway with the "volume" which is dominant will be able to invoke the cellular response needed.

Some instruments such as a bagpipe rely on a background tone, or drone (Nordquist & Ayers, 2009). The highland bagpipe has a tonic note (that is the base note of the scale) of A. Therefore, the other notes are played over this, but the tonic gives the constant tone to the music. It is the other notes which will dominate to give the tune and harmonies. In cells there are various parameters upon which activities and functions of proteins and signalling components will need to contend. In cellular compartments pH is crucial, but as well as this there is the redox state of the environment. In the cytoplasm for example, there is a high concentration of reduced glutathione which will endeavour to maintain the redox state relatively constant (Schafer & Buettner, 2001). This is important because proteins contain reactive side groups which may be affected by the redox of the medium in which they function. Such groups include the thiol groups of cysteine residues for instance. Here, two cysteines may react together in an oxidation reaction to create a disulphide bridge which may stabilise the protein. Alternatively they may react with signalling molecules such as nitric oxide (to be S-nitrosylated) or hydrogen peroxide (to be oxidised). Disruption of the redox state of the cell towards the oxidised state is referred to as oxidative stress, a condition of cells with is extremely important not only to control cell function but also to regulate processes such as apoptosis. Oxidative stress has been implicated in numerous diseases, including degenerative disease (Kadenbach *et al.*, 2009). Therefore it is extremely important for this basal redox state to be maintained, much like a basal note of the bagpipe. It needs to be there, allowing continuity of the harmony of the cell. However, there does need to be the involvement of signals such as hydrogen peroxide and nitric oxide. It may be that the basal redox state in some cases maintains the thiol groups in a state to enable compounds such as nitric oxide to react and have its effect. This would be like the pitched notes on the bagpipe being strident above the background tonal level. On the other hand, if the background is disturbed, perhaps during oxidative stress, such thiols would have already reacted with for example hydrogen peroxide and be no longer available for a reaction it would normally partake in. A disruption of the basal background harmony has altered the effect of the other signals, and the overall effect is quite different. For normal signalling to resume, the background "tonal" redox state would need to be restored.

14. Timing and phasing: Oscillations and waves

One of the intriguing aspects of cell signalling is the timing of the signals and how they fit together temporally. To get a full understanding of signalling needs a full appreciation of both the spatial and temporal aspects of any signal, but particularly how they might be working together in time and space. Early work in this area concentrated on calcium ion signalling, and it was reported that calcium ions were not only altered transiently in some systems but this transient change in ion concentrations actually followed an oscillating pattern. A superb example of this is shown by Alberts *et al.* (1994). Here, the oscillations are dependant on the concentration of the initial signal added. It is not the amplitude of the change which seems to be important in this signalling, but rather the frequency of the oscillations. However, temporal fluctuations on the concentrations of signals are not unique to calcium ions. The biphasic nature of other signalling systems has also been reported, for example with reactive oxygen species (Bleeke *et al.,* 2004) and also with insulin signalling (Rorsman *et al.*, 2000). If hydrogen peroxide levels are followed for example, they increase quickly but transiently, but after a period of relatively low activity the levels once again rise, often to be sustained for the second period. This may be reflected in levels of other signals too, such as nitric oxide. Therefore, at any moment in time the levels of signals may be rising and falling, and it is probably the combined nature of such changes which brings about the desired response in the cell. It is pattern of change which should be considered, rather than the individual changes which might be being recorded.

Music is often written in a pattern. As discussed above, Lane likens biological variations and patterns to musical variations (Lane, 2010). But musical patterns are often phased too. A prime example here is the fugue. Oxford Dictionaries describe a fugue as being written in such as way that "…a short melody or phrase (the subject) is introduced by one part and successively taken up by others and developed by interweaving the parts." A superlative example of such a work is the fugue in the *Toccata and Fugue in D minor*, BWV 565, by Johann Sebastian Bach.

Cell signalling in some cases needs to be thought of in this manner. Hydrogen peroxide and nitric oxide can be considered as two lines of music, one being interwoven with the other. One rising and falling in unison with the changes in the other. But of course it would be naïve to think in these terms for just two signals such as these. Both nitric oxide and hydrogen peroxide can impinge on calcium signalling, so phasing of changes in calcium ion concentrations will need to be considered too. But a myriad of other signals will be employed at any moment in time in a cell so the overall response, or set of responses needs to be orchestrated by the phasing and overall shifting pattern of signals being employed by the cell.

15. Setting up for the future

In music there are often times when the phrasing and harmony just does not sound quite right. This is usually very transient and the harmonies resolve very quickly. Perhaps a composer has asked for a F and a G to be played together. If they were the adjacent notes the resulting discord would be very obvious, but often such notes are played with two or three octaves between them – in those cases the discord is not so blatant. However, often that harsh nature of the discord will mean that when the music does harmonise the end result is more pleasing than it would have been without the disharmony. The composer has set the scene for the final resolution. Again using Beethoven as an example, he does this to great effect in the "*Moonlight Sonata*", where the listener is treated to slight disharmony and one is waiting in anticipation for the resolution, which when it comes is delightful. It brings depth and feeling to the work. Therefore the composer is setting up for the future, ensuring that what subsequently arrives results in a success.

Cells need constantly to be setting the scene and making sure that they are ready for the future. And of course cell signalling is the key to doing this. Signalling often leads to adaptation, where the cell sets itself up for future possible events (Neill *et al.*, 2002). Music often sets the scene in a strident and discordant, or stressful, way to allow for future harmony. In cells exposure to one stress can lead to cells being able to cope better with subsequent stress in the future, and not only to the same stress. Temperature stress in plants for example can lead to adaption to future stress by other abiotic and biotic stress factors. Instead of viewing sub-lethal stress as a negative thing in cell signalling perhaps we should be more like the composer who is prepared to chose a discord to ensure the future has a better outcome. The composer is adapting our ear, just as cell signalling is adapting the functioning of the cell for future events.

In cell signalling adaptation and preparing for the future may require long term activity, and will no doubt involve the control of gene expression with an alteration of the complement of proteins in the cell. Perhaps the cell will alter its levels of certain receptors or signalling proteins, and this will enable the cell to have a faster or more tailored response in the future. This would be like a disc jockey in the night club being asked for a certain genre of music, but realising that he didn't have it. There would be a period of short term stress. Instead of being caught out in the future, a trip to the music shop would ensure that he would be "adapted" for future requests, so lessening the chance of stress, and also allowing alternatives to be played which may relieve an otherwise stressful situation. Our cell signalling would "take a trip to our genome" and so ensure that they are ready for future "requests" from their environment.

16. Signalling dysfunction

Dysfunction of cell signalling can have catastrophic consequences. This is certainly one of the main reasons why the topic needs to be more fully understood. Dysfunction can lead to either the lack of functionality or indeed too much activity, with either situation being undesirable. If insulin signalling is taken as an example, a dysfunction of the insulin receptor would mean that the arrival of insulin at the cell surface would not be recognised and no insulin response by the cell would be mounted. Obvious effects of such a dysfunction are conditions such as diabetes. One the other hand, if the G protein Ras is taken as an example, mutation of the coding for amino acids at position 12, 13, or 61 in the sequence leads to a protein which has impaired GTPase activity and therefore can not be turned off (Figure 8). This leads to the signalling pathways in which RAS is involved being in the permanently active state, regardless of the lack of continued initiation of the signalling pathway. *RAS* mutations are found in about one third of human malignancies (Riely *et al.*, 2009). These are just two examples and there are many more, highlighting the importance and impact of correct cell signalling.

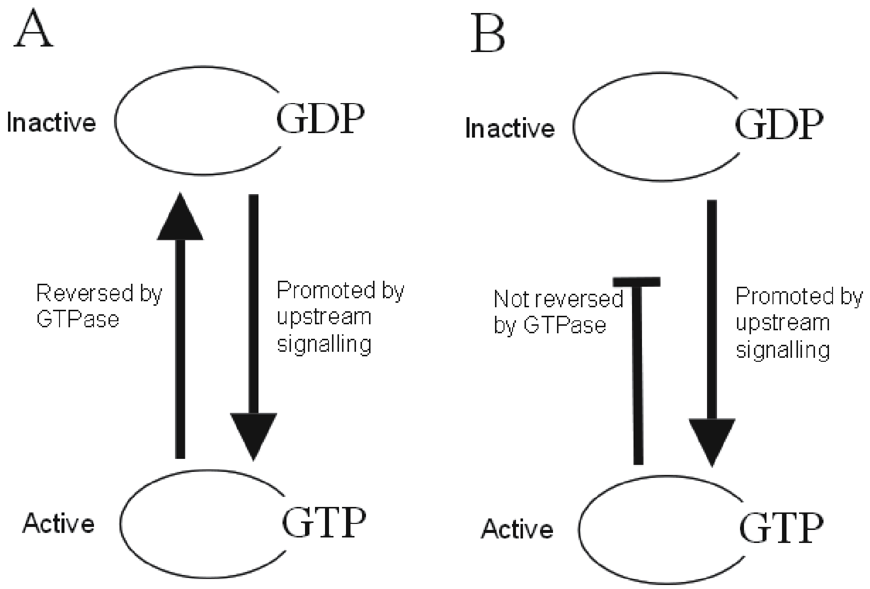


Fig. 8. G protein dysfunction can lead to continuous signalling. G proteins can be thought of as being molecular switches, with an "on" and an "off" state. In the inactive state they are bound to GDP, but on activation this is exchanged for GTP, so leading to a conformational change in the protein which allows it to signal (A). The GTP bound form of the G protein will then signal downstream to the next effector in the chain. In the case of the G protein Ras, the next signalling component in the transduction pathway could be the protein kinase Raf. To then inactivate the G protein its intrinsic GTPase activity will convert the GTP back to GDP and inactive the protein – through a reversal of the conformation change. However, in proteins such as Ras, a mutation can disrupt the GTPase active site so stopping the conversion of GTP back to GDP (B). In this case, the protein will continue to be bound to GTP, and continue to be in the conformation that signals. Therefore, even if all the up-stream signalling is reversed or halted, the G protein will continue to signal downstream regardless. Because Ras is often on pathways which are invoked by growth factors, continuous G protein signalling can lead to a continuous "grow" signal, and hence lead to tumour growth and cancer.

Dysfunction can and does happen in music too. The discussion above emphasises the fact that many cell signalling events will be taking place in the cell at the same time. There may be many effects, often in different parts of the cell. However, there should be minimal interference of one pathway over the other if they are controlling completely independent effects. At many music venues in recent times there may be several events taking place all at the same time, and the Glastonbury Festival is a good example. People need to listen to the band of their choice without hearing the others in such a way that it disrupts their enjoyment. However, if one band starts to dominate, or the equipment on one band loses its volume, then the effect that the festival envisaged will be compromised, just as the overall signalling network of the cell would be compromised. An example of this would be signals moving into a cell through the gap junctions. The signalling of one cell may overwhelm the signalling in one of its neighbours if second messengers move on mass through the gap junctions. Clearly there needs to be control of such movement and gap junction function and regulation is clearly important to understand (Evans and Martin, 2002).

Although it is less common now physical recordings can cause problems with both vinyl records and CDs able to "jump". This can render the music so bad that it can't be listened too. Other equipment can fail too, including digital instruments and amplifiers. But even more classical equipment can have problems. A sonata played on a piano with a broken key or hammer may make the music very poor. It may not stop the piece being played altogether, and the musician may be able to continue. However, the concert is unlikely to get good reviews and both the musician and venue may struggle to have a future event. Just like in a cell, a small dysfunction may render the longer term future to be in doubt.

17. Future and evolution

It would be naïve to think that cell signalling has evolved to the point where it will evolve no further. Organisms continue to evolve and the proteins involved in cell signalling will no doubt evolve too, and certainly will not stay the same for eternity. Likewise it would be foolish to think the same about music and instruments.

Over billions of years since life began, cell signalling proteins have mutated and changed to give the polypeptides that we can find today. Since the creation of the first cells some form of signalling was required, both to sense the environment of the cell and to coordinate adaption as the world changed. From an oxygen free atmosphere to the present climate of Earth cells and their signalling have had to adapt along the way. They will continue to change, adapt and no doubt the proteins involved in signalling will increase in number in the future. Perhaps an example of such a change can be seen with the enzyme nitric oxide synthase (NOS). Very recently a NOS has been characterised from a very primitive green algae *Ostreococcus tauri* (Foresi *et al.*, 2010). Perhaps this photosynthetic organism inherited its gene for this enzyme from a more primitive cell, one which gave rise to both plants and animals. This is likely as the *O. tauri* amino acid sequence is 45% similar to that of a human gene for NOS. Therefore, the human gene has changed considerably compared to the *O. tauri* gene over time. Perhaps more striking is the fact that although *O. tauri* is a primitive plant, higher plants do not seem to have a form of this NOS gene at all. Either it has been lost altogether, or it has been mutated to a form which has yet to be identified. Either way, evolution has been hard at work on this gene, and will no doubt continue such work into the future.

Mutation and duplication of gene can lead to families of proteins. Certainly in signalling families of protein isoforms can be recognised, some with added domains, some with extra phosphorylation sites and some in truncated form. Good examples are phosphatase proteins that remove the phosphate groups from proteins (Cohen *et al.,* 1990).

Musical instruments also evolve, and in many ways in a similar manner to proteins. Some have certainly been around for a long time, but even those that form part of current orchestrates are different from those used by great composers such as Mozart. Furthermore he would never have imagined the possibility of an electric violin, but today his music is often played on such an instrument. Music itself evolves, with successive composers building on the work of those who went before them.

Instruments have changed over the years in a way that resembles that of proteins. Protein isoforms can be created when a gene is copied so there are two versions, and then those genes mutate after a period of time to two separate genes which are able to be characterised, and they would give rise to different but related proteins. Musical instruments are the same. A violin is like a copy of a viola, except one is bigger and plays different notes. Copy it again and make it bigger still and a cello is created, and so on so there is a family of instruments which are recognisable as being related, and yet they have different roles. They could be thought of as isoforms perhaps, just like proteins. Using the piano as an example and again one can see "isoforms" which are all recognisable as pianos, that is the concert grand, the baby grand, the upright, the studio piano and so on. Over the years the piano has been adapted to the place it needs to be placed and the audience it is aimed at. There have formats which are no longer seen, like genes which have disappeared during evolution, and there are new versions being developed and used. There now seems to be a vast array of electric pianos and electronic keyboards, with the idea of using a piano keyboard layout being copied and mutated to develop new instruments.

Music and the instruments used to create it will continue to develop and evolve, just as the proteins which are involved in cell signalling. Especially in the advance of climate change, organisms will need to adapt and it will be cell signalling which coordinates such changes, but the signalling pathways will change too. The future will see the development of new cellular components, and no doubt new musical instruments, especially as new digital technologies are adopted. Not all changes will be beneficial, with mutations in genes no only allowing the future evolution of species but creating dysfunctional proteins along the way causing disease in individuals. No doubt not all new music innovations will be successful either, and the future will be littered with new proteins and musical instruments abandoned by nature and the music industry respectively.

18. Conclusion

Cell signalling is both enormously important to the understanding of how cells works and immensely complicated. Therefore ideas which can be used to aid in the teaching and study of the subject would be extremely helpful, and an analogy is often a good tool. Music initiates cell signalling events in organisms, but music terminologies are often used to explain aspects of cell signalling. However, music as an analogy for cell signalling events can be an interesting and useful way to look at the principles of signalling and transduction pathways. Such an analogy will be useful to those teaching and studying cell signalling.

19. Acknowledgements

I would like to thank Annabel Hancock for supplying the music used in Figure 6.

20. References

Agosta, W.C. (1992) *Chemical communications: the language of pheromones*. Scientific American Library, New York. USA,.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J.D. (1994) *Molecular Biology of the Cell* 3rd edn. Garland Press, New York.

Bleeke, T., Zhang, H., Madamanchi, N., Patterson, C. & Faber, J.E. (2004) Catecholamine-induced vascular wall growth is dependent on generation of reactive oxygen species. *Circ. Res.,* Vol. 94, pp. 37–45.

Chikahisa, S., Sei, H., Morishma, M., Sano, A., Kitaoka, K., Nakaya, Y. & Morita, Y. (2006) Exposure to music in the perinatal period enhances learning performance and alters BDNF/TrkB signalling in mice as adults. *Behav. Brain Res.*, Vol. 169, pp. 312–9.

Cohen, P.T.W., Brewis, N.D., Hughes, V. & Mann, D.J. (1990) Protein serine/threonine phosphatases: an expanding family. *FEBS Lett.,* Vol. 268, pp. 355–359.

Colucci-Guyon, E., Portier, M.M., Dunia, I., Paulin, D., Pournin, S. & Babinet, C. (1994) Mice lacking vimentin develop and reproduce without an obvious phenotype. *Cell*, Vol. 79, pp.679-694.

Evans, W.H. & Martin, P.E.M. (2002) Gap junctions: structure and function (Review). *Molecular Membrane Biology,* Vol. 19, No. 2, pp. 121-136.

Dawkins, R. *The Selfish Gene* (1989) 2nd edn. Oxford University Press, Oxford, UK

Filmore, D. (2004) It's a GPCR world, cell-based screening assays and structural studies are fueling G-protein-coupled receptors as one of the most important classes of investigational drug targets. *Modern Drug Discovery,* Vol. 7, pp. 24–28.

Foresi, N., Correa-Aragunde, N., Parisi, G., Caló, G., Salerno G. & Lamattina L. (2010) Characterisation of a nitric oxide synthase from the plant kingdom: NO generation from the green algae *Ostreococcus tauri* is light irradiance and growth phase dependent. *The Plant Cell*, Vol. 22, pp. 3816-3830.

Fung, B. K.-K. & Stryer, L. (1980) Photolyzed rhodopsin catalyses the exchange of GTP for bound GDP in retinal rod outer segments. *Proc. Nat. Acad. Sci. USA*, Vol. 77, pp. 2500-2504.

Jenkins, J.S. (2001) The Mozart effect. *J. R. Soc. Med.,* Vol. 94, pp. 170-172.

Hancock, J.T. (2009) Cell signalling is the music of life. *Brit. J. Biomed. Sci.*, Vol. 65, pp. 205-208.

Hancock, J.T. (2010) *Cell Signalling*. 3rd edn. Oxford Univeristy Press, Oxford, UK.

Hancock, J.T. (2005) *Cell signalling*. 2nd edn. Oxford University Press, Oxford UK.

Hancock, J.T. (2003) Principles of cell signalling. In *On growth, form and computers*. Kumar, S., Bentley, P.J. eds. pp. 64–81, Academic Press, Oxford.

Jeffrey, C.J. (2009) Moonlighting proteins—an update. *Mol. BioSyst.*, Vol. 5, pp. 345-350.

Jiang, X. & Wang, X. (2004) Cytochrome C-mediated apoptosis. *Annu. Rev. Biochem*., Vol. 73, pp. 87-106.

Kadenbach, B., Ramzan, R. & Vogt, S. (2009) Degenerative diseases, oxidative stress and cytochrome c oxidase function. *Trends in Molecular Medicine*, Vol. 15, No.4, pp. 139-147.

Kell, D.B., Kaprelyants, A.S. & Grafen, A. (1995) Pheromones, social behaviour and the functions of secondary metabolism in bacteria. *Trends Ecol. Evol.* Vol. 10, pp. 126–9.

Krauss, G. (2008) *Biochemistry of Signal Transduction and Regulation*. Wiley-VCH, Chichester, UK.

Lambeth, J.D., Kawahara, T. & Diebold, B. (2007) Regulation of Nox and Duox enzymatic activity and expression. *Free Radic. Biol. Med.*, Vol. 43, pp. 319-331.

Lane, N. (2010) *Life Ascending: The Ten Great Inventions of Evolution*. Profile Books Ltd, London, UK.

Lewitt-Bentley, A. & Rety, S. (2000) EF-hand calcium-binding proteins. *Current Opinion in Structural Biology*, Vol. 10, pp. 637-643.

Neill, S.J., Desikan, R., Clarke, A., Hurst, R. & Hancock JT. (2002) Hydrogen peroxide and nitric oxide as signalling molecules in plants. *J. Exp. Bot.*, Vol. 53, pp. 1237–1247.

Nordquist, P.R. & Ayers, R.D. (2009) Tuning and tone quality of bagpipe drones. *Acoust. Soc. Am.*, Vol. 125, pp. 2652-2652.

Oxford Dictionaries. http://oxforddictionaries.com/

Pereyra, E., Mizyrycki, C. & Moreno, S. (2000) Threshold level of protein kinase A activity and polarized growth in *Mucor rouxii*. *Microbiology*, Vol. 146, pp. 1949-1958.

Polo, S. & de Fiore, P.P. (2008) Endocytosis conducts the cell signalling orchestra. *Cell,* Vol. 124, pp. 897–900.

Qin, Y.C., Lee, W.C., Choi, Y.C. & Kim, T.W. (2003) Biochemical and physiological changes in plants as a result of different sonic exposures. *Ultrasonics*, Vol. 41, pp. 407-411.

Rachmilewitz, J. & Lanzavecchia, A. (2002) A temporal and spatial summation model for T-cell activation: signal integration and antigen decoding. *Trends Immunol.*, Vol 23, pp. 592-595.

Riely, G.J., Marks, J. & Pao, W. (2009) *KRAS* mutations in Non–Small Cell Lung cancer. *The Proceedings of the American Thoracic Society*, Vol. 6, pp. 201-205.

Reya, T. & Clevers, H. (2005) Wnt signalling in stem cells and cancer. *Nature*, Vol. 434, pp. 843-850.

Rorsman, P., Eliasson, L., Renström, E., Gromada, J., Barg, S. & Göpel, S. (2000) The cell physiology of biphasic insulin secretion. *News Physiol. Sci.,* Vol 15, pp. 72–77.

Schafer, F.Q. & Buettner, G.R. (2001) Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radic. Biol. Med.* Vol. 30, No. 11, pp. 1191-1212.

Trappe, H.J. (2010) The effects of music on the cardiovascular system and cardiovascular health, *Heart*, Vol. 96, pp. 1868-1871.

Tristan, C., Shahani, N., Sedlak, T.W. & Sawa, A. (2011) The diverse functions of GAPDH: views from different subcellular compartments. *Cell Signal*. Vol. 23, pp. 317-323.

Wan, X., Steudle, E. & Hartung, W. (2004) Gating of water channels (aquaporins) in cortical cells of young corn roots by mechanical stimuli (pressure pulses): effects of ABA and of $HgCl_2$. *J. Exp. Bot.* Vol. 55, pp. 411–422.

Zhang, R., Tsai, F.Y. & Orkin, S.H. (1994) Hematopoietic development of vav-/- mouse embryonic stem cells. *Proc. Nat. Acad. Sci. USA*, Vol. 91, pp. 12755-12759.

3

Chemical Carcinogenesis: Risk Factors, Early Detection and Biomedical Engineering

John I. Anetor[1], Gloria O. Anetor[2], Segun Adeola[1] and Ijeoma Esiaba[1]
[1]Department of Chemical Pathology, College of Medicine, University of Ibadan,
[2]Department of Human Kinetics and Health Education, Faculty of Education,
University of Ibadan,
Nigeria

1. Introduction

Cancer is now recognized in both humans and in other multicellular animals as arising from a number of different causes, including specialized viruses, radiation, chemicals, certain highly irritative parasites (inflammation) and a number of other factors, such as specific genetic defects present in individual humans and possibly in every member of a colony of specially bred animal models. Cancer from non genetic causes largely from environmental factors, of which chemicals have a disproportionate share, is believed to contribute nearly 70% of all cancer cases. Chemical carcinogenesis originally derives from experimental induction of malignant skin tumor in mice with chemicals. Early studies indicated some agents such as polycyclic aromatic hydrocarbons (PAH) could cause cancer of the skin if they were painted on to mice in high doses. These early studies also showed that the induction of cancer was dose dependent; in low dosage they would not cause cancer but would render the skin susceptible to developing cancer on exposure to another agent, which, on its own would not induce cancer.

Thus at the dawn of the 20th century, it was recognized that chemicals cause cancer; though individual cancer causing molecules had not yet been identified, nor their cellular targets clearly known. It was however clearly understood that carcinogenesis, at the cellular level, was predominantly an irreversible process. Knowledge of the mechanisms by which chemicals cause cancer and the molecular changes that characterize tumor progression was lacking. The origin of the understanding that cancer had a cause was first pointed out by the Italian investigator, Ramazini in 1700. Seven and a half decades later, the British Surgeon, Percival Pott made the connection between exposure to soot, rich in hydrocarbons and scrotal cancer (Pott, 1775). It is now known that at the most fundamental level, cancer is caused by abnormal gene expression. This abnormal gene expression occurs through a number of mechanisms, including direct damage to the DNA, and inappropriate transcription and translation of cellular genes. The contribution of chemicals to the carcinogenic process is well known to have increased given the parallel between industrialization with associated increased chemical production and utilization and the prevalence of cancer.

Increasing use of chemicals, particularly in the industrializing developing countries (Pakin et al., 1993; Pearce et al., 1994) places new demands on these countries, as they have limited resources to adequately regulate exposure to chemicals. Majority of the chemicals cause mutation in DNA among others. The consequences of increased exposure to chemicals, risk of cancer, early detection of chemical-induced neoplastic changes and the prominent role of biomedical engineering is poorly recognized generally and particularly in the developing countries where chemical carcinogenesis is believed to be currently more prevalent (Huff and Rall,1992). Cancer is classically viewed as the result of series of mutations, including dominantly acting oncogenes and recessively acting tumor- suppressor genes. Each mutation leads to the selective overgrowth of a monoclonal population of tumor cells, and each significant tumor property (invasiveness, metastasis and drug resistance) is accounted for by such mutation (figure 1). The seminal observation that carcinogenesis is a multistage process helps to explain why some chemical carcinogens lack apparently important properties exhibited by others. Such agents may act as promoters on tissues that have been previously initiated or have the ability to produce naturally occurring tumors without treatment.

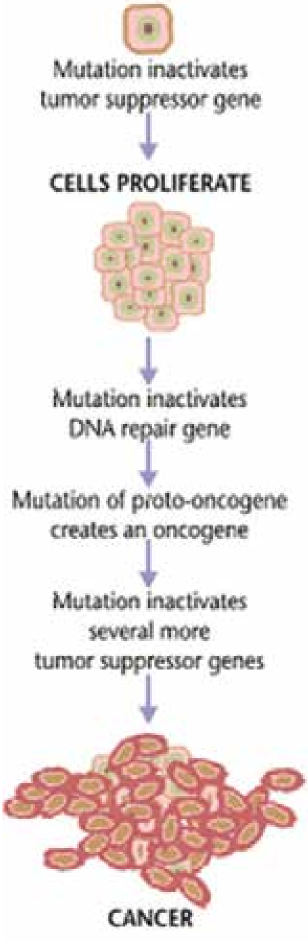

Mutation inactivates tumor suppressor gene
CELLS PROLIFERATE
Mutation inactivates DNA repair gene
Mutation of proto-oncogene creates an oncogene
Mutation inactivates several more tumor suppressor genes
CANCER

Fig. 1. Stages of the carcinogenic process

DNA-reactive	Activation-independent	Nitrogen mustards, chlorambucil
		Alkylating agents
		Epoxides: ethylene oxide
	Activation-dependent	Aliphatic halides: vinyl chloride
		Aromatic amines: monocyclic-o-toludine; polycyclic-4-aminobiphenyl, benzidine
		Nitroaromatic compounds: 1-nitropyrene, 3-nitrofluoranthene
		Heterocyclic amines: 2-amino-3-methylimidazo [4,5-b]pyridine(Ph1P)
		Aminoazo dyes: dimethylaminoazobenzene
		Polycyclic aromatic hydrocarbons: benzo(a)pyrene
		Substituted polycyclic aromatic hydrocarbons: 3-methlycholanthrene
		N-nitroso compounds: dialkyl-dimethylnitrosamine, diethylnitrosamine; cyclic-N-nitrosonornicotine (NNK), nitrosomorphaline triazines, hydrazines, azoxymethane, methylazoxymethanol, benzene
		Mycotoxins: aflatoxin B1, aflatoxin G1
		Plant products: pyrrolizidine alkaloids, aristolochic acid, cycasin
		Pharmaceuticals: cyclophosphamide, phenacetin, tamoxifen
Epigenetic	Promoter	Liver enzyme-inducer type hepatocarcinogens: chlordane, DDT, pentachlorophenol, phenobarbital, polybrominated biphenyls, polychlorinated biphenyls
		Kidney: nitrilotriacetic acid
		Bladder: sodium saccharin
		Forestomach-butylated hydroxyanisole
	Endocrine-modifier	Hormones: estrogens-17-estradiol; catechol estrogens-4-hydroxyestradiol, 2-hydroxyestradiol
		Estrogen agonists: 17-ethinyl estradiol, diethylstillbestrol (DES)
		Prolactin inducers: chloro-s-triazines-atrazine
		Antiandrogens: finasteride, vinclozolin
		Antithyroid thyroid tumor enhancers
		Thyroperoxidase inhibitors: amitrole, sulfamethazine
		Thyroid hormone conjugation enhancers: phenobarbital,
		Gastrin-elevating inducers of gastric neuroendocrine and glandular tumors: lansoprazole,omeprazole, pantoprazole
	Immunosuppressor	Cyclosporin
		Purine analogs
	Cytotoxin	Mouse forestomach toxicants: propionic acid, diallyl phthalate, ethyl acrylate
		Rat nasal toxicants: chloracetanilide herbicides
		Rat renal toxicants: potassium bromate, nitrilotriacetic acid
		Male rat kidney _2u-globulin nephropathy inducers: D-limonene, p-dichlorobenzene
	Peroxisome proliferator	Hypolipidemic fibrates: ciprofibrate, clofibrate, gemfibrozil
		Phthalates: di(2-ethylhexyl)phthalate(DEHP), di(isononyl)phthalate
		Lactofen
Minerals and Metals		Minerals: asbestos
		Metals: arsenic, beryllium, cadmium, chromium (IV), nickel, silica
Unclassified		Acrylamide, acrylonitrile, dioxane

Table 1. Classification of chemicals with carcinogenic activity

Foulds (1969) suggested that cancer development consisted of three, rather than two processes: (1) initiation, or the conversion of normal cells to a potentially precancerous form (2) promotion, or the expansion of the clones of initiated cells to form tumors; and (3) progression, or the development of tumors to increasing levels of malignancy. The original view was based on Foulds' wealth of experience with both clinical and experimental cancer. It has however, been expanded greatly since it was first propounded by Foulds (1982). Other investigators have also made significant contributions to the understanding of the process of carcinogenesis by suggesting that there are two major cell-based processes essential to the formation of tumors (Ames and Gold, 1981).The first, or initiating stage, is due to mutation; alteration of the DNA of the affected cell through permanent modification of the DNA. These mutations take place at specific locations on the DNA, referred to as oncogenes and tumor suppressor genes, if these individual cells are to serve as precursors of cancer (Willis, 1960; Klein and Klein, 1984). This area remains intensely investigated in the last couple of decades. What is perhaps worthy of note is that while the activation of an oncogene requires mutation at a specific single base (arrangement of the amines making up the DNA) pair on the DNA template, inhibition of a tumor suppressor gene may be achieved by a much wider range of damaging interactions.

In current research, emphasis is laid on the identification of the genes that are involved in the mutation and subsequent molecular events. The failure in the control mechanisms regulating the expression of and response to tissue growth factors is of considerable interest in chemical carcinogenesis. This contributes to the risk of chemical carcinogenesis and is in turn attributable to a number of factors that will be discussed subsequently. A critical process in carcinogenesis is promotion. This involves cellular proliferation, which involves the division of cells to form two unusually identical cells. This may increase the number of both "normal" and neoplastic mutated or preneoplastic cells, enhancing the chance of a tumor being expressed in a clinically observable form. Surprisingly, such increased levels of cellular proliferation may not be apparent in normal cells of a particular tissue but may occur only in pretumor cells thus making early detection difficult. Tumors are well known to increase in their degree of malignancy with time, a process named "progression" by Foulds. Cohen and Elwein (1991) have suggested that progression is the result of a cascade of further critical mutations in the neoplastic cell population followed by further cell proliferation to increase the number of genetically altered cells and the chance of their forming an increasingly malignant, clinically apparent cancer.

2. Summary of stages of chemical carcinogenesis

Studies indicate that three stages of chemical carcinogenesis can be defined; initiation, promotion and progression as briefly alluded previously.

Initiation: This is concerned with the induction of genetic changes in cells, leading to genome instability. This can be accentuated by micronutrient deficiency disorders. The nature of the initial changes is still incompletely elucidated (Satoh, 1988)

Promotion: This largely involves the induction or commencement of cell proliferation. In this phase of carcinogenesis a promoting agent or enabling microenvironment, brings about increased cell proliferation. This stage is very important as it is reversible if the promoting agents or risk factor(s) are withdrawn. Probably also, if some genome stabilizing micronutrients are abundant. This stage has been exploited considerably for both therapeutic and chemopreventive measures that are in part dependent on micronutrients.

Progression: If cell proliferation is sustained then initiated cells acquire secondary genetic abnormalities in oncogenes which first lead to dysregulation and finally to autonomous growth characteristic of cancer. The ultimate end-point of progression is development of invasive neoplasm.

While many environmental agents can be considered to be chemical carcinogens; some act as both initiators and promoter (complete carcinogens). The understanding of molecular aspects of chemical carcinogenesis has lead to development of the concept of chemoprevention (anticarcinogenesis). This process proposes strategies for intervention at the phase of malignancy using drugs or natural or synthetic agents to reverse or halt the evolution of carcinogenesis which is dependent on recognition of risk factors and early detection.

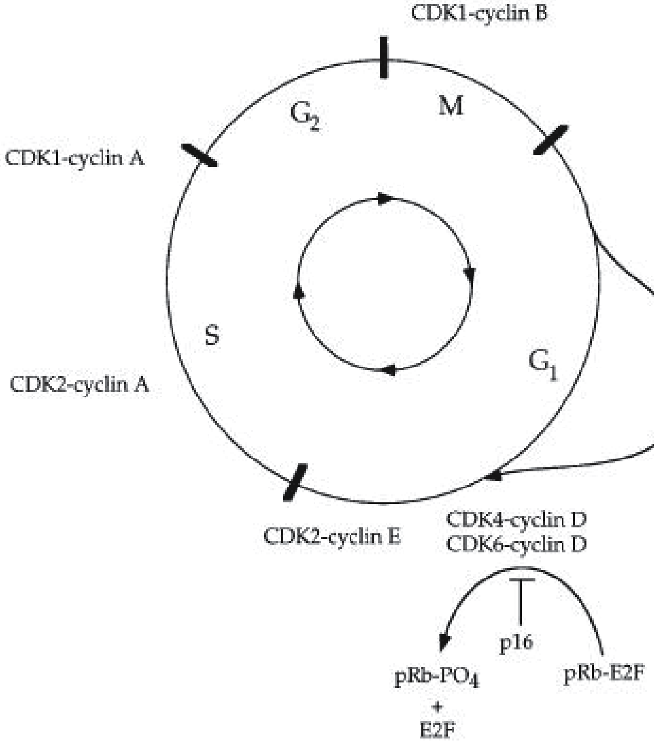
2.1 Brief history of chemical carcinogenesis

The history of chemical carcinogenesis is punctuated by key epidemiologic observations and animal experiments that identified cancer-causing chemicals and that led to increasingly insightful experiments to establish molecular mechanisms and to reduction of human exposure to chemicals. In 1914, Boveri (1914) made key observations of chromosomal changes, including aneuploidy. His analysis of mitosis in frog cells and his extrapolation to human cancer is an early example of a basic research finding generating an important hypothesis (the somatic mutation hypothesis). The first experimental induction of cancer in rabbits exposed to coal tar was performed in Japan by Yamagiwa and Ichikawa (1918) and was a confirmation of Pott's epidemiologic observation of scrotal cancer in chimney sweeps in the previous century (Potts, 1775). Owing to the fact that coal tar is a complex mixture of chemicals, a search for specific chemical carcinogens was undertaken. British chemists, including Kennaway (1930), took on this challenge and identified polycyclic aromatic hydrocarbons (PAHs), such as, benzopyrene, which was shown to be carcinogenic in mouse skin by Cook and his colleagues in 1933. The fact that benzopyrene and many other carcinogens were polyaromatic hydrocarbons lead the Millers (1947) to postulate and verify that many chemical carcinogens required activation to electrophiles (electron seeking moieties) to form covalent adducts with cellular macromolecules. This in turn prompted Conney and the Millers (1956) to identify microsomal enzymes (P450s) that activated many drugs and chemical carcinogens.

3. Cell regulatory mechanisms: The cell cycle

Carcinogenesis, or the sequence of events leading to cancer, is a multistep process involving both intrinsic and extrinsic factors. In the normal tissue, there are numerous regulatory signals that instruct cells when to replicate and when to die. In a cancer cell these regulatory mechanisms become disabled and the cell is allowed to grow and replicate unchecked. Thus at the most fundamental level cancer is caused by abnormal gene expression. This abnormal gene expression occurs through a number of mechanisms including direct damage to the DNA, and inappropriate transcription and translation of cellular genes. Carcinogenesis has been demonstrated abundantly to be induced or at least caused by exposure to certain types of chemicals (carcinogens). The mechanisms are elaborated on subsequently. The cell cycle plays an important role in this regard. It is concerned with the processes that govern the life and death of cells and through transient delay in G_0 phase or outright apoptosis (programmed cell death) might be able to prevent damage in the DNA of a cell that may proceed to carcinogenesis. In the normal cell, replication of the DNA and cell division is stimulated by the presence of growth factors that bind receptors at the cytoplasmic membrane and initiate a cascade of intracellular signals. Once these signals reach the nucleus they induce transcription of a complex array of genes producing proteins that mediate progression of the cell cycle culminating in mitosis or cell division. One remarkable contribution of biomedical engineering is the introduction of flow cytometer equipment which enable stages of the cell cycle to be followed and disorders or disruptions there of detected.

The cell cycle is conventionally divided into four (4) phases, although there is the subsidiary G_0 phase. The duration of each of these phases varies depending on factors such as cell type and localized conditions within a given tissue (microenvironment). At the end of mitosis (M) daughter cells enter gap 1 (G_1) phase.



A schematic representation of the mammalian cell cycle. In each cell division cycle, chromosomes are replicated once (DNA synthesis or S-phase) and segregated to create two genetically identical daughter cells (mitosis or M-phase). These events are spaced by intervals of growth and reorganization (gap phases G_1 and G_2). Cells can stop cycling after division, entering a state of quiescence (G_0). Commitment to traverse an entire cycle is made in late G_1. Progress through the cycle is accomplished in part by the regulated activity of numerous CDK–cyclin complexes, indicated here and described in the text.

Fig. 2. Schematic representation of Cell Cycle.

If conditions are favourable cells enter the synthetic (S) phase of the cycle where the entire genome is replicated during DNA synthesis. Following 'S' phase, cells enter the gap 2 (G_2) phase before proceeding through mitosis again. A critical phase boundary exists early in the G_1 phase called the restriction point. This is the point at which the cell must decide to either enter the cell cycle once more or to secondly move into a state of quiescence; G_0 phase. Once committed to this pathway, the cell can either remain in this state of replicative quiescence until it receives a signal to divide again. Alternatively the cell can proceed down a path that leads either to terminal differentiation or to apoptosis. Movement of a cell through the cell cycle is regulated by an enormously complex array of proteins.

The proteins include:

- Cyclins
- Cyclin dependent kinases (CDKs)

- Cyclin activating kinases (CAKs)

- CDK inhibitory proteins.

Binding of an appropriate growth factor at the cell surface starts a signaling cascade that ultimately leads to the expression of the G1 phase cyclins. It is important to remark that in normal cells, external stimuli (factors) such as growth factors are absolutely needed for the cell to proceed beyond the restriction point. Beyond this point, the cell is committed to DNA replication and cell division. Interference with the normal signal transduction pathway by chemical carcinogens, the mechanism notwithstanding can transform a cell into a state of proliferation that is not regulated by normal physiological controls (carcinogenesis). This basically is broadly the molecular basis of carcinogenesis.

Table 2. Proteins Controlling the Cell Cycle.

It is note worthy that all phases of the cell cycle are regulated by the micronutrient zinc. Thus zinc deficiency common in many developing countries (WHO, 2002; Ames, 2010) can be risk factors in chemical carcinogenesis. Ho et al (2003) have elegantly demonstrated this in their studies. This is an area where biomedical engineering has contributed significant in the last five or more decades by the production of flame absorption spectrophotometers (FAAS) and later the graphite furnace (carbon rod) (GFAAS). This was followed by inductively coupled plasma mass spectrometer which allows for simultaneous multi element analysis. These equipment have exquisite sensitivities which enable the status of zinc and many other micronutrients to be detected and indirectly play a preventive role; reducing risk of cancer.

4. Molecular biology and chemical carcinogenesis

The discovery of DNA as the genetic material by Avery, MacLeod, and McCarthy (1944) and the description of the structure of DNA by Watson and Crick (1953) indicated that DNA was the cellular target for activated chemical carcinogens and that mutations (alterations in the sequence of bases making up amino acids) were key to understanding mechanisms of cancer. This led to defining the structure of the principal adducts in DNA (complexes of

DNA and a carcinogen or its metabolite) by benzo (*a*) pyrene (Carrel et al., 1997) and aflatoxin B_1 (Croy et al., 1978). The concepts developed in investigating mechanisms of chemical carcinogenesis also led to discoveries that are relevant to other human conditions in addition to cancer, including atherosclerosis, cirrhosis, and aging. The fact that genetic changes in individual cancer cells are essentially irreversible and that malignant changes are transmitted from one generation of cells to another strongly points to DNA as the critical cellular target modified by environmental chemicals. DNA damage by chemicals occurs randomly; the phenotypes of associated carcinogenic changes are determined by selection.

Epidemiologic studies from all over the world have identified environmental and occupational chemicals as potential carcinogens. The most definitive epidemiologic studies have been those in which a small group is exposed to a tremendously large amount of a specific chemical, such as aniline dyes. The table below (table 3) lists some of the fairly well characterized chemicals sites where they have induced cancer.

Carcinogens		**Site of cancer**
Chemical mixtures	Soots, tars, oils	Skin, lungs
	Cigarette smoke	Lungs
Industrial chemicals	Benzidine	Urinary bladder
	Nickel compound	Lungs, nasal sinuses
	Arsenic	Skin, Lungs
	Vinyl choride	Liver
Drugs	Mustard gas	Lungs
	Phenacetin	Renal pelvis
Naturally occuring compounds	Cyclomates	Bladder
	Nitroso compounds	Oesophagus, Liver, Kidney, stomach

Table 3. List of Chemical Carcinogens.

5. Mechanisms of chemical carcinogenesis

As part of daily existence, DNA frequently sustains damage. If unrepaired, this can lead to mutations that replicate resulting in abnormal and cancerous development. Some biological mechanisms usually inhibit this process. An enzyme 8-oxoguanine DNA glycosylase (OGG1) among others repairs DNA by excising damaged nitrogen bases constituting the DNA. DNA damage may occur through exposure to chemicals present in cigarette smoke, ionizing radiation and oxidative stress, which can be induced by a number of chemicals such as cadmium and the polycyclic aromatic hydrocarbons. The levels of OGG1 can thus be used to predict an individual's risk of developing cancer.

At least four fundamentally different mechanisms of cancer induction by chemicals have been identified. These may lead to cancer as individual processes or on occasion, the same agent may exert its effects through two or more processes to lead to tumor formation. The importance of the mutation/proliferation approach to the development of cancer lies in its ability to encompass each of these mechanisms within a single frame work. This as has been demonstrated (Clayson, 2001), means that if we can measure changes in mutation and proliferation frequencies, due to a specific carcinogen, there may be no need to elucidate the detailed mechanism of carcinogenesis for every chemical carcinogen before attempting to calculate accurately the risk it may carry for exposed human subjects. This requires exquisite and very sensitive instruments, and appears to be one challenge to specialists in biomedical engineering. Fortunately by the wide range of flow cytometry and mass spectrometry based techniques, the field of biomedical engineering appears to be rising to the challenge. It needs not be emphasized that a great deal of thought and effort will be required if mutation rates and proliferation in specific cell types are to be measured in humans by non-invasiveness.

A number of chemical carcinogens now appear to exert their primary effect on the mutational part of the carcinogenic process, while some others seem to be relatively devoid of the ability to interact with DNA and appear to work mainly through a mechanism of induction of cellular proliferation. Mutation on the other hand appears to be induced by chemical carcinogens by at least two major modes. The modes involve direct interaction with the DNA through the formation of highly reactive, positively charged entities known as "electrophiles" This entity is capable of reacting chemically with many different cellular constituents, including the genetic material, DNA (Miller et al, 1961). The adducts formed with DNA the interaction products of such carcinogen-derived electrophiles with DNA, are not regarded as genetic lesions in their own right. They only represent the first stage in the formation of a mutational event. The adducts may be effectively repaired by the DNA repair enzyme system found in the cell nucleus as earlier indicated. In the alternative, if they are not repaired they may affect important sites on the DNA and consequently die, or if the DNA replicates while they are still present, may lead to mutations through base-mispairing or other errors, that is culminating in true genetic lesions. This has been broadly illustrated by the figure below (figure 3).

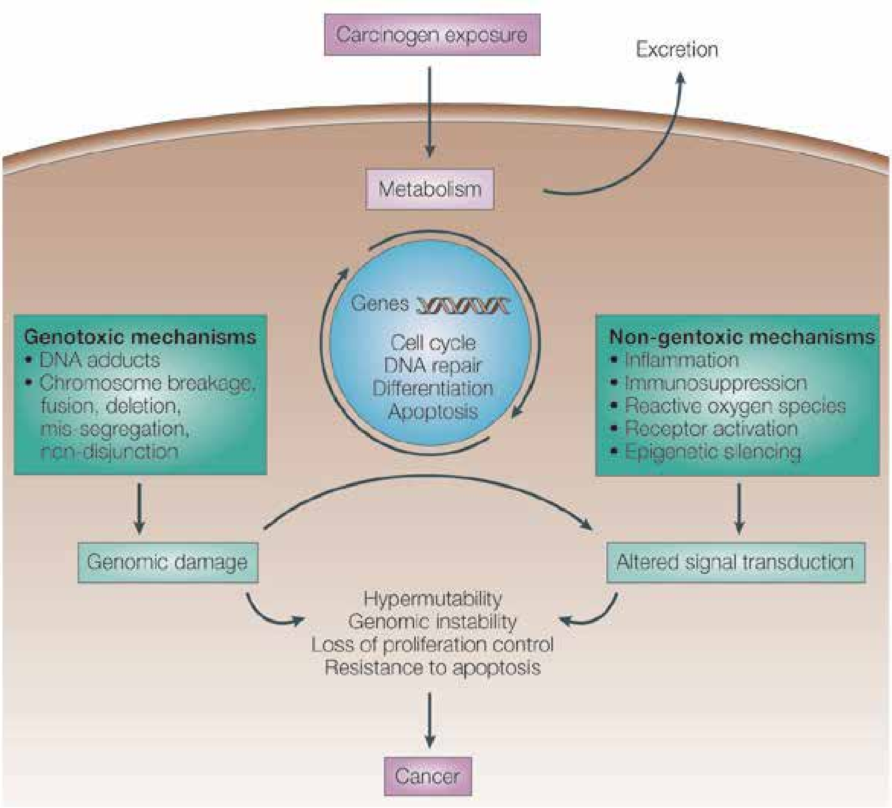

Carcinogen exposure
Excretion
Metabolism
Genes
Cell cycle
DNA repair
Differentiation
Apoptosis
Genotoxic mechanisms
• DNA adducts
• Chromosome breakage, fusion, deletion, mis-segregation, non-disjunction
Non-gentoxic mechanisms
• Inflammation
• Immunosuppression
• Reactive oxygen species
• Receptor activation
• Epigenetic silencing
Genomic damage
Altered signal transduction
Hypermutability
Genomic instability
Loss of proliferation control
Resistance to apoptosis
Cancer

Fig. 3. Overview of genotoxic and non-genotoxic effects of carcinogens.

If the induced mutations occur at one or a relatively few critical sites on the DNA, then the cells may be converted from a "normal" to a preneoplastic state. Chemically-induced mutations are not limited to critical genes, different pretumor cells may demonstrate a variety of different growth potentials due to the range of altered " non-critical" genes, thus enabling those cells with most favourable properties to transform most rapidly to clinically apparent tumours. Alternatively, the carcinogens may act indirectly through the formation of reactive oxygen species or nitrogen radicals, some types of which are also highly reactive with macromolecules such as DNA. The process of raised cellular proliferation is also multi-factorial in its genesis. It may arise from, for instance, direct hormone-like stimulation of specific cell types, from perturbation of tissue processes that lead to a balance between cell proliferation and cell death (apoptosis), it may alternatively arise from massive cell-killing or cytotoxicity followed by proliferative regeneration to maintain the physiological functioning of the affected tissues. A yet further way by which excess, tumor-inducing cellular division may be induced is that exhibited by the urinary bladder. Oyasu and his colleagues (1981) and in his subsequent studies (1995) showed by using heterotopical transplant rat bladder technique that urine by itself, but not water can induce proliferation. The mechanism by which this happens is not quite clear.

It was however conceived that urine contained epithelial growth factors that stimulate cell division and that such factors would penetrate the epithelium should it be injured by the presence of a foreign agent in the bladder. A fourth type of mechanism of carcinogenesis may arise from the ability of the agent to form a complex with a specific protein. This complex (ligand-protein) may have the property of altering the expression of specific and important region in the DNA. There is the emerging but incompletely understood involvement of epigenetics; alteration in the genetic processes not involving DNA base sequence. Many epimutagens have already been identified and a number of existing chemicals such as cadmium are also known to act through this pathway.

6. The environment and cancer

Cancers caused by environmental agents frequently occur in tissues with the greatest surface of exposure to the agents: lung, gastrointestinal tract, and skin. Recently, the study of chemical carcinogenesis has merged with studies on the molecular changes in cancer cells, thus generating biological markers to assess altered metabolic pathways and providing new targets for therapy. Although these are exciting areas, they may be peripheral to attacking the primary causes of the most common human cancers. As more and more mutations are catalogued in cancer cells and more and more changes in transcription regulation, it becomes increasingly apparent that we need to understand what generates these changes. The fact that chemicals cause random changes in our genome immediately implies that our efforts need to be directed to quantifying these changes, reducing exposure, and developing approaches to chemoprevention (Extensively reviewed by Pereira, 1997).

7. Mechanism of oxidative stress and DNA damage due to micronutrient deficiency

Micronutrients are referred to as nutrients required in very small amounts that do not have calorific values but are extremely important for the maintenance of health. They comprise the vitamins and trace elements that regulate vital metabolic and molecular pathways and processes. Some of them play very vital roles in DNA and RNA metabolism either as coenzymes or cofactors involved in their metabolism or as components of systems intimately interacting with these molecular regulators. The micronutrients are basically supplied in the diet or may be taken as supplements. Very few of them are synthesized endogenously enough to meet physiological requirements hence they are mostly essential.

Based on public data emanating from the Healthy People 2010 Project, it has been estimated that about 80% of colon and prostate cancers, may be influenced by diet, nutrition, and life style. It has been proposed that DNA damage induced by dietary micronutrient deficiency accounts for about 33% of preventable cancers (Ames, 2001; Ames and Wakimoto, 2002). Owing to the fact that micronutrient deficiencies can induce DNA damage in a manner similar to those induced by ionizing radiation and reactive oxygen species (ROS), it has been suggested that oxidative stress and the associated DNA breaks are critical targets for nutritional control of carcinogenesis (Cheng, 2009) and perhaps a marker for early detection. When DNA lesions are left unrepaired they can promote accumulation of mutations that facilitate the process of carcinogenesis. Micronutrients may act directly on the genome to prevent mutations, or indirectly as enzyme cofactors in cellular processes that modulate transformation (Hanahan and Weinberg, 2000; Sjoblom et al., 2006). By yet incompletely defined mechanisms, micronutrients at levels higher than nutritional requirements may also activate DNA damage response or senescence, which are processes that are recognised to eliminate cancer cells or limit the progression of precancerous cells (Gorgoulis et al., 2005; Bartkova et al., 2005, 2006). The essential microminerals copper, iron, selenium and zinc play important roles in genome stability. In particular, these microminerals have significant impact on oxidative DNA damage and the corresponding repair pathways.

Micronutrient intakes below recommended levels are known to be unusually widespread in poor countries, though also some segments of the population in economically advanced nations such as the United States, especially among the poor, children, adolescents (Anetor, 2009; Ames, 2010). It has been hypothesised that two of the many insidious but measurable consequences of moderate micronutrient inadequacy are increased DNA damage (precursor of cancer) and mitochondria decay which can cause mutagenic oxidant release also involved in future carcinogenesis. Studies indicate that sensitive assays targeted at these end points have a high probability of detecting changes in individuals with micronutrient deficiencies(Ames, 2010) This again appears instructive to scientists and technologists in the biomedical engineering field.

8. Technology and chemical carcinogenesis

Many technological advances have allowed conceptual ideas to be experimentally tested, including the sensitive detection of chemical carcinogens by high-pressure liquid chromatography (Esaka, et al., 2003) and mass spectrometry (Sigh and Farmer,2006), detection of DNA adducts by postlabeling (Randerath et al, 1981) and by specific antibodies (Poirier et al., 1977), transcriptional profiling by arrays (Kallioniemi et al., 1992, Schena et al., 1995), and quantitation of mutagenicity of carcinogens using bacterial genetics (Ames et al., 1973). The testing of certain concepts in chemical carcinogenesis awaited the development of new technologies. For example, the concept of somatic mutations in cancer preceded by 40 years the establishment of DNA as the genetic material and by 67 years the development of DNA sequencing methods that directly showed clonal mutations in human cancer cells. Also, the mutator phenotype hypothesis formulated in 1974 has been only recently experimentally verified. The list below shows some commonly employed analytical techniques requiring the attention of biomedical engineering and technology.

Some analytical areas that may be particularly relevant in chemical carcinogenesis will include:

- 32P-post labelling
- Fluorescent-based techniques
- Immunoassay-based techniques
- Electrochemical detectors
- Electron microscopy
- High performance liquid chromatography/ Ultra high performance liquid chromatography
- Comet (Single cell gel electrophoresis), imaging system
- Mass spectrometry
- Atomic absorption spectrophotometry (AAS)
- Inductively coupled plasma- mass spectrometry (ICP-MS)
- Advanced spectrophotometry
- DNA microarray systems

The latter is particularly promising as it enables the simultaneous measurement of transcription of thousands of genes using microchips containing thousands of probes of complementary DNA (cDNA) immobilized in predetermined array. But suffers the caveat of being very expensive especially for the developing countries that appear to be mostly in need of it currently.

9. Early studies in chemical carcinogenesis and risk factors

Early in the field of chemical carcinogenesis, investigators recognized that perturbation of the normal microenvironment by physical means, such as wounding of mouse skin or partial hepatectomy in rodents (Hennings and Boutwell, 1970; Fausto etal, 2006) or chemical agents, such as exposure of the mouse skin to certain phorbol esters (Berenblum, 1941), can drive clonal expansion of the initiated cells toward cancer. In the second stage, tumor promotion results in proliferation of the initiated cells to a greater extent than normal cells and enhances the probability of additional genetic damage, including endogenous mutations that accumulate in the expanding population. This classic view of two-stage carcinogenesis (Berenblum, 1941) has been conceptually important but also an oversimplification of the increasing understanding of the multiplicity of biological processes that are deregulated in cancer. In addition, an active debate continues on the relative contribution of procarcinogenic endogenous mechanisms—for example, free-radical–induced DNA damage (Halliwell and Aruoma, 1991), DNA depurination (Lindahl and Nyberg, 1972), DNA polymerase infidelity (Loeb et al, 1974), and deamination of 5-methylycytosine (Lindahl and Nyberg, 1974)—compared with exposure to exogenous environmental carcinogens (Ames et al, 1973).

The enhancement of carcinogens by epigenetic mechanisms such as halogenated organic chemicals and phytoestrogens (Martin et al, 2007), as well as the extrapolation of results from animal bioassays for identifying carcinogens to human cancer risk assessment, are also difficult to quantify (Swenberg et al., 1998). As discussed below, this debate is not merely an academic event, in that societal and regulatory decisions critical to public health are at issue. The identification of chemical carcinogens in the environment and occupational settings [benzo (*a*) pyrene and tobacco-specific nitrosamines in cigarette smoke, aflatoxin B_1 (AFB_1), residues from fossil fuel, vinyl chloride, pesticides and benzene] has led to regulations that have reduced the incidence of cancer. Further reduction or near total elimination may be achieved by sensitive instruments that enable early detection of up- stream changes that may culminate in cancer. This needless to say has heavy reliance on biomedical engineering.

9.1 Risk factors and early detection: role for micronutrient deficiency and oxidative stress

The risk of contracting cancer generally increases as the population grows older; this has been reported to be directly proportional to the number of years raised to the fourth power (Tolonen, 1990). This may be modified in the case of chemical carcinogenesis to include dose of chemicals or environmental agent to which the population is exposed and the stoichiometric bioavailability of protective factors such as the micronutrients. In animal models, such as the mice, the life-span of the models has been increased and cancer prevented by calorie restriction (reducing oxygen intermediates) and feeding them antioxidants. This suggests that excess calorie and antioxidant deficit (oxidative stress) are risk factors that may enhance the carcinogenic process. The natural dietary antioxidants, selenium, zinc, vitamins A, C, and E plus β-carotene protect against free radicals, lipid peroxidation (Tolonen, 1990) and thus the risk of chemical carcinogenesis. Ames (1983) has greatly emphasised the role of antioxidants largely derived from micronutrients as anticarcinogenesis. Vitamin C for instance is well known to counteract carcinogenic nitrous amine in the stomach. Urban population are more exposed to the risk of cancer than rural dwellers. This perhaps can be explained by the probability that the rural population is exposed to fewer carcinogens present in the environment.

Additionally, there is the often neglected element of greater host resistance in that rural populations are more likely to consume diet replete with antioxidant micronutrients some of which will also enhance the immune system. This is important in that its contact with carcinogens either in the diet or environment is inevitable. Thus it may be possible for us to avoid the most prominent risk factors the most pragmatic option appears to be reinforcing host resistance .This may be enhanced by the use of biomarkers. Biomarkers are playing an increasing role in the assessment of human exposure to hazardous environmental pollutants or chemicals and in risk assessment to these compounds. Biomarkers may be applied at any stage in the toxicological process, ranging from measurement of the external dose as an indicator of exposure to determine altered structure and function of cells as a marker of effect- carcinogenesis. Genetic carcinogens interact with nucleic acids to produce adducts, measurement of which is an indicator of the dose of active material which has reached the cells in question, termed biologically active dose (BAD), in the individual. This consequently incorporates the inter-individual variation in absorption, metabolism, and excretion of the compound which may affect risk assessment.

10. DNA repair: Introduction

To maintain the genomes of organisms, they have evolved a network of DNA repair pathways to excise altered residues from DNA. A major consideration is the relative contribution of environmental and endogenous DNA damage to carcinogenesis. DNA damage by environmental agents would have to be extensive and exceed that produced by normal endogenous reactive chemicals to be a major contributor to mutations and cancer.

This consideration underlines the difficulty in extrapolating risk of exposure to that which would occur at very low doses of carcinogens

10.1 DNA repair and role of micronutrients as biomarkers of susceptibility to DNA damage

Very many factors including nutritional factors have been shown to delay the carcinogenic process. Thus this can be exploited to reverse, delay or prevent the carcinogenic process. Maintenance of genome stability is of fundamental importance for counteracting carcinogenesis. Many human genome instability syndromes exhibit a predisposition to cancer. An increasing body of epidemiological evidence has suggested a link between nutrient status and risk of cancer. Populations in developing countries that are deficient in these protective micronutrients (WHO, 2002; Ames, 2010) and are increasingly exposed to chemicals owing to progressive or rapid industrialization are thus at increased risk (Anetor et al., 2008). Based on public data from the healthy people 2010 project, it is estimated that up to 80% of colon and prostate cancers may be influenced by diet, nutrition and life styles. As earlier indicated, it has been proposed that DNA damage induced by dietary micronutrient deficiency accounts for one-third of preventable cancers. Because micronutrient deficiencies can induce DNA damage in forms similar to those induced by ionizing radiation and reactive oxygen species (ROS), it has been suggested that oxidative stress and associated DNA breaks are critical targets for nutritional control of carcinogenesis. If left unrepaired, DNA lesions can promote accumulation of mutations that facilitate the process of carcinogenesis.

Micronutrients may act directly on the genome to prevent mutations, or indirectly as enzyme cofactors in cellular processes that modulate transformation. Thus micronutrient status may serve as biomarkers of risk of carcinogenesis. For instance low selenium status is a biomarker of risk of many cancers including cancer of the prostate. This should be particularly appealing to industrializing developing countries. Human cells possess an armamentarium of mechanisms for DNA repair that counter the extensiveness of DNA damage caused both by endogenous and environmental chemicals. These mechanisms include base excision repair (BER) that removes products of alkylation and oxidation (Duncan et al, 1976; Roth and Samson, 2002; Gersson, 2002); nucleotide excision repair (NER) that excises oligonucleotide segments containing larger adducts (Setlow and Carrier, 1963); mismatch repair that scans DNA immediately after polymerization for misincorporation by DNA polymerases (Modrich, 1991); and oxidative demethylation (Sedgwick, 2004), transcription-coupled repair (TCR) that preferentially repairs lesions that block transcription (Hanawalt, 1994); double-strand break repair and recombination that avoids errors by copying the opposite DNA strand (Friedberg et al, 2005); as well as mechanisms for the repair of cross-links between strands (Kuraoka et al, 2000; Zheng et al, 2005) that yet need to be established. Micronutrients deficiency disorder may inhibit DNA repair, thus acting as risk factors. Determinations of the levels of micronutrients may therefore serve as biomarker of susceptibility to DNA damage using the various instruments provided by biomedical engineering. Micronutrient deficiency for instance is inversely correlated with the level of 8-hydrodeoxyguanosine (8-OHdG), a marker of oxidative DNA damage, which is mutagenic and has to be removed by protective enzymes such as the human oxo-guanine DNA glycosylase (hOGG1).

Most DNA lesions are subject to repair by more than one pathway. As a result, only a minute fraction of DNA lesions which escapes correction are present at the time of DNA replication and can direct the incorporation of noncomplementary nucleotides resulting in mutation. Unrepaired DNA lesions initiate mutagenesis by stalling DNA replication forks or are copied over by error-prone *trans*-lesion DNA polymerases (McCulloch and Kunkel, 2008). Alternatively, incomplete DNA repair can result in the accumulation of mutations and mutagenic lesions, such as abasic sites (Loeb, 1985). Maintenance of genome stability is crucial for avoiding carcinogenesis. A number of human cancers display a range of chromosomal abnormalities; a characteristic now termed genome instability. The relationship between cancer and genome instability is well recognized, but the causes of genome instability in the evolution of human cancers is incompletely elucidated. The DNA damage response safeguards the integrity of the genome by detecting alterations, halting cell cycle progression and repairing damaged DNA. Zinc which plays a role in all the phases of cell cycle when deficient can be critical. (Anetor et al, 2008) Cells with defective DNA damage responses are characterized by high level of genome instability (Cheng, 2009).

It is known that in particular, cells in s-phase are vulnerable to agents, such as chemicals in the environment that cause DNA damage and induce DNA replication fork arrest. Since such events can adversely affect genomic stability, cells have evolved S-phase DNA response cascades, including checkpoint responses and DNA repair mechanisms, to fix DNA damage (Bartek et al., 2004). In response to DNA damage, check points are activated that coordinate DNA damage signaling, cell cycle arrest and DNA repair. Cells have developed elaborate systems to repair varieties of DNA damage. Abundant evidence has linked defects in DNA repair to carcinogenesis. As a way of avoiding mutagenic events, the DNA base excision repair (BER) pathway copes with oxidatively modified DNA (Xu et al., 1997; Kungland et al., 1999), and nucleotide excision repair can deal with bulky DNA adducts including DNA cross-links (Cleaver, 2005). It is noteworthy that micronutrient deficiency significantly affects DNA damage repair. Zinc (Zn) deficiency is common in children and adults (WHO, 2002; Moshfegh et al., 2005). Human cell culture studies demonstrating severe Zn deficiency causes complex IV deficiency and the release of oxidants, resulting in significant oxidative DNA damage (Ho and Ames, 2002). Zinc deficiency has also been reported to cause chromosome breaks in rats (Bell et al, 1975) which has been associated with cancer in both animal models and humans (Fong et al., 2005).

These reports strengthen the significant effect of micronutrient deficiency on DNA damage repair and by extrapolation on risk of carcinogenesis particularly chemical carcinogenesis in populations exposed to chemicals. Zinc deficiency in human cells has also been shown to inactivate Zn-containing proteins such as the tumor suppressor protein, p53 which plays a significant role in genome protection (Lane, 1992) and the DNA base excision repair enzyme,apyrimidinic/apurinic endonuclease, with a resulting synergistic effect on genetic damage (Ho and Ames, 2002;Ho Courtemanche and Ames, 2003).

11. Integrative cell biology and chemical carcinogenesis

Damage to DNA by chemical carcinogens activates checkpoint signaling pathways leading to cell cycle arrest and allows time for DNA repair processes (Sweasy et al, 2006). In the absence of repair, cells can use special DNA polymerases that copy past DNA adducts (Masutani et al., 2000) or undergo apoptosis by signaling the recruitment of immunologic and inflammatory host defense mechanisms. The demonstration that each methylcholanthrene-induced tumor has a unique antigenic signature provided one of the earliest glimpses into the stochastic nature of cellular responses to carcinogens. The immunologic and inflammatory responses facilitate not only engulfment and clearance of damaged cells but also the resulting generation of reactive oxygen (Klebanoff, 1988) and nitrogen radicals (Ohshima and Batsch, 1994) that further damage cellular DNA.

12. Chemical carcinogens and induced somatic mutations as biomarkers in molecular epidemiology

Extensive experience of laboratory research in chemical carcinogenesis has provided a solid foundation for the analysis of chemical-specific macromolecular adducts and related somatic mutations in humans as biomarkers of carcinogen exposure. A paradigm for validating causal relationships between biomarkers of carcinogens exposure and a cancer risk biomarker is shown in a number of observations (Wogan et al, 2005). The metabolite, AFB_1, a fungal toxin, is a prototypical example of an environmental chemical carcinogen that has been validated using this strategy. Benzo (a) pyrene, a polycyclic aromatic hydrocarbon 4-aminobiphenyl (Vineis and Pirastu, 1997), an aromatic amine dye, and 4-(N-methyl-N-nitrosamino)-1-(3-pyridyl)-1-butanone, a tobacco-specific N-nitrosamine (Hetch, 1999), are other key examples. The prevalence of Aflatoxin B1 in developing countries makes host resistance mechanism based on micronutrients of great import to reduction of risk of chemical carcinogenesis.

13. Impact of new technologies or analytical techniques and biomedical engineering

Recent advances in molecular methodologies are phenomenal, and they increasingly are being applied to understanding the interaction of chemical carcinogens with cellular constituents and metabolism. Cloning of DNA has facilitated the identification of specific genes mutated in human cancers. Chemical methods, including mass spectrometry, allow us to measure carcinogen alteration with unprecedented sensitivity and specificity, particularly the ultra high performance liquid chromatography (UHPLC) coupled to MS. Mass spectrometry is being coupled with many other site-specific techniques to study mutagenesis; to define how specific alterations in DNA produce cognate mutations. Sequencing of the human genome and the identification of DNA restriction enzymes opens up the field of molecular epidemiology, focusing in part on individual susceptibility to carcinogens.

A very useful technique in cancer studies is the single cell protein electrophoresis (COMET ASSAY). This assay is the only direct method for the detection of DNA damage in cells. It is used in cancer research, in genotoxicity studies on environmental mutagens, and for screening compounds for cancer therapeutics.

DNA microarray technology is a powerful tool that allows the activity of several genes to be monitored simultaneously. DNA microarray has been especially useful for detecting genes that respond to a specific chemical or physical signal in the same way. A DNA microarray or DNA chip consists of a solid surface, usually glass, to which DNA fragments are attached. The copies of each kind of fragments are attached to the glass surface at a specific site to regular pattern or array. DNA fragments attached to the chip act as probes that can hybridize with complimentary DNA or RNA molecules (targets) in solution. DNA microarray technology has been used to study a wide variety of gene expression problems such as tissue specific, cell cycle specific, and tumor- specific gene activation and repression.

Once a similar pattern expression profile has been established for a group of genes, it seems reasonable to assume that the profile is at least partly caused by similar transcription regions. Therefore, if information is available about a regulatory region in one gene within this group, it may provide clues to the regulatory regions of other members of the group as well as to the protein factors that activate or inhibit gene expression. The initial stimulus may be a given carcinogen. Array technology facilitates analysis of carcinogen-induced alterations in the expression of both protein coding and noncoding genes. These are all areas where biomedical engineering is expected to play a pivotal role in risk identification and early detection. On the horizon are techniques that can measure single molecules of carcinogens in cells, random mutations in individual cells, analysis of the dynamics of how molecules exist and work, and bioinformatics and genetic maps to delineate complex interacting functional pathways in cells. Underlying this progress in understanding chemical carcinogenesis is a cascade of advances in molecular biology that makes it feasible to quantify DNA damage by chemical agents, mutations, and changes in gene expression.

This calls for very intimate collaboration between biotechnology and biomedical engineering, a partnership that promises to be very rewarding in the fight against chemical carcinogenesis. Determining the structure of DNA, DNA sequencing, and the PCR revolutionized cell biology, including carcinogenesis. Advances in detection of DNA damage, including postlabeling of DNA (Randerath et al, 1981), immunoassays (Poirier et al, 1977) and mass spectrometry as earlier discussed (Singh and Farmer, 2006) have allowed the detection of a single altered base in 10^9 nucleotides using human nuclear DNA. This technology can be extended to analyze DNA or RNA in a single cell (Klein, 2005). Advances in cell biology, including array technology (Schena et al, 1995) and proteomics (Anderson et al, 1984; Aebersold et al, 1987), make it feasible to assess global changes in RNA and protein expression during carcinogenesis. Together, these technologies underlie systems biology, making it increasingly feasible to map biochemical pathways in cancer cells from DNA, to RNA, to proteins, to function. This again calls for greater involvement of the biomedical engineering field.

The ultimate application of techniques of biomedical engineering in the field of chemical carcinogenesis holds great promise, but faces several formidable challenges requiring ingenuity in matching technology with a social obligation to ensure that those most affected can afford its dividend to combat the menace of chemical carcinogenesis. This calls for sensitive methods for the early detection. The complex interplay among these and the great promise they hold for human health, particularly in the rapidly industrializing developing countries has not been adequately addressed. This contribution largely sees this as needing to be addressed- carcinogenesis, integrating chemical exposure, nutritional; mainly micronutrient modulation and the yawning gap biomedical engineering should fill in non-invasive applications. Although the field of biomedical engineering as regards cancer studies is still relatively in its early stages and will be a long-term effort and the magnitude of the task far greater than the physical resources and intellectual capacity currently available, the challenge is to focus on sensitive, specific or selective, cost effective and efficient techniques with dereliction for resource poor rapidly industrializing developing countries that appear to bear a greater brunt of the increasing chemical exposure and attendant carcinogenic risk.

14. References

Aebersold R. H., Leavitt. J., Saavedra, R. A, Hood, L. E. and Kent, S. B. (1987). Internal amino acid sequence analysis of proteins separated by one- or two-dimensional gel electrophoresis after in situ protease digestionon nitrocellulose. *Proc. Natl. Acad. Si. USA*. 84:6970-6974

Ames, B. N. and Wakimoto, P. (2002). Are vitamin and mineral deficiencies a major cancer risk? *Nat. Rev. Cancer.* 2:694-704.

Ames, B. N. (2001). DNA damage from micronutrient deficiencies is likely to be a major cause of cancer, *Mutat. Res*. 475:7-20.

Ames, B.N., Gold, L. S. and Willet, W. C. (1995). The causes and prevention of cancer. *Proc. Natl. Acad. Sci.* USA 92: 5258–5265.

Ames, B. N. and Gould, C. S. (1991). Endogenous mutagenesis and the causes of aging and cancer. *Mutations Research*, 250: 3.

Ames, B. N. (1983). Dietary carcinogens and anticarcinogenesis. *Science*. 221:1256-1264.

Ames, B. N. Durston, W. E., Yamasaki, E. and Lee, F.D. (1973). Carcinogens are mutagens: A simple test system combining liver homogenates for activation and bacteria for detection. *Proc. Natl. Acad. Sci.* USA. 70: 2281-2285.

Anderson, N. L., Hofmann, J. P., Gemmell, A. and Taylor, J. (1984). Global approaches to quantitative analysis of gene-expression patterns observed by use of two-dimensional gel electrophoresis. *Clin. Chem*., 30: 2031-2036.

Anetor J. I, Anetor G. O, Udah,D.C, Adeniyi F. A. A(2008): Chemical carcinogesis and chemoprevention:Scientific priority area in rapidly industrializing developing countries. African Journal of Environ. Sci. Tech. 2.(7):150-156

Avery, O. T., MacLeod, C. M. and McCarty M.(1944). Studies on the chemical nature of the substance inducing transformation of pneumoccal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumoccus type III. *J Exp Med*.79:137–58.

Bartek, J., Lukas, C and Lukas, J. (2004). Checking on DNA damage in S phase. *Nat. Rev. Mol. Cell Biol*.5:792-804.

Bartkova, J., Rezaei, N., Liontos, M., Karakaidos, P et al., (2006). Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature.* 444:633-637.

Bartkova, J., Horejsi, Z., Koed, K., Kramer,A.,et al., (2005). DNA damage response as a candidate anticancer barrier in early human tumorigenesis. *Nature.* 434:864-870.

Bell, L. T., Branstrator, M., Roux, C., and Hurley, L. S. (1975). Chromosomal abnormalities in maternal and fetal tissues of magnesium or zinc deficient rats. *Teratology*. 12(3):221-226.

Berenblum, I. (1941). The mechanism of carcinogenesis. A study of the significance of cocarcinogenic action and related phenomena. *Cancer Research*, 1: 807.

Block, G., B. Patterson, B. and Subar, A. (1992). Fruit, vegetables and cancer prevention: A review of the epidemiologic evidence. *Nutr. Cancer* 18: 1–29.

Burcham, P. C. (1998). Genotoxic lipid peroxidation products: Their DNA damaging properties and role in formation of endogenous DNA adducts. *Mutagenesis* 13: 287–305.

Cairns, J. (1981). The origin of human cancers. *Nature*, 289: 353-357.

Carrell, J. C., Carrell, T. G., Carrell, H. L., Prout, K. and Glusker, P.(1997). Benzo[a]pyrene and its analogues: structural studies of molecular strain. *Carcinogenesis*, 18:415–22.

Cheng, W. H. (2009). Impact of inorganic nutrients on maintenance of genomic stability. *Environ. Mol. Mutagenesis*. 50:349-360.

Clayson, D.B. (2001).*Toxicological carcinogenesis*. Lewis Publishers, Florida.

Cleaver, J. E. (2005). Cancer in xeroderma pigmentosum and related disorders of DNA repair. *Nat. Rev. Cancer*. 5:564-573.

Cohens, S. M. and El Wein, C. B. (1991). Genetic erors, cell proliferation and carcinogenesis. *Cancer Research*, 51: 6493.

Croy, R. G., Essigman, J. M., Reinhold, V. N. and Wogan, G. N.(1978). Identification of the Principal Aflatoxin B1-DNA Adduct Formed *in vivo* in Rat Liver. *PNAS*, 75:1745–9.

Duncan, J., Hamilton, L. and Friedberg, E. C. (1976). Enzymatic degradation of uracil-containing DNA. *J. Virology*, 19: 338-345.

Eden, A., Gaudet, F., Waghmare, A. and Jaenisch, R.(2003). Chromosomal instability and tumors promoted by DNA hypomethylation, *Science* 300:455.

Esaka, Y., Inagaki, S. and Goto, M. (2003). Seperation procedures capable of revealing DNA adducts. *J. Chromatography B.*, 797: 321-329.

Fausto, N., Camobekk, J. S. and Riehle, K. J. (2006). Liver regeneration. *Hepatology*, 43: S45-53.

Fenech, M. (2001). The role of folic acid and Vitamin B12 in genomic stability of human cells, *Mutat. Res.* 475: 57–67.

Fong, L. Y. Y., Zhang, L., Jiang, Y., and Farber, J. L. (2005). Dietary Zinc modulation of COX-2 expression and lingual and esophageal carcinogenesis in rats. *J. Natl. Cancer Instit.* 97(1):40-50.

Foulds, L. (1961). Neoplastic development. Vol. 1. Academic press, London.

Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A. and Ellenberger, T. (2005). DNA repair and mutagenesis 2. ASM press, Washington D. C.

Gerson, S. L. (2002). Clinical relevance of MGMT in the treatment of cancer. *J. Clinical Oncology,* 20: 2388-2399.

Gerster, H. (1995). Beta-carotene, vitamin E and vitamin C in different stages of experimental carcinogenesis. *Eur. J. Clin. Nutr.* 49: 155–168.

Gorgoulis V. G.,Vassiliou, L. F., Karakaidos, P., Zacharatos, P., et al., (2005). Activation of the DNA damge checkpoint and genomic instability in human precancerous lesions. *Nature.* 434:907-913.

Hagen, T. M., Ingersoll, R. T., Liu. J., Lykkesfeldt. J., Wehr. C. M., Vinarsky, V., Barthelomew, J. C. and Ames, B. N. (1998). (*R*)-α-Lipoic acid-supplemented old rats have improved mitochondrial function, decreased oxidative damage, and increased metabolic rate. *FASEB J.* 13: 411–418.

Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*.100:57-70.

Halliwell, B. and Aruoma, O.I. (1991). DNA damage by oxygen-derived species. Its mechanism and measurement in mammalian systems. *FEBS Lett.,* 291: 9-19.

Hanawalt, P. C. (1994). Transcription-coupled repair and human disease. *Science,* 266: 1957-1958.

Hecht, S.S. (1999). Tobacco smoke carcinogens and lung cancer. *J. Natl. Cancer Inst.* 91: 1194-1210.

Henning, S. M., Swendseid, M. E. and Coulson, W. F. (1997). Male rats fed methyland folate-deficient diets with or without niacin develop hepatic carcinomas associated with decreased tissue NAD concentrations and altered poly(ADP-ribose) polymerase activity. *J. Nutr.* 127: 30–36.

Hennings, H. and Boutwell, R. K. (1970). Studies on the mechanism of skin tumor promotion. *Cancer Research,* 30:312-20.

Ho, E and Ames, B.N. (2002). Low intracellular zinc induces oxidative DNA damage, disrupts p53, NFkappa B, AP1 DNA binding and affects DNA repair in a rat glioma cell line. *Proc. Natl. Acad. Sci. USA.* 99:16770-16775.

Ho, E., Courtemanche, C., and Ames, B. N. (2003). Zinc deficiency induces oxidative DNA damage and increases p53 expression in human lung fibroblast. *J. Nutr.* 133:2543-2548.

International Agency for Research on Cancer. (1971-2006). IARC monographs on the evaluation of carcinogenic risks to humans. Overall evaluation of carcinogenicity. Monographs Volumes 1 to 99. http://monographs.iarc.fr.

Kallioniemi A., Kallioniemi, O. P., Sudar, D., et al., (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science.* 258:818-821.

Klebanoff, S. J. (1988). Phagocytic cells: Products of oxygen metabolism. In: Gallin, J. I., Goldstein, I. M. and Snyderman, R. (editors). Inflammation: Basic principles and clinical correlates. Raven press Ltd, New York. Pp391-444.

Klein, C. A. (2005). Single cell amplification methods for the study cancer and cellular ageing. *Mech. Ageing Dev.,* 126: 147-151.

Klein, G. (1987). The approaching era of the tumor suppressor genes. *Science,* 238: 1539.

Klein, G. and Klein, E. (1984). Commentry: Oncogene activation and tumor progression. *Carcinogenesis,* 5: 429.

Klungland, A., Rosewell, I., Hollenbach, S., Larsen, E., et al., (1999). Accumulation of premutagenic DNA lesions in mice defective in removal of oxidative base damage. Proc. *Natl. Acad. Sci. USA.* 96:13300-13305.

Kuraoka, I., Kobertz, W. R., Ariza, R.R., Biggerstaff, M., Essigmann, J. M. and Wood, R. D. (2000). Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease. *J. Biol. Chem.,* 275: 26632-26636.

Lindahl, T. and Nyberg, B. (1974). Heat-induced deamination of cytosine esidues in deoxyribonucleic acid. Biochem. 13:3405-3410.

Lindahl, T. and Nyberg, B. (1972). Rate of depurination of native deoxyribonucleic acid. *Biochemistry.* 11: 3610-3618.

Loeb, L. A. (1985). Apurinic sites as mutagenic intermediates. *Cell.* 40:483-484

Martin, J. H., Crotty, S. and Nelson, P. N. (2007). Phytoestrogens: Perpetrators or protectors? *Future Oncology,* 4: 3007.

Masutani, C., Kusumoto, R., Iwai, S. and Hanaoka, F. (2000). Mechanisms of accurate translesion synthesis by human DNA polymerase eta. *EMBO J.,* 19: 3100-3109.

McCulloch, S. D. And Kunkel, T. A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res. 18:148-161

Miller, E.C., Miller, J. A. and Hartman, J. A. (1961). N-Hydroxy-2-acetyllaminofluorine: A metabolite of 2-actylaminofluorine with increased carcinogenic activity in the rat carcinogenesis. *Cancer Research,* 2:815.

Modrich, P. (1991). Mechanisms and biological effects of mismatch repair. *Annu. Rev. Genet.,* 25: 229-253.

Moshfegh, A., Goldman,J., and Cleveland, L. (2005). "What we eat in America. NHANES 2001-2002: Usual nutrient intakes from food compared to dietary reference intakes," U. S. Department of Agriculture, Agricultural Research Service.

Ohshima, H. and Bartsch, H. (1994). Chronic infections and inflammatory processes as cancer risk factors: possible role of nitric oxide in carcinogenesis. *Mutat. Res.* 305: 253-264.

Oyasu, R. (1995). Epithelial tumors of the lower urinary tract in humans and rodents. Food and *Chemical Toxicology,* 33: 747.

Oyasu, R. Hirao, I. and Izumi, K. (1981). Enhancement cement by urine of urinary bladder carcinogenesis. *Cancer Research,* 41: 478.

Perera, F. P.(1997): Environment and Cancer: Who are susceptible?. Science vol. 278:1068-1073.

Poirier, M. C., Yuspa, S. H., Weinstein, I. B. and Blobstein, S. (1977). Detection of carcinogen-DNA adducts by radioimmunoassay. *Nature*, 270: 186-188.

Randerath, K., Reddy, M. V. and Gupta, R. C. (1981). 32P-labelling test for DNA damage. *Proc. Natl. Acad. Sci.* USA., 78: 6126-6129.

Roth, R. B. and Samson, L.D. (2002). 3-Methyladenine DNA glycosylase-deficient Aag null mice display unexpected bone marrow akylation resistance. *Cancer Research,* 62: 656-660.

Schena, M., Shalon, D., Davies, R. W. and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science,* 270: 467-470.

Sedgwick, B. (2004). Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.,* 5: 148-157.

Setlow, R. B. and Carrier, W. L. (1963). The disappearance of thymine dimmers from DNA: An error-correcting mechanism. *Proc. Natl. Acad. Sci.* USA. 51: 226-231.

Singh. R. and Farmer, B.P. (2006). Liquid chromatography-electro-spray-ionization-mass spectrometry: The future of DNA adduct detection. *Carcinogenesis*, 27: 178-196.

Sjoblom T., Jones. S., Wood, L. D. et al. (2006). The concensus coding sequences of humanbreast and colorectal cancers. *Science.* 314: 268-274.

Solt, D. and Farber, E. (1976). New principle for the analysis of chemical carcinogenesis. *Nature*, 263:70b

Sweasy, J. B., Lauper, J. M. and Eckert, K. A. (2006). DNA polymerases and human diseases. *Radiat. Res*. 166: 693-714.

Swenberg, J.A., Richardson, F.C., Boucheron, J. A. et al., (1987). High- to low-dose extrapolation: Critical determinants involved in the dose response of carcinogenic substances. *Environmental Health Perspectives,* 76: 57-63.

Tolonen, M. (1990). *Vitamins and minerals in health and nutrition.* Ellis Horwood Ltd. Chichester , West Sussex, England.

Vineis, P. and Pirastu, R. (1997). Aromatic amines and cancer. *Cancer Causes Control*, 8: 346-355.

Watson, J. D. and Crick, F. H.(1953). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid.*Nature*, 171:737–8.

Weinberg, E (1981). Review: Iron and neoplasia. *Biol. Trace Element Res.* 3:55 - 80.

Weinberg, E. (1984). Iron withholding: A defense against infection and neoplasia. *Physiol. Rev.* 64(1):65- 102.

WHO (2002). The world health report 2002. Reducing risk, promoting healthy lifestyle. World Health Organisation, Geneva.

Willett, W.C. and Trichopoulos, D.(1996). Nutrition and cancer: A summary of the evidence. *Cancer Causes and Control* 7: 178–180.

Wogan, G. N., Hecht, S. S., Felton, J. S., Conney, A. H. and Loeb, L. A. (2004). Environmental and chemical carcinogenesis. *Semin. Cancer Biol.*, 14: 473-486.

Xu, Y., Moore, D. H., Broshears,J., Liu, L., Wilson, T. M. and Kelley M.R. (1997). The apurinic/apyrimidinic endonuclease (APE/ref-1) DNA repair enzyme is elevated in premalignant and malignant cervical cancer. Anticancr Res. 17:3713-3719.

Yamagiwa, K. and Ichikawa, K. (1918). Experimental study of the pathogenesis of carcinoma. J. Cancer Research, 3:1-21.

Zheng, H., Wang, X., Legerski, R. J., Glazer, P. M. and Li, L. (2006). Repair of DNA interstrand cross-links: Interactions between homology-dependent and homology-independent pathways. *DNA Repair,* 5: 566-574.

4

AGE/RAGE as a Mediator of Insulin Resistance or Metabolic Syndrome: Another Aspect of Metabolic Memory?

Hidenori Koyama and Tetsuya Yamamoto
Department of Internal Medicine, Division of Endocrinology and Metabolism,
Hyogo College of Medicine,
Japan

1. Introduction

Large randomized studies in diabetes have established that early intensive glycemic control reduces the risk of diabetic microvascular complications, with less impact on macrovascular complications [1, 2]. In type 2 diabetic patients, further intensive therapy to target normal glycated hemoglobin levels also failed to reduce mortality and major cardiovascular events [3, 4], while it may be rather harmful [5]. However, follow-up data of these trials reveal a long-term influence of early metabolic control on longer cardiovascular outcomes, even though the influence on glycemic control has been immediately disappeared after the trials [6, 7]. This phenomenon has recently been defined as "metabolic memory". In at-risk patients with type 2 diabetes, intensive intervention with multiple drug combinations and behavior modification had similar sustained beneficial effects with respect to vascular complications and on rates of death from any cause and from cardiovascular causes [8]. Similarly in patients with end-stage renal disease (ESRD), intensive interventions to the general risk factors, such as high LDL-cholesterol or C-reactive protein, have not been successful in improving their cardiovascular outcomes [9, 10], suggesting that the beneficial effect of risk reduction may be overwhelmed by accumulated “metabolic memory” by long-term exposure to oxidative stress during the progression of renal failure.

Potential mechanisms for propagating this "memory" are the non-enzymatic glycation of cellular and tissue proteins which are conceptualized as advanced glycation end-products (AGEs), the generation of which has been implicated to be deeply associated with increased oxidative stress as well as hyperglycemia. AGEs, with their receptor (receptor for AGEs, RAGE), potentially mediate molecular and cellular pathway leading to metabolic memory. Moreover, interaction of the RAGE with AGEs leads to crucial biomedical pathway generating intracellular oxidative stress and inflammatory mediators, which could result in further amplification of the pathway involved in AGE generation.

By utilizing genetically engineered mouse models, emerging evidence suggests that AGE/RAGE axis is also found to be profoundly associated with non-diabetic, non-uremic pathophysiological conditions including 1) atherogenesis, 2) angiogenic response, 3) vascular injury, and 4) inflammatory response (see review in [11]), many of which are now implicated in metabolic syndrome. Numerous truncated forms of RAGE have also been described, and the

C-terminally truncated soluble form of RAGE has received much attention. Soluble RAGE consists of several forms including endogenous secretory RAGE (esRAGE) which is a spliced variant of RAGE [12], and a shedded form derived from cell surface RAGE [13, 14]. These heterogeneous forms of soluble RAGE, carrying all of the extracellular domains but devoid of the transmembrane and intracytoplasmic domains, bind ligands including AGEs, and may antagonize RAGE signaling in vitro and in vivo. ELISA systems to measure plasma esRAGE and total soluble RAGE have been developed, and decreased plasma esRAGE is found to be associated with insulin resistance, obesity and metabolic syndrome [15]. Moreover, our recent observation highlights the direct role of RAGE in adiposity; RAGE deficiency is associated with less weight gain, less abdominal fat mass, less adipocyte size, less atherosclerotic lesion formation and higher plasma adiponectin than wild type control [16].

Insulin resistance is the primary mechanism underlying the development of type 2 diabetes and is a central component defining the metabolic syndrome, a constellation of abnormalities including obesity, hypertension, glucose intolerance, and dyslipidemia. Insulin resistance or metabolic syndrome has been defined to be associated with low-grade inflammation, and therefore inflammation could contribute in large part to its development [17], implicating an intriguing possibility that this pathophysiological condition is also an additional face of metabolic memory driven by RAGE axis. Although insulin resistance has been characterized by complex factors including genetic determinants, nutritional factors, and lifestyle, growing evidence suggests that mediators synthesized from inflammatory cells are critically involved in the regulation of insulin action. In brief, insulin binding to its specific receptor stimulates tyrosine phosphorylation of insulin receptor substrate (IRS) proteins, which is a crucial step for insulin signaling system. Many inflammatory signals appear to induce serine phosphorylation of IRS, which could be involved in disruption of insulin-receptor signaling [17]. In this chapter, we would like to summarize the recent findings regarding pathophysiological roles of RAGE and soluble RAGE in insulin resistance and metabolic syndrome.

2. AGEs

2.1 AGE formation by glucose and its derivatives

AGEs are proteins generated by a series of reactions termed the Maillard Reaction. Classically, AGE formation has been described by a nonenzymatic reaction between proteins and glucose [18, 19]. AGEs derive from the spontaneous reaction of carbohydrates with amino group of proteins, which undergo from the formation of reversible products (Schiff base adducts) to the generation of more stable products (Amadori products). Subsequently, complex reactions occur including intermolecular crosslink formation, and cleavage through oxidation, dehydration, condensation, cyclization, and other reactions follows, with generation of AGEs through a late reaction characterized by fluorescent and brown coloration and molecular crosslinkage. Recently, it was confirmed that AGEs are also formed by non-enzymatic reaction of reactive carbonyl compounds such as 3-deoxyglucosone, methylglyoxal resulting from persisting high blood glucose level, and oxidative stress associated with the amino residues of proteins.

2.2 AGEs formation independent of hyperglycemia

There is also increasing evidence that AGEs are also formed through lipid-derived intermediates, resulting in advanced lipoxidation products [20]. AGEs might be formed directly by autoxidation of free glucose [21, 22]. In this pathway, known as autoxidative glycosylation, such reactive oxygen species as hydrogen peroxide were identified as both products and catalysts of autoxidation of sugars. Other than diabetes mellitus patients, high plasma and tissue levels of AGEs are observed in patients with ESRD. It has been reported that no difference was noted in blood AGEs levels between those with and without diabetes mellitus among chronic renal failure patients on hemodialysis, which is believed to enhance production and accumulation of AGEs in conditions other than hyperglycemia. Local accumulation of AGEs is also observed in patients with Alzheimer disease, rheumatoid arthritis, arteriosclerosis, cancer, and other diseases, suggesting the involvement of inflammation and oxidative stress in the formation of AGEs.

2.3 Orally absorbed AGEs

In addition to the endogenously formed, AGEs are abundant in exogenous sources such as foods, especially when prepared under elevated temperatures [23]. Vlassara's group has extensively examined the role of the exogenous AGEs in several pathological conditions (see review in [24]). After ingestion, 10% of orally-administered AGEs are absorbed into the circulation [25-27], majorities of which are shown to be accumulated in tissues. Among them are tissue-reactive α, β-dicarbonyl-containing intermediate products, such as methylglyoxal, which has been linked to cellular oxidant stress and apoptosis [28], and terminal products, such as $^{\varepsilon}$ *N*-carboxymethyllysine (CML), which is formed by glycoxidation as well as by lipoxidation [29-31]. Both methylglyoxal and CML have been identified in vivo and are shown to be associated with oxidant stress and tissue damage [31-33].

3. AGEs and endogenous RAGE ligand in insulin resistance or metabolic syndrome

3.1 AGEs in insulin producing and acting tissues

Importantly, among the multiple targets of bioactive AGEs are also such diverse tissues as the pancreatic islet [34, 35], the adipose tissue (adipocyte) [36] and skeletal muscle cells [37], major tissues involved in insulin secretion and its actions. Pharmacological inhibition of glycoxidation protects against damage to either tissue [33, 38]. AGEs are shown to inhibit glucose-stimulated insulin secretion from islet through iNOS-dependent nitric oxide production [35]. Reduced intake of dietary AGEs has also been shown to decrease the incidence of type 1 diabetes in NOD mice [39] as well as the formation of atherosclerotic lesions in diabetic apolipoprotein E-deficient mice [40].

3.2 AGEs and insulin resistance in vivo

AGEs burden is also shown to be associated with impaired endothelial function [41-44]. Endothelial dysfunction could be profoundly associated with less insulin delivery to the skeletal muscle interstitium, leading to decreased insulin-stimulated glucose uptake by the skeletal muscle [45-49], which is implicated in the pathogenesis of insulin resistance. More directly, the restriction of the AGE content in standard mouse diets was found to markedly improve insulin resistance in obese *db/db*$^{(++)}$ mice [50]. More recent observation shows targeted reduction of the advanced glycation pathway improved renal function in obesity [51]. This interesting observation was further supported by the findings that the development of insulin resistance and type 2 diabetes during prolonged high-fat feeding are linked to the excess AGEs/advanced lipoxidation end products inherent in fatty diets [52].

3.3 AGEs and insulin resistance in vitro

Several evidences also suggest that AGEs affect the function of insulin-target cells in vitro. AGEs interact with CD36 in mouse 3T3 and human subcutaneous adipocytes, which is associated with down-regulation of leptin expression in adipocyte through reactive oxygen species (ROS) system [53]. Miele et al showed in L6 skeletal muscle cells that AGEs affect glucose metabolism by impairing insulin-induced insulin receptor substrate (IRS) signaling through protein kinase Cα-mediated mechanism [37]. The same research group also showed in the muscle cells that methylglyoxal, an essential source of intracellular AGEs, hampers a key insulin signaling molecule [54]. Recent observations by Unoki et al also showed that AGEs impair insulin signaling in adipocytes by increasing generation of intracellular ROS [55]. Thus, AGEs may not only induce the debilitating complications of diabetes, but may also contribute to the impairment of insulin signaling in insulin-target tissues which could be involved in pathophysiology of insulin resistance, metabolic syndrome and diabetes.

3.4 Endogenous RAGE ligands, insulin resistance and metabolic syndrome

RAGE also interacts with other endogenous non-glycated peptide ligands including S100/calgranulin [56], amphoterin (also termed as high mobility group box 1 protein, HMGB1) [57, 58], amyloid fibrills [59], transthyretin [60], and a leukocyte integrin, Mac-1 [61], many of which are important inflammatory regulators. Some of these inflammatory ligands for RAGE may be involved in pathogenesis of obesity and metabolic syndrome. Early studies show expression of S100B protein in pre- and mature- adipocyte and is induced during adipogenesis [62, 63]. Physiological S100B levels appear to closely reflect adipose tissue mass or insulin resistance in humans [64-66]. HMGB1 is also found to be expressed in human adipose tissue with the expression level associated with the fat mass and obesity-associated gene [67]. Moreover, growing evidences suggest that infiltration of inflammatory cells, including macrophages, play fundamental roles in adiposity and metabolic syndrome [68-70]. MAC-1, an integrin expressed in macrophage, can act as a RAGE ligand [61], and may be involved in adipogenesis through interaction with RAGE.

4. RAGE and its potential link with insulin resistance and metabolic syndrome

4.1 Structure and function of RAGE

RAGE is a multiligand cell-surface protein that was isolated from bovine lung in 1992 by the group of Schmidt and Stern [71, 72]. RAGE belongs to the immunoglobulin superfamily of cell surface molecules and has an extracellular region containing one "V"-type immunoglobulin domain and two "C"-type immunoglobulin domains [71, 72] (Figure 1). The extracellular portion of the receptor is followed by a hydrophobic trans-membrane-spanning and then by a highly charged, short cytoplasmic domain which is essential for intracellular RAGE signaling. RAGE is initially identified as a receptor for CML-modified proteins [73], a major AGE in vivo [74]. Three-dimensional structure of the recombinant AGE-binding domain by using multidimensional heteronuclear NMR spectroscopy revealed that the domain assumes a structure similar to those of other immunoglobulin V-type domains [75, 76]. Three distinct surfaces of the V domain were identified to mediate AGE-V domain interactions [75]. The site-directed mutagenesis studies identified the basic amino acids which play a key role in the AGE binding activities [76]. As mentioned in the previous sentence, RAGE also interacts with other endogenous non-glycated peptide ligands, many of which are important inflammatory regulators. The common characteristics of these ligands are the presence of multiple β-sheets [61, 77, 78]. RAGE is thought to interact with these ligands through their shared three-dimensional structure.

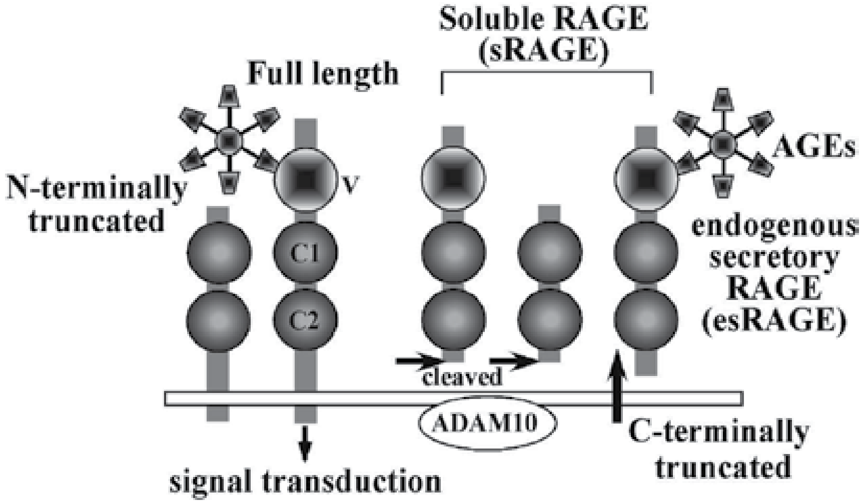


Fig. 1. Numerous truncated forms of RAGE. There are three major spliced variants of RAGE: full length, N-terminally truncated, and C-terminally truncated. The C-terminally truncated form of RAGE is secreted from the cell and is named endogenously secreted RAGE (esRAGE). esRAGE has a V-domain, which is essential for binding with ligands, and is capable of competing with RAGE signaling as a decoy receptor. There are other forms of soluble RAGE (sRAGE) that are cleaved from cell-surface RAGE by matrix metalloproteinases. The ELISA assay for sRAGE measures all soluble forms including esRAGE in human plasma, while the ELISA for esRAGE measures only esRAGE, using polyclonal antibody raised against the unique C-terminus of the esRAGE sequence.

4.2 Inflammatory signaling mediated by RAGE

Ligand engagement of RAGE leads to prolonged inflammation, resulting in a RAGE-dependent expression of proinflammatory mediators such as monocyte chemoattractant protein-1 (MCP-1) and vascular cell adhesion molecule-1 (VCAM-1) [79, 80]. RAGE-mediated proinflammatory signals could potentially converge with insulin signaling system (Figure 2). The engagement of RAGE has been reported to induce activation of the transcription factor nuclear factor-κB (NF-κB). Recent reports by Harja et al demonstrate that RAGE mediates upregulation of VCAM-1 in response to S100b and oxLDL and JNK MAP kinase underlies the RAGE ligand-stimulated molecular events [81]. It is not known at present whether this is also the case in classical insulin target cells. JNK activity is strikingly increased in critical metabolic sites (eg. adipose and liver tissues) [82], and is shown to be crucial in IRS-1 phosphorylation and consequently insulin resistance [82, 83]. Moreover, the main pathological consequence of RAGE ligation is the induction of intracellular reactive oxygen species (ROS) via NAD(P)H oxidases and other identified mechanisms such as mitochondrial electron transport chain [84], which consequently results in oxidative stress in the cells [85]. Oxidative stress is emerging as a feature of obesity and an important factor in the development of insulin resistance [86, 87]. Both the NF-kB and JNK pathways can be activated under the conditions of oxidative stress, and this may be important for the ability of ROS to mediate insulin resistance. RAGE has a short cytosolic portion that contains 43 amino acids [72]. So far, adaptors and/or scaffold proteins that interact with the cytosolic tail of RAGE has barely been identified. The RAGE mutant lacking the 43-residue C-terminal tail fails to activate NF-κB, and expression of the mutant receptor results in a dominant negative effect against RAGE-mediated production of proinflammatory cytokines from macrophages [56, 57].

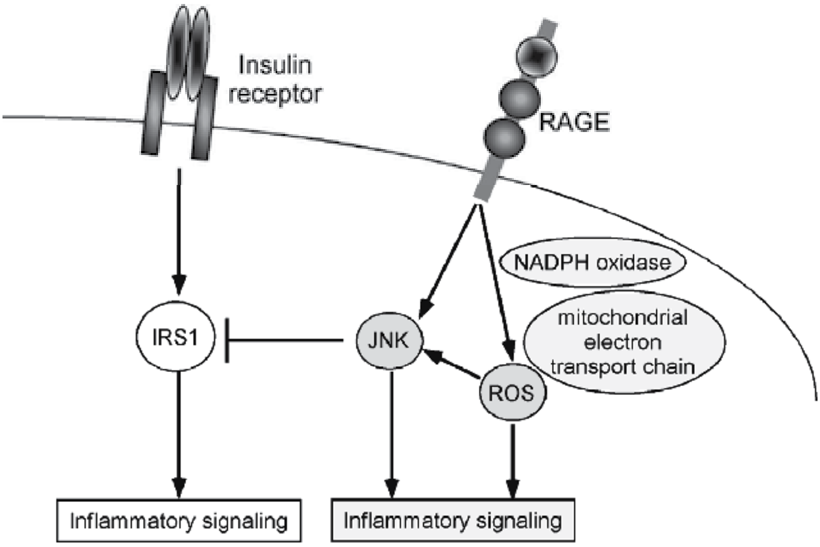


Fig. 2. RAGE and insulin signaling. RAGE is known to activate JNK pathway, which could phosphorylate serine-residue of insulin receptor substrate (IRS) and inhibit its activity. RAGE mediated generation of reactive oxygen spices (ROS) may alternatively influence insulin signaing.

4.3 RAGE and obesity

A. RAGE, adiposity and atherosclerosis in mouse model

Recent reports suggest that RAGE could be involved in progression of obesity. Recent study in humans shows RAGE mRNA expression in subcutaneous adipose tissues [88]. Although this study does not delineate which cells in adipose tissue express RAGE, our current animal study shows RAGE expression in adipocyte as wells as endothelial cells in adipose tissues [16]. We have shown by using apo E/RAGE double knockout mice that progression of atherosclerosis is closely associated with RAGE-regulated adiposity in non-diabetic conditions [16]. As shown in Figure 3, apoE$^{-/-}$RAGE$^{-/-}$ mice fed either with standard or atherogenic diet exhibited significantly decreased atherosclerotic plaque area in aorta as compared with apoE$^{-/-}$RAGE$^{+/+}$ mice. Importantly, apoE$^{-/-}$RAGE$^{-/-}$ mice also exhibited significantly less body weight, epididymal fat weight and epididymal adipocyte size than apoE$^{-/-}$RAGE$^{+/+}$ mice at 20-weeks of age (Figure 4). Decreased body weight, epididymal fat weight, and adipocyte size are associated with higher plasma adiponectin levels and decreased atherosclerosis progression. RAGE is involved in adiposity even in apo E$^{+/+}$ genetic background. At 20-weeks of age, epididymal adipocyte size of RAGE$^{-/-}$ mice was significantly smaller than that of RAGE$^{+/+}$ mice (data not shown).

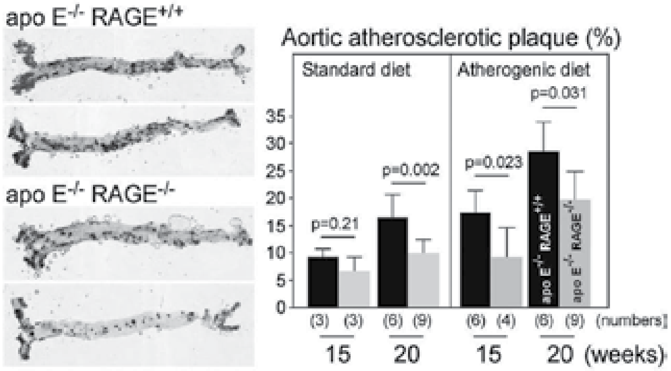


Fig. 3. RAGE deficiency suppresses atherosclerotic progression in apoE deficient mice. Representative aortas from apoE$^{-/-}$RAGE$^{+/+}$ and apoE$^{-/-}$RAGE$^{-/-}$ mice (20-weeks old) fed with atherogenic diet were shown in left panel. Right panel summarizes the quantitative analyses. Plaque area was represented as percentages of the total plaque area. Columns represent mean ± standard deviation. Black columns represent apoE$^{-/-}$RAGE$^{+/+}$ mice, and grey columns, apoE$^{-/-}$RAGE$^{-/-}$ mice. P values were analyzed by Student's unpaired t-test. Reproduced from ref [16].

B. Roles of inflammatory cells?

RAGE is also known to play fundamental role in functions of inflammatory cells [61, 89, 90], raising an intriguing possibility that RAGE's function on adiposity may be mediated through its function in inflammatory cells infiltrated in adipose tissues. In our study in apoE$^{-/-}$ genetic background fed with atherogenic diet, numbers of Mac-3-positive inflammatory cells infiltrated in the epididymal adipose tissues of RAGE$^{+/+}$apoE$^{-/-}$ mice and RAGE$^{-/-}$apoE$^{-/-}$ did not show significant differences, and crown-like structure were barely detected in epididymal adipose tissue in both groups even at 20-week of age. In standard diet-fed mice, even though the adiposity was significantly different between RAGE$^{+/+}$apoE$^{-/-}$ and RAGE$^{-/-}$apoE$^{-/-}$ mice, crown-like structure were not detected in epididymal adipose tissues in both groups even at 20-week of age. Further in apoE$^{+/+}$ genetic background at 10 week of age when significantly different pattern of gene expression was observed between WT and RAGE$^{-/-}$ mice, no marked differences in expressions of macrophage markers were observed as analyzed by gene microarray. At that age, macrophage infiltration in adipose tissues is also reported to be scant [91]. Thus, it appears infeasible to RAGE acting primarily at inflammatory cells at least in early phase of adiposity, while RAGE expressed in endothelial cells or adipocyte might play fundamental roles.

C. RAGE-regulated genes in adipose tissue: gene chip analysis

To explore potential mechanisms underlying RAGE-regulation of adiposity, mRNA expression profile in epididymal adipose tissue was compared between RAGE$^{+/+}$ and RAGE$^{-/-}$ mice using Affymetrix GeneChip Mouse Genome 430 2.0. We isolated total RNA from epididymal adipose tissue at 10-weeks of age, at which phenotypic change in adipocyte size was not observed. Using 3 µg of total RNA, 59.8% and 61.4% of 45,037 genes were revealed to be present in RAGE$^{+/+}$ and RAGE$^{-/-}$ adipose tissue, respectively. Comparison analysis of the genes (RAGE$^{+/+}$ adipose tissue as base line) revealed that 10.3% of the total genes were decreased, while 11.7% increased in RAGE$^{-/-}$ adipose tissue. As compared with RAGE$^{+/+}$ adipose tissue, 623 genes were downregulated to less than a half, and 2,470 genes upregulated more than 2 fold in RAGE$^{-/-}$ adipose tissue.

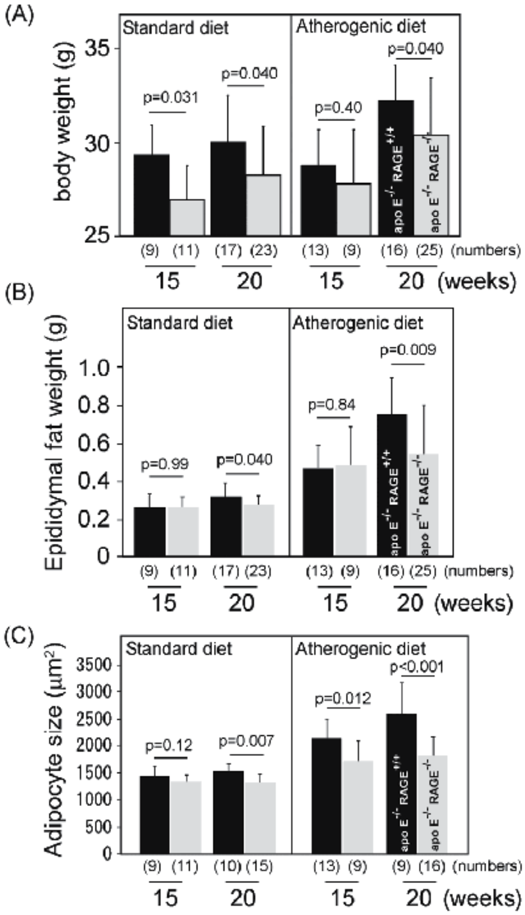


Fig. 4. RAGE deficiency is associated with decreased body weight, epididymal fat weight and adipocyte size in apolipoprotein E (apoE)-deficient genetic background. (A) Comparisons of body weight between apoE$^{-/-}$RAGE$^{+/+}$ and apoE$^{-/-}$RAGE$^{-/-}$ mice fed with standard or atherogenic diet. (B) Comparisons of epididymal fat weight between apoE$^{-/-}$RAGE$^{+/+}$ and apoE$^{-/-}$RAGE$^{-/-}$ mice fed with standard or atherogenic diet. (C) Comparisons of adipocyte size in epididymal adipose tissues. Columns represent mean ± standard deviation. P values were analyzed by Student's t-test. Modified from ref [16].

D. RAGE-regulated genes in adipose tissue: ontology analysis

To mine specific group of genes involved in adiposity regulated by RAGE, gene ontology analyses were performed. Downregulated genes in RAGE$^{-/-}$ adipose tissue were significantly accumulated in the ontology terms of metabolic process including acetyl-CoA biosynthetic process, neutral lipid biosynthetic process, pyruvate metabolic process, gluconeogenesis, glycogen biosynthetic process, and NADPH regeneration. Interestingly, genes involved in fat cell differentiation were also identified to be accumulated as down-regulated in RAGE$^{-/-}$ adipose tissue. Ontology terms of glucose transport and neutral amino acid transport were also significantly extracted as downregulated in RAGE$^{-/-}$ adipose tissue. Insulin receptor signaling pathway was a highly significant ontology term downregulated in RAGE$^{-/-}$ adipose tissue. On the contrary, many of the genes upregulated in RAGE$^{-/-}$ adipose tissue were accumulated in ontology terms including cell adhesion, endocytosis, T cell activation, prostaglandin biosynthesis, protein binding, protein folding, processing and glycoprotein biosynthetic process, many of which are known be associated with cellular mechanisms for inflammation and defensive process. Nitrogen compound metabolic process, including amino acid metabolic process, was also identified to be a significant ontology term upregulated in RAGE$^{-/-}$. Interestingly, upregulated genes in RAGE$^{-/-}$ tissue were also significantly accumulated in ontology term for cell redox homeostasis process.

E. RAGE-regulated genes in adipose tissue: pathway analysis

To further identify potential pathways involved in RAGE-regulation of adiposity, KEGG pathway analyses were performed (Table 1). In accordance with the ontology analyses, insulin signaling pathway, pyruvate metabolism, fatty acid biosynthesis and gluconeogenesis were identified to be downregulated pathways in RAGE$^{-/-}$ adipose tissue. PPAR signaling and adipocytokine signaling were also identified to be downregulated in RAGE$^{-/-}$ adipose tissue. Similar to gene ontology analyses, inflammatory pathways including cell adhesion molecules and leukocyte transendothelial migration were the significant pathways upregulated in RAGE$^{-/-}$ mice. Pathways including amino acid metabolic pathways, nitrogen metabolism, glycan biosynthesis, structure and degradation were the pathways significantly upregulated in RAGE$^{-/-}$ adipose tissues.

WT>RAGE-/- (>= 3 fold)	count	P value
Insulin signaling pathway	4/137	0.0002
Fatty acid biosynthesis	1/6	0.0146
ErbB signaling pathway	2/87	0.0179
Ethylbenzene degradation	1/10	0.0242
1- and 2-methylnaphthalene degradation	1/20	0.0478
Jak-STAT signaling pathway	2/151	0.0500
WT>RAGE (>= 2 fold)		
Alanine and aspartate metabolism	7/33	<0.0001
Pyruvate metabolism	7/41	<0.0001
Insulin signaling pathway	12/137	0.0001
Adipocytokine signaling pathway	6/71	0.0059
Fatty acid biosynthesis	2/6	0.0077
Glycerophospholipid metabolism	5/61	0.0134
Glycerolipid metabolism	4/41	0.0148
Glutathione metabolism	4/42	0.0160
Ethylbenzene degradation	2/10	0.0216
Valine, leucine and isoleucine biosynthesis	2/10	0.0216
Type II diabetes mellitus	4/46	0.0218
Metabolism of xenobiotics by cytochrome P450	5/71	0.0245
Sulfur metabolism	2/11	0.0260
PPAR signaling pathway	5/74	0.0287
Propanoate metabolism	3/30	0.0320
Glycolysis / Gluconeogenesis	4/53	0.0345
Circadian rhythm	2/13	0.0358

Table 1. Pathway analyses of the genes differentially expressed in WT vs. RAGE-/- epididymal adipose tissue.

WT<RAGE-/- (>= 2 fold)		
Arginine and proline metabolism	6/33	0.0008
Glycine, serine and threonine metabolism	7/47	0.0010
Glycerolipid metabolism	5/41	0.0128
N-Glycan degradation	3/15	0.0136
Cyanoamino acid metabolism	2/6	0.0163
One carbon pool by folate	3/16	0.0164
Glycosphingolipid biosynthesis - ganglioseries	3/16	0.0164
Polyunsaturated fatty acid biosynthesis	3/17	0.0194
Ether lipid metabolism	4/32	0.0234
Cell adhesion molecules (CAMs)	10/147	0.0301
Prostate cancer	7/88	0.0317
Nitrogen metabolism	3/21	0.0343
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	3/21	0.0343
Glycan structures - biosynthesis 1	8/114	0.0428
WT<RAGE-/- (>= 3 fold)		
Arginine and proline metabolism	4/33	0.0053
Polyunsaturated fatty acid biosynthesis	3/17	0.0054
Glycerolipid metabolism	4/41	0.0114
Leukocyte transendothelial migration	7/115	0.0118
Glutathione metabolism	4/42	0.0124
Cell adhesion molecules (CAMs)	8/147	0.0137
O-Glycan biosynthesis	3/27	0.0199
Thyroid cancer	3/28	0.0219
Glyoxylate and dicarboxylate metabolism	2/14	0.0358
Glycan structures - biosynthesis 1	6/114	0.0364
One carbon pool by folate	2/16	0.0459
Pantothenate and CoA biosynthesis	2/16	0.0459

Table 1. Pathway analyses of the genes differentially expressed in WT vs. RAGE-/- epididymal adipose tissue (continuation).

F. RAGE-regulated genes in adipose tissue: real time RT-PCR confirmation

Adipogenesis related genes including, lipin 1, peroxisome proliferator-activated receptor (PPAR)-γ, adipose differentiation related protein, were shown to be downregulated in RAGE$^{-/-}$ mice. Fatty acid binding protein 5, 1-acylglycerol-3-phosphate O-acyltransferase 2, diacylglycerol O-acyltransferase 2, monoacylglycerol O-acyltransferase 1, acetoacetyl-CoA synthetase, acetyl-coenzyme A carboxylase α were downregulated in RAGE$^{-/-}$ adipose tissue, which could be an essential mechanisms for decreased adiposity in RAGE$^{-/-}$ mice. In insulin signaling, phosphatidylinositol 3-kinase (p85α), adaptor protein with pleckstrin homology and src (APS), sorbin and SH3 domain containing 1 (CAP), insulin receptor substrate (IRS) 1 and 3, thymoma viral proto-oncogene 2 / similar to serine/threonine kinase (Akt), Protein phosphatase 1 regulatory (inhibitor) subunit 3C, facilitated glucose

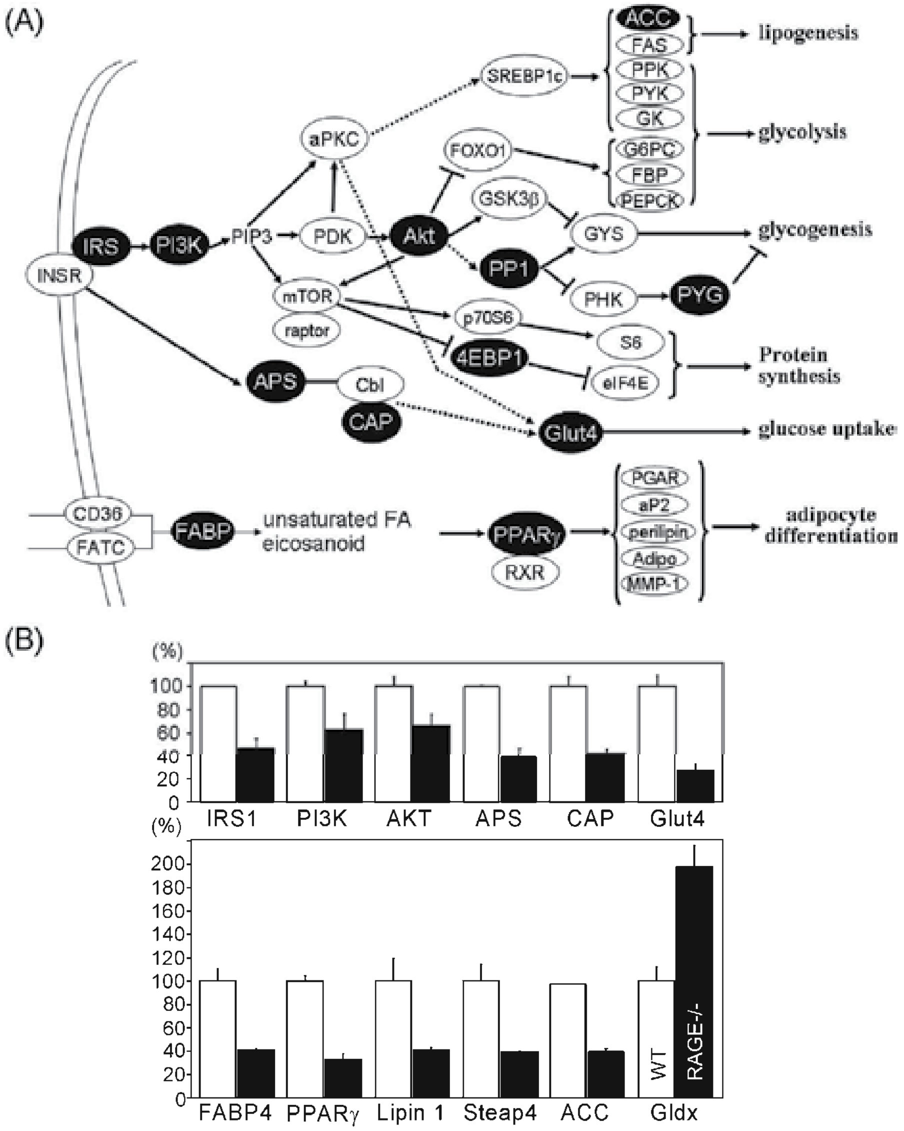

(A)
ACC
FAS
lipogenesis
SREBP1c
PPK
PYK
GK
glycolysis
aPKC
FOXO1
G6PC
FBP
PEPCK
GSK3β
GYS
glycogenesis
IRS
PI3K
PIP3
PDK
Akt
PP1
INSR
mTOR
raptor
PHK
PYG
p70S6
S6
4EBP1
eIF4E
Protein synthesis
APS
Cbl
CAP
Glut4
glucose uptake
PGAR
aP2
perilipin
Adipo
MMP-1
CD36
FATC
FABP
unsaturated FA
eicosanoid
PPARγ
RXR
adipocyte differentiation
(B)
(%)
100
80
60
40
20
0
IRS1
PI3K
AKT
APS
CAP
Glut4
(%)
200
180
160
140
120
100
80
60
40
20
0
FABP4
PPARγ
Lipin 1
Steap4
ACC
Gldx
WT
RAGE-/-

Fig. 5. Gene microarray, Ontology and KEGG pathway analyses suggest that insulin signaling and adipocyte differentiation are the potential pathways regulated by RAGE. (A) Figure summarizes the results of Ontology and KEGG pathway analyses. Genes suppressed in RAGE$^{-/-}$ adipose tissue were described in black circles. (B) Changes in mRNA expression obtained by gene microarray analyses were confirmed by real-time quantitative RT-PCR analyses. All changes in gene expression were statistically significance ($p<0.05$, Student's t-test). IRS-1: insulin receptor substrate 1, PI3K: phosphatidylinositol 3-kinase (p85α), AKT: thymoma viral proto-oncogene 2 / similar to serine/threonine kinase, APS: adaptor protein with pleckstrin homology and src, CAP: sorbin and SH3 domain containing 1, Glut4: facilitated glucose transporter member 4, FABP4: fatty acid binding protein 4, PPAR-γ: peroxisome proliferator-activated receptor, Steap4: six-transmembrane epithelial antigen of prostate 4, ACC: acetyl-coenzyme A carboxylase α, Gldx: glutaredoxin.

transporter member 4 (Glut 4) were identified to be downregulated in RAGE$^{-/-}$ adipose tissue. Figure 5A shows genes specifically suppressed in RAGE$^{-/-}$ adipose tissue (closed circles) in insulin signaling and adipocyte differentiation pathways. Real-time quantitative RT-PCR analyses confirmed the genes in the pathways were indeed down-regulated in RAGE$^{-/-}$ adipose tissue (Figure 5B). These results altogether suggest direct role of RAGE in adiposity. Although in which cell types RAGE is principally working, insulin signaling and adipocyte signaling pathway in adipose tissue appear to play important part in RAGE regulation of adiposity.

4.4 RAGE, endothelial dysfunction and insulin resistance

Impaired insulin action, when assessed by fasting serum insulin levels or the homeostasis model assessment of insulin resistance (HOMA-IR) [92], is associated with atherosclerosis and an increased risk of myocardial infarction. Insulin resistance is associated with endothelial dysfunction [93] and may serve as a link between insulin resistance and atherosclerosis. Recent findings by Harja et al highlighted the involvement of RAGE in endothelial dysfunction [81]. Endothelium-dependent vasorelaxation was tested in isolated mouse aortic rings from *apoE*$^{-/-}$ and *apoE*$^{-/-}$*RAGE*$^{-/-}$ mice, and relaxation response to acetylcholine was significantly improved in the RAGE deficient mouse. Similarly, impaired endothelial function in diabetic obese mice was also shown to be mediated by AGEs/RAGE system, since blockade of AGE-RAGE interaction by soluble RAGE significantly improved endothelial function [94]. Recent clinical observations by Linden et al [44] also implies AGEs/RAGE system is involved in impaired endothelial function in patients with chronic kidney diseases. Thus, not only by the interaction at the cellular signaling level, but RAGE appears to impair endothelial function and potentially blood flow in insulin target tissues, leading to insulin resistance in vivo.

5. C-terminally truncated form of RAGE (soluble RAGE, sRAGE) as potential biomarkers for cardiovascular diseases, metabolic syndrome and insulin resistance

5.1 Truncated form of RAGE

Numerous truncated forms of RAGE have recently been described [12, 95-98] (Figure 1). Two major spliced variants of RAGE mRNA, N-terminal and C-terminal truncated forms, have been most extensively characterized [12]. The N-truncated isoform of RAGE mRNA codes for a 303-amino-acid protein lacking the N-terminal signal sequence and the first V-like extracellular domain. The N-truncated form is incapable of binding with AGEs, since the V-domain is critical for binding of the ligand [71]. The N-truncated form of RAGE appears to be expressed on the cell surface similar to the full-length RAGE, although its biological roles remain to be elucidated [99]. It has been suggested that this form of RAGE could be involved in angiogenic regulation in a fashion independent of the classical RAGE signaling pathway [99].

5.2 Endogenous secretory RAGE (esRAGE)

The C-terminal truncated form of RAGE lacks the exon 10 sequences encoding the transmembrane and intracytoplasmic domains [12]. This spliced variant mRNA of RAGE encodes a protein consisting of 347 amino acids with a 22-amino-acid signal sequence, and is released from cells. This C-truncated form is now known to be present in human circulation and is named endogenous secretory RAGE (esRAGE) [12]. Regulation of alternative splicing of the RAGE is recently shown to be regulated through G-rich cis-elements and heterogenous nuclear ribonucleoprotein H [100]. esRAGE was found to be capable of neutralizing the effects of AGEs on endothelial cells in culture [12]. Adenoviral overexpression of esRAGE in vivo in mice reverses diabetic impairment of vascular dysfunction [101]. Thus, the decoy function of esRAGE may exhibit a feedback mechanism by which esRAGE prevents the activation of RAGE signaling.

5.3 Soluble RAGE generated by shedding

It has also been suggested that some sRAGE isoforms that could act as decoy receptors may be cleaved proteolytically from the native RAGE expressed on the cell surface [102], suggesting heterogeneity of the origin and nature of sRAGE. This proteolytic generation of sRAGE was initially described as occurring in mice [103]. Recent studies suggest that ADAM10 and MMP9 to be involved in RAGE shedding [13, 14]. ADAM is known as a shedase to shed several inflammatory receptors and can be involved in regulation of RAGE/sRAGE balance. A RAGE gene polymorphism is shown to be strongly associated with higher sRAGE levels, although the mechanism by which the polymorphism alters the sRAGE levels remains to be elucidated [104]. Thus, the molecular heterogeneity of the diverse types of sRAGE in human plasma could exert significant protective effects against RAGE-mediated toxicity. However, the endogenous action of sRAGE may not be confined to a decoy function against RAGE-signaling. In HMGB1-induced arthritis model, for example, sRAGE is found to interact with Mac-1, and act as an important proinflammatory and chemotactic molecule [105]. Further analyses are warranted to understand more about the endogenous activity of sRAGE.

5.4 Circulating sRAGE and esRAGE in diseases

A. Circulating sRAGE and cardiovascular diseases

Since sRAGE and esRAGE may be involved in feedback regulation of the toxic effects of RAGE-mediated signaling, recent clinical studies have focused on the potential significance of circulating sRAGE and esRAGE in a variety of pathophysiological conditions, including atherosclerotic disorders, diabetes, hypertension, Alzheimer's dieases and chronic kidney diseases (Table 2). First, Falcone et al [106] reported that total sRAGE levels are significantly lower in patients with angiographically proven coronary artery disease (CAD) than in age-matched healthy controls. The association between circulating sRAGE and angiographic observations was shown to be dose-dependent, with individuals in the lowest quartile of sRAGE exhibiting the highest risk for CAD. Importantly, this cohort consisted of a non-diabetic population, suggesting that the potential significance of sRAGE is not confined to diabetes. Falcone et al also showed that the association between sRAGE and the risk of CAD was independent of other classical risk factors. Their findings are reproduced later by several research groups in larger numbers of subjects, and are also extended to other atherosclerotic diseases, such as carotid atherosclerosis, cerebral ischemia, and aortic valve stenosis (Table 2). Patients with

Alzheimer disease have also lower levels of sRAGE in plasma than patients with vascular dementia and controls, suggesting a role for the RAGE axis in this clinical entity as well [107].

sRAGE		references
CAD (non-DM)	decreased	106, 153, 154
	increased	124
Calcified aortic valve stenosis	decreased	155
Carotid atherosclerosis	decreased	156
Cerebral ischemia	decreased	157 158 127
Alzheimer's disease	decreased	107 159
Endothelial dysfunction	decreased	160
Diabetes (type 1)	increased	122
Diabetes (type 2)	increased	123, 124
	decreased	120, 121
Hypertension	decreased	117
NASH	decreased	118 119
Chronic kidney disease	increased	109, 123, 129, 161 162
Oxidative stress and inflammatory markers	positively associated	163, 164
	inversely associated	121
esRAGE		
Insulin resistance	inversely associated	15
Metabolic syndrome	decreased	15
Diabetes (type 1)	decreased	108, 110
Diabetes (type 2)	decreased	15, 113
Hypertension	decreased	15
NASH	decreased	119
Carotid atherosclerosis	decreased	15, 110-112
	no association	109
CAD	decreased	113, 153
Altzheimer's disease	decreased	165
Chronic kidney disease	increased	114, 161 162

Table 2. Levels of circulating soluble RAGE in cardiovascular and metabolic diseases.

B. Circulating esRAGE and cardiovascular diseases

Following development of an ELISA system to specifically measure human esRAGE [108], we measured plasma esRAGE level and cross-sectionally examined its association with atherosclerosis in 203 type 2 diabetic and 134 non-diabetic age- and gender-matched subjects [15]. esRAGE levels were inversely correlated with carotid and femoral atherosclerosis, as measured as intimal-medial thickness (IMT) by arterial ultrasound. Stepwise regression analyses revealed that plasma esRAGE was the third strongest and an independent factor associated with carotid IMT, following age and systolic blood pressure[15]. Importantly however, when non-diabetic and diabetic groups were separately analyzed, inverse correlation between plasma esRAGE level and IMT was significant in non-diabetic population only, suggesting a potential significance of esRAGE in non-diabetic condition. No association of plasma esRAGE with IMT in diabetes was also reported in other study with 110 Caucasian type 2 diabetic subjects [109]. Another Japanese research group found an inverse correlation between plasma esRAGE and carotid atherosclerosis in type 1 [110] and type 2 diabetic subjects [111]. Recently, the same research group also longitudinally examined the predictive significance of plasma esRAGE and sRAGE on progression of carotid atherosclerosis, and found that low circulating esRAGE level as well as sRAGE level was an independent risk factor for the progression of carotid IMT in type 1 diabetic subjects [112]. In Chinese type 2 diabetic patients, plasma esRAGE is recently shown to be decreased in angiographically-proved patients with coronary artery disease than those without it [113].

C. Low circulating sRAGE as a predictor of cardiovascular diseases

We also reported an observational cohort study in patients with end-stage renal disease (ESRD) and longitudinally evaluated the effect of plasma esRAGE on cardiovascular mortality [114]. The cohort in that study included 206 ESRD subjects, who had been treated by regular hemodialysis for more than 3 months. Even though the plasma esRAGE levels at baseline were higher in ESRD subjects than in those without kidney disease, the subjects in the lowest tertile of plasma esRAGE levels exhibited significantly higher cardiovascular mortality, but not non-cardiovascular mortality. Importantly, even in the subpopulation of non-diabetic subjects alone, low circulating esRAGE level was a predictor of cardiovascular mortality, independent of the other classical risk factors. Thus, low circulating esRAGE or sRAGE level is a potential predictor for atherosclerosis and cardiovascular diseases even in non-diabetic population.

D. Circulating sRAGE, esRAGE and metabolic syndrome

Several components of metabolic syndrome have been shown to be associated with altered plasma sRAGE or esRAGE levels. We first reported that plasma esRAGE levels are already decreased in patients with impaired glucose tolerance as compared with those with normal glucose tolerance (Figure 6A). Moreover, patients with metabolic syndrome showed significantly lower plasma esRAGE than those without it (Figure 6A). Plasma esRAGE levels are inversely correlated with many of the components of metabolic syndrome including body mass index (Figure 6B), blood pressures, fasting plasma glucose, serum triglyceride, and lower HDL-cholesterol levels [15]. The majorities of these correlations remained significant even when the non-diabetic or type 2 diabetic subpopulation was extracted for analyses. An inverse correaltion between esRAGE (or sRAGE) and body mass index was also found for control subjects [115], those with type 1 diabetes [116], and those with ESRD [114]. Patients with hypertension have been found to have lower plasma sRAGE or esRAGE levels [15, 117]. Importantly, our findings also showed that plasma esRAGE was also inversely associated with insulin resistance index, HOMA (Figure 6B), suggesting esRAGE and sRAGE as potential biomarkers for metabolic syndrome and insulin resistance, which could be associated with altered cardiovascular outcomes. Both sRAGE and esRAGE are found to be decreased in patients with liver steatosis [118, 119], which is know to be deeply associated with visceral fat accumulation and insulin resistance.

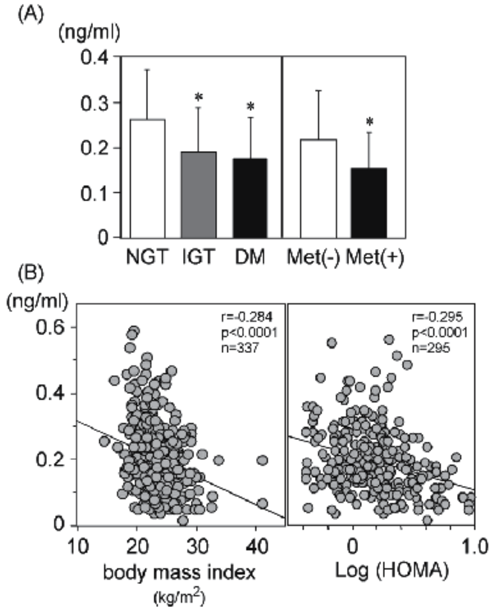


Fig. 6. Plasma esRAGE levels are decreased in glucose intolerance, metabolic syndrome, obesity and insulin resistance (A) Left panel demonstrates the levels of plasma esRAGE in subjects with normal glucose tolerance (NGT) (n=118), impaired glucose tolerance (IGT) (n=16), and type 2 diabetes (DM) (n=203). Right panel compares the plasma esRAGE levels in subjects with (n-53) or without (n=282) metabolic syndrome (Met) as characterized by modified NCEP criteria. * $p<0.05$, ANOVA with multiple comparison (Scheffe's type). (B) Plasma esRAGE levels were inversely associated with body mass index or HOMA insulin resistance index. Logarithm-transformed HOMA index was used for the analyses because of the skewed distribution. Modified from ref [15].

E. Circulating sRAGE and esRAGE in diabetes

The findings regarding plasma levels of the soluble form of RAGE in diabetes are quite confusing. We and other groups have found that plasma esRAGE level is significantly lower in type 1 and type 2 diabetic patients than in non-diabetic controls [15, 110]. Plasma sRAGE levels have also been shown to be decreased in diabetic subjects [120, 121], although conflicting findings have also been reported for type 1 [122] and type 2 diabetes [123, 124]. We examined plasma sRAGE levels by different ELISA system using esRAGE as a standard protein and different sets of antibodies against whole RAGE molecule [125]. In our hand, type 2 diabetic subjects without overt nephropathy (0.60 ± 0.28 ng/ml) exhibited significantly ($p<0.001$, Student's t-test) lower plasma sRAGE level than non-diabetic controls (0.77 ± 0.34 ng/ml) [11]. Of note, when diabetic subjects alone were extracted for analyses, a direct association was not observed between plasma soluble RAGE (both sRAGE and esRAGE) levels and the status of glycemic control (i.e. glycated hemoglobin A1c) [15, 109, 116, 120, 126]. Thus, these complex findings in diabetic subjects suggest that levels of plasma soluble forms of RAGE are not determined simply by status of glycemic control, and that even plasma esRAGE and sRAGE levels may be under the control of distinct mechanisms. Recent study suggests that sRAGE levels may be significantly influenced by ethnicity [127], which may partially explain controversial findings.

F. Circulating sRAGE and esRAGE in CKD

Another important component that can affect plasma sRAGE is the presence of chronic kidney disease. It has been shown that, in peripheral monocytes from subjects with varying severities of CKD, RAGE expression is closely associated with worsening of CKD and is strongly correlated with plasma levels of pentosidine, a marker for AGEs [128]. Circulating sRAGE levels have been shown to be increased in patients with decreased renal function, particularly those with ESRD [109, 123, 129]. Our observations revealed that plasma esRAGE levels in type 2 diabetic subjects without CKD are lower than non-diabetic controls, which is gradually elevated in accordance with progression of CKD [11]. Plasma sRAGE levels in diabetic subjects without CKD also exhibited significantly lower than those of non-diabetic controls [11]. Thus, plasma sRAGE and esRAGE are markedly affected by the presence of CKD, which might make the interpretation of the role of soluble RAGE quite complicated [130]. It remains to be determined whether the increase in plasma esRAGE in CKD is caused by decreased renal function alone or whether esRAGE levels are upregulated to protect against toxic effects of the RAGE ligands. Successful kidney transplantation resulted in significant decrease in plasma sRAGE [131], implying that the kidneys play a role in sRAGE removal.

6. RAGE and Soluble RAGE as a therapeutic target against metabolic syndrome, insulin resistance and cardiovascular disease?

6.1 Soluble RAGE as a therapeutic tool in animal disease models

Potential usefulness of soluble RAGE for prevention and treatment of inflammatory diseases has been demonstrated in many animal models. Blockade of RAGE by administration of genetically engineered sRAGE successfully prevented the development of micro- [132, 133] and macrovascular complications in diabetes [134-136]. We have also shown that adenoviral overexpression of esRAGE successfully restored the impaired angiogenic response in diabetic mice [101]. Sakaguchi et al found that administration of sRAGE markedly suppressed neointimal formation following arterial injury in non-diabetic mice [137]. Soluble RAGE has also been shown to effectively prevent the development of diabetes [138], protect against tumor growth and metastasis [58], improve the outcome of colitis [56], restore impaired wound healing [139], and suppress Alzheimer disease-like conditions [140]. These effects of soluble RAGE in animal models could be explained by its decoy function, inhibiting RAGE interaction with its proinflammatory ligands, which might be applicable to human diseases as well. Since our findings strongly suggest the role of RAGE in adiposity, metabolic syndrome and atheroslcerosis [16], RAGE/soluble RAGE axis could also be a potential therapeutic target against these pathophysiological conditions.

6.2 Potential regulatory mechanisms of circulating soluble RAGE

So far, limited findings are available regarding the mechanisms of regulation of circulating esRAGE or sRAGE in humans. A tissue microarray technique using a wide variety of adult normal human preparations obtained from surgical and autopsy specimens revealed that esRAGE was widely distributed in tissues, including vascular endothelium, monocyte/macrophage, pneumocytes, and several endocrine organs [141]. However, it is unclear at present from which organ or tissue plasma sRAGE or esRAGE originate. Circulating AGEs may be involved in regulation of the secretion or production of soluble RAGE, since AGEs are known to upregulate RAGE expression in vitro [142]. esRAGE could be simultaneously upregulated by AGEs and act as a negative feedback loop to compensate for the damaging effects of AGEs. We and others have found positive correlations between plasma sRAGE or esRAGE and AGEs [11, 114-116, 123]. Significant positive correlation between plasma esRAGE and pentosidine was observed both in hemodialysis and non-hemodialysis subjects [11]. However, plasma CML did not significantly correlated with plasma esRAGE both in hemodialysis and non-hemodialysis subjects. AGEs-mediated regulation of soluble RAGE is also supported by the findings that the suppression of sRAGE expression in diabetic rat kidney is reversed by blockade of AGEs accumulation with alagebrium [143]. Other inflammatory mediators, such as S100, tumor necrosis factor α, and C-reactive protein, could also be potential candidates for regulation of the plasma level of soluble RAGE in humans [120, 142, 144]. Moreover, Geroldi et al [145] showed that high serum sRAGE is associated with extreme longevity, suggesting that understanding the intrinsic regulation of RAGE and soluble RAGE is important for longevity/anti-aging strategies. Without doubt, further understanding of the regulation of soluble RAGE will be most helpful in delineating potential targets for therapeutic application of soluble RAGE.

6.3 Pharmacological agents regulate circulating sRAGE and esRAGE

A. Angiotensin-converting enzyme inhibitor

It would be essential to determine whether currently available pharmacological agents can regulate plasma sRAGE or esRAGE. Potential agents that may affect circulating soluble RAGE include the angiotensin-converting enzyme (ACE) inhibitor [146], thiazolidinediones (TZD) [147] and statins [148-150], which are known to modulate the AGEs-RAGE system in culture. Forbes et al [146] showed that inhibition of angiotensin-converting enzyme (ACE) in rats increased renal expression of sRAGE, and that this was associated with decreases in expression of renal full-length RAGE protein. They also showed that plasma sRAGE levels were significantly increased by inhibition of ACE in both diabetic rats and in human subjects with type 1 diabetes. Thus, one attractive scenario is that the protective effect of ACE inhibition against progression of renal dysfunction is mediated through regulation of RAGE versus soluble RAGE production.

B. Statin

Tam et al recently reported changes in serum levels of sRAGE and esRAGE in archived serum samples from a previous randomized double-blind placebo-controlled clinical trial that explored the cardiovascular effects of atorvastatin in hypercholesterolemic Chinese type 2 diabetic patients, and found that atorvastatin can increase circulating esRAGE levels [150].

C. Thiazolidinedione

For thiazolidinedione, a randomised, open-label, parallel group study was performed with 64 participants randomised to receive add-on therapy with either rosiglitazone or sulfonylurea to examine the effect on plasma soluble RAGE [151]. At 6 months, both rosiglitazone and sulfonylurea resulted in a significant reduction in HbA1c, fasting glucose and AGE. However, significant increases in total sRAGE and esRAGE were only seen in the rosiglitazone group. In a recent study in type 2 diabetes mellitus patients, pioglitazone, but not rosiglitazone, significantly raised sRAGE levels [152], suggesting that all thiazolidinedione may not act similarly. Nevertheless, thiazolidinedione could be one promising candidate which increase circulating levels of esRAGE and sRAGE, and RAGE/soluble RAGE regulation may be involved in thiazolidinedione-mediated improvement of insulin resistance. Finally, we have started the randomized clinical trial comparing the effect of pioglitazone with glimepiride on plasma sRAGE and esRAGE, expression of RAGE on peripheral mononuclear cells, and RAGE shedase gene expression in type 2 diabetic patients (UMIN000002055). This study will be of particular importance to understand the regulatory mechanisms of sRAGE and esRAGE in clinical setting.

7. Summary

The findings discussed here implicated pivotal role of RAGE system in initiation and progression of metabolic syndrome, insulin resistance and atherosclerosis. Provided that continuous RAGE activation represents the concept of "metabolic memory", metabolic syndrome might be conceptualized as memorized long-term subtle inflammation and oxidative stress using RAGE as an inflammatory scaffold. In this system, endogenous inflammatory RAGE ligands may be profoundly involved (Figure 7). Further, sRAGE or esRAGE could serve as a biomarker as well as a therapeutic target for these disease conditions. Obviously there are many missing parts to be veiled to further understand the role of RAGE/soluble RAGE axis in metabolic syndrome and insulin resistance. However, we believe our findings and this concept would open up a new research field which could further precede our understanding of the RAGE biology.

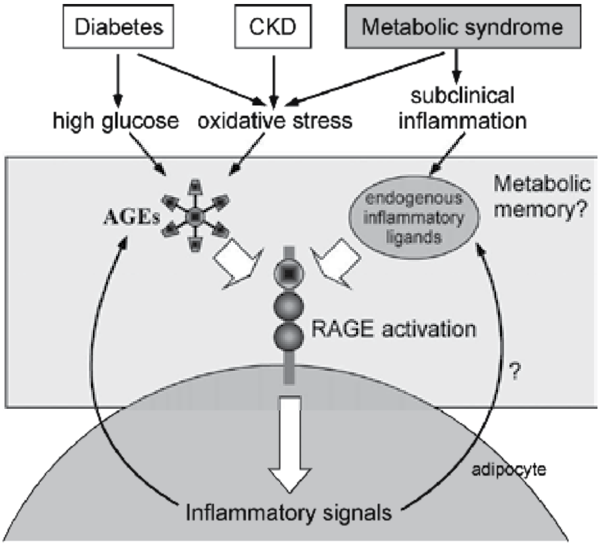


Fig. 7. Metabolic syndrome may be an aspect of "metabolic memory" conceptualized as prolonged RAGE activation through subclinical information.

8. Acknowledgments

The authors thank all colleagues in the Hyogo College of Medicine, Osaka City University Graduate School of Medicine, and Kanazawa University Graduate School of Medical Science for their unflagging support to our projects. We apologize to all colleagues whose work we could not cite other than indirectly through other publications, due to limitation of space. This work was supported in part by a Grant-in-aid for Scientific Research from the Japan Society for the Promotion of Science (20591067 and 23591329 to H.K), and a Grant-in-Aid for Promotion of Technological Seeds in Advanced Medicine, Hyogo College of Medicine (T.Y.).

9. References

[1] The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. *N Engl J Med.* 1993;329(14):977-986.

[2] Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. *Lancet.* 1998;352(9131):837-853.

[3] Patel A, MacMahon S, Chalmers J, Neal B, Billot L, Woodward M, Marre M, Cooper M, Glasziou P, Grobbee D, Hamet P, Harrap S, Heller S, Liu L, Mancia G, Mogensen CE, Pan C, Poulter N, Rodgers A, Williams B, Bompoint S, de Galan BE, Joshi R, Travert F. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *N Engl J Med.* 2008;358(24):2560-2572.

[4] Duckworth W, Abraira C, Moritz T, Reda D, Emanuele N, Reaven PD, Zieve FJ, Marks J, Davis SN, Hayward R, Warren SR, Goldman S, McCarren M, Vitek ME, Henderson WG, Huang GD. Glucose control and vascular complications in veterans with type 2 diabetes. *N Engl J Med.* 2009;360(2):129-139.

[5] Gerstein HC, Miller ME, Byington RP, Goff DC, Jr., Bigger JT, Buse JB, Cushman WC, Genuth S, Ismail-Beigi F, Grimm RH, Jr., Probstfield JL, Simons-Morton DG, Friedewald WT. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med.* 2008;358(24):2545-2559.

[6] Nathan DM, Cleary PA, Backlund JY, Genuth SM, Lachin JM, Orchard TJ, Raskin P, Zinman B. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N Engl J Med.* 2005;353(25):2643-2653.

[7] Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HA. 10-year follow-up of intensive glucose control in type 2 diabetes. *N Engl J Med.* 2008;359(15):1577-1589.

[8] Gaede P, Lund-Andersen H, Parving HH, Pedersen O. Effect of a multifactorial intervention on mortality in type 2 diabetes. *N Engl J Med.* 2008;358(6):580-591.

[9] Wanner C, Krane V, Marz W, Olschewski M, Mann JF, Ruf G, Ritz E. Atorvastatin in patients with type 2 diabetes mellitus undergoing hemodialysis. *N Engl J Med.* 2005;353(3):238-248.

[10] Fellstrom BC, Jardine AG, Schmieder RE, Holdaas H, Bannister K, Beutler J, Chae DW, Chevaile A, Cobbe SM, Gronhagen-Riska C, De Lima JJ, Lins R, Mayer G, McMahon AW, Parving HH, Remuzzi G, Samuelsson O, Sonkodi S, Sci D, Suleymanlar G, Tsakiris D, Tesar V, Todorov V, Wiecek A, Wuthrich RP, Gottlow M, Johnsson E, Zannad F. Rosuvastatin and cardiovascular events in patients undergoing hemodialysis. *N Engl J Med.* 2009;360(14):1395-1407.

[11] Koyama H, Yamamoto H, Nishizawa Y. RAGE and soluble RAGE: potential therapeutic targets for cardiovascular diseases. *Mol Med.* 2007;13(11-12):625-635.

[12] Yonekura H, Yamamoto Y, Sakurai S, Petrova RG, Abedin MJ, Li H, Yasui K, Takeuchi M, Makita Z, Takasawa S, Okamoto H, Watanabe T, Yamamoto H. Novel splice variants of the receptor for advanced glycation end-products expressed in human vascular endothelial cells and pericytes, and their putative roles in diabetes-induced vascular injury. *Biochem J.* 2003;370(Pt 3):1097-1109.

[13] Raucci A, Cugusi S, Antonelli A, Barabino SM, Monti L, Bierhaus A, Reiss K, Saftig P, Bianchi ME. A soluble form of the receptor for advanced glycation endproducts (RAGE) is produced by proteolytic cleavage of the membrane-bound form by the

sheddase a disintegrin and metalloprotease 10 (ADAM10). *FASEB J.* 2008;22(10):3716-3727.

[14] Zhang L, Bukulin M, Kojro E, Roth A, Metz VV, Fahrenholz F, Nawroth PP, Bierhaus A, Postina R. Receptor for advanced glycation end products is subjected to protein ectodomain shedding by metalloproteinases. *J Biol Chem.* 2008;283(51):35507-35516.

[15] Koyama H, Shoji T, Yokoyama H, Motoyama K, Mori K, Fukumoto S, Emoto M, Tamei H, Matsuki H, Sakurai S, Yamamoto Y, Yonekura H, Watanabe T, Yamamoto H, Nishizawa Y. Plasma level of endogenous secretory RAGE is associated with components of the metabolic syndrome and atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2005;25(12):2587-2593.

[16] Ueno H, Koyama H, Shoji T, Monden M, Fukumoto S, Tanaka S, Otsuka Y, Mima Y, Morioka T, Mori K, Shioi A, Yamamoto H, Inaba M, Nishizawa Y. Receptor for advanced glycation end-products (RAGE) regulation of adiposity and adiponectin is associated with atherogenesis in apoE-deficient mouse. *Atherosclerosis.* 2010;211(2):431-436.

[17] Hotamisligil GS. Inflammation and metabolic disorders. *Nature.* 2006;444(7121):860-867.

[18] Monnier VM, Cerami A. Nonenzymatic browning in vivo: possible process for aging of long-lived proteins. *Science.* 1981;211(4481):491-493.

[19] Dyer DG, Blackledge JA, Thorpe SR, Baynes JW. Formation of pentosidine during nonenzymatic browning of proteins by glucose. Identification of glucose and other carbohydrates as possible precursors of pentosidine in vivo. *J Biol Chem.* 1991;266(18):11654-11660.

[20] Thorpe SR, Baynes JW. Maillard reaction products in tissue proteins: new products and new perspectives. *Amino Acids.* 2003;25(3-4):275-281.

[21] Wolff SP, Dean RT. Glucose autoxidation and protein modification. The potential role of 'autoxidative glycosylation' in diabetes. *Biochem J.* 1987;245(1):243-250.

[22] Harding JJ, Beswick HT. The possible contribution of glucose autoxidation to protein modification of diabetes. *Biochem J.* 1988;249(2):617-618.

[23] O'Brien J, Morrissey PA. Nutritional and toxicological aspects of the Maillard browning reaction in foods. *Crit Rev Food Sci Nutr.* 1989;28(3):211-248.

[24] Vlassara H, Striker G. Glycotoxins in the diet promote diabetes and diabetic complications. *Curr Diab Rep.* 2007;7(3):235-241.

[25] Koschinsky T, He CJ, Mitsuhashi T, Bucala R, Liu C, Buenting C, Heitmann K, Vlassara H. Orally absorbed reactive glycation products (glycotoxins): an environmental risk factor in diabetic nephropathy. *Proc Natl Acad Sci U S A.* 1997;94(12):6474-6479.

[26] He C, Sabol J, Mitsuhashi T, Vlassara H. Dietary glycotoxins: inhibition of reactive products by aminoguanidine facilitates renal clearance and reduces tissue sequestration. *Diabetes.* 1999;48(6):1308-1315.

[27] Uribarri J, Peppa M, Cai W, Goldberg T, Lu M, He C, Vlassara H. Restriction of dietary glycotoxins reduces excessive advanced glycation end products in renal failure patients. *J Am Soc Nephrol.* 2003;14(3):728-731.

[28] Che W, Asahi M, Takahashi M, Kaneto H, Okado A, Higashiyama S, Taniguchi N. Selective induction of heparin-binding epidermal growth factor-like growth factor by methylglyoxal and 3-deoxyglucosone in rat aortic smooth muscle cells. The involvement of reactive oxygen species formation and a possible implication for atherogenesis in diabetes. *J Biol Chem.* 1997;272(29):18453-18459.

[29] Fu MX, Requena JR, Jenkins AJ, Lyons TJ, Baynes JW, Thorpe SR. The advanced glycation end product, Nepsilon-(carboxymethyl)lysine, is a product of both lipid peroxidation and glycoxidation reactions. *J Biol Chem.* 1996;271(17):9982-9986.

[30] Wells-Knecht KJ, Brinkmann E, Wells-Knecht MC, Litchfield JE, Ahmed MU, Reddy S, Zyzak DV, Thorpe SR, Baynes JW. New biomarkers of Maillard reaction damage to proteins. *Nephrol Dial Transplant.* 1996;11 Suppl 5:41-47.

[31] Baynes JW, Thorpe SR. Role of oxidative stress in diabetic complications: a new perspective on an old paradigm. *Diabetes.* 1999;48(1):1-9.

[32] Miyata T, van Ypersele de Strihou C, Kurokawa K, Baynes JW. Alterations in nonenzymatic biochemistry in uremia: origin and significance of "carbonyl stress" in long-term uremic complications. *Kidney Int.* 1999;55(2):389-399.

[33] Singh R, Barden A, Mori T, Beilin L. Advanced glycation end-products: a review. *Diabetologia.* 2001;44(2):129-146.

[34] Kaneto H, Fujii J, Myint T, Miyazawa N, Islam KN, Kawasaki Y, Suzuki K, Nakamura M, Tatsumi H, Yamasaki Y, Taniguchi N. Reducing sugars trigger oxidative modification and apoptosis in pancreatic beta-cells by provoking oxidative stress through the glycation reaction. *Biochem J.* 1996;320 (Pt 3):855-863.

[35] Zhao Z, Zhao C, Zhang XH, Zheng F, Cai W, Vlassara H, Ma ZA. Advanced glycation end products inhibit glucose-stimulated insulin secretion through nitric oxide-dependent inhibition of cytochrome c oxidase and adenosine triphosphate synthesis. *Endocrinology.* 2009;150(6):2569-2576.

[36] Kuniyasu A, Ohgami N, Hayashi S, Miyazaki A, Horiuchi S, Nakayama H. CD36-mediated endocytic uptake of advanced glycation end products (AGE) in mouse 3T3-L1 and human subcutaneous adipocytes. *FEBS Lett.* 2003;537(1-3):85-90.

[37] Miele C, Riboulet A, Maitan MA, Oriente F, Romano C, Formisano P, Giudicelli J, Beguinot F, Van Obberghen E. Human glycated albumin affects glucose metabolism in L6 skeletal muscle cells by impairing insulin-induced insulin receptor substrate (IRS) signaling through a protein kinase C alpha-mediated mechanism. *J Biol Chem.* 2003;278(48):47376-47387.

[38] Brownlee M. Negative consequences of glycation. *Metabolism.* 2000;49(2 Suppl 1):9-13.

[39] Peppa M, He C, Hattori M, McEvoy R, Zheng F, Vlassara H. Fetal or neonatal low-glycotoxin environment prevents autoimmune diabetes in NOD mice. *Diabetes.* 2003;52(6):1441-1448.

[40] Lin RY, Choudhury RP, Cai W, Lu M, Fallon JT, Fisher EA, Vlassara H. Dietary glycotoxins promote diabetic atherosclerosis in apolipoprotein E-deficient mice. *Atherosclerosis.* 2003;168(2):213-220.

[41] Stirban A, Negrean M, Stratmann B, Gawlowski T, Horstmann T, Gotting C, Kleesiek K, Mueller-Roesel M, Koschinsky T, Uribarri J, Vlassara H, Tschoepe D. Benfotiamine prevents macro- and microvascular endothelial dysfunction and oxidative stress following a meal rich in advanced glycation end products in individuals with type 2 diabetes. *Diabetes Care.* 2006;29(9):2064-2071.

[42] Negrean M, Stirban A, Stratmann B, Gawlowski T, Horstmann T, Gotting C, Kleesiek K, Mueller-Roesel M, Koschinsky T, Uribarri J, Vlassara H, Tschoepe D. Effects of low- and high-advanced glycation endproduct meals on macro- and microvascular endothelial function and oxidative stress in patients with type 2 diabetes mellitus. *Am J Clin Nutr.* 2007;85(5):1236-1243.

[43] Uribarri J, Stirban A, Sander D, Cai W, Negrean M, Buenting CE, Koschinsky T, Vlassara H. Single oral challenge by advanced glycation end products acutely impairs endothelial function in diabetic and nondiabetic subjects. *Diabetes Care.* 2007;30(10):2579-2582.

[44] Linden E, Cai W, He JC, Xue C, Li Z, Winston J, Vlassara H, Uribarri J. Endothelial dysfunction in patients with chronic kidney disease results from advanced

glycation end products (AGE)-mediated inhibition of endothelial nitric oxide synthase through RAGE activation. *Clin J Am Soc Nephrol.* 2008;3(3):691-698.

[45] Sherwin RS, Kramer KJ, Tobin JD, Insel PA, Liljenquist JE, Berman M, Andres R. A model of the kinetics of insulin in man. *J Clin Invest.* 1974;53(5):1481-1492.

[46] Yang YJ, Hope ID, Ader M, Bergman RN. Insulin transport across capillaries is rate limiting for insulin action in dogs. *J Clin Invest.* 1989;84(5):1620-1628.

[47] Jansson PA, Fowelin JP, von Schenck HP, Smith UP, Lonnroth PN. Measurement by microdialysis of the insulin concentration in subcutaneous interstitial fluid. Importance of the endothelial barrier for insulin. *Diabetes.* 1993;42(10):1469-1473.

[48] Barrett EJ, Eggleston EM, Inyard AC, Wang H, Li G, Chai W, Liu Z. The vascular actions of insulin control its delivery to muscle and regulate the rate-limiting step in skeletal muscle insulin action. *Diabetologia.* 2009;52(5):752-764.

[49] Kubota T, Kubota N, Kumagai H, Yamaguchi S, Kozono H, Takahashi T, Inoue M, Itoh S, Takamoto I, Sasako T, Kumagai K, Kawai T, Hashimoto S, Kobayashi T, Sato M, Tokuyama K, Nishimura S, Tsunoda M, Ide T, Murakami K, Yamazaki T, Ezaki O, Kawamura K, Masuda H, Moroi M, Sugi K, Oike Y, Shimokawa H, Yanagihara N, Tsutsui M, Terauchi Y, Tobe K, Nagai R, Kamata K, Inoue K, Kodama T, Ueki K, Kadowaki T. Impaired insulin signaling in endothelial cells reduces insulin-induced glucose uptake by skeletal muscle. *Cell Metab.* 2011;13(3):294-307.

[50] Hofmann SM, Dong HJ, Li Z, Cai W, Altomonte J, Thung SN, Zeng F, Fisher EA, Vlassara H. Improved insulin sensitivity is associated with restricted intake of dietary glycoxidation products in the db/db mouse. *Diabetes.* 2002;51(7):2082-2089.

[51] Harcourt BE, Sourris KC, Coughlan MT, Walker KZ, Dougherty SL, Andrikopoulos S, Morley AL, Thallas-Bonke V, Chand V, Penfold SA, de Courten MP, Thomas MC, Kingwell BA, Bierhaus A, Cooper ME, Courten BD, Forbes JM. Targeted reduction of advanced glycation improves renal function in obesity. *Kidney Int.* 2011 in press.

[52] Sandu O, Song K, Cai W, Zheng F, Uribarri J, Vlassara H. Insulin resistance and type 2 diabetes in high-fat-fed mice are linked to high glycotoxin intake. *Diabetes.* 2005;54(8):2314-2319.

[53] Unno Y, Sakai M, Sakamoto Y, Kuniyasu A, Nakayama H, Nagai R, Horiuchi S. Advanced glycation end products-modified proteins and oxidized LDL mediate down-regulation of leptin in mouse adipocytes via CD36. *Biochem Biophys Res Commun.* 2004;325(1):151-156.

[54] Riboulet-Chavey A, Pierron A, Durand I, Murdaca J, Giudicelli J, Van Obberghen E. Methylglyoxal impairs the insulin signaling pathways independently of the formation of intracellular reactive oxygen species. *Diabetes.* 2006;55(5):1289-1299.

[55] Unoki H, Bujo H, Yamagishi S, Takeuchi M, Imaizumi T, Saito Y. Advanced glycation end products attenuate cellular insulin sensitivity by increasing the generation of intracellular reactive oxygen species in adipocytes. *Diabetes Res Clin Pract.* 2007;76(2):236-244.

[56] Hofmann MA, Drury S, Fu C, Qu W, Taguchi A, Lu Y, Avila C, Kambham N, Bierhaus A, Nawroth P, Neurath MF, Slattery T, Beach D, McClary J, Nagashima M, Morser J, Stern D, Schmidt AM. RAGE mediates a novel proinflammatory axis: a central cell surface receptor for S100/calgranulin polypeptides. *Cell.* 1999;97(7):889-901.

[57] Hori O, Brett J, Slattery T, Cao R, Zhang J, Chen JX, Nagashima M, Lundh ER, Vijay S, Nitecki D, Morser J, Stern D, Schmidt AM. The receptor for advanced glycation end products (RAGE) is a cellular binding site for amphoterin. Mediation of neurite outgrowth and co-expression of rage and amphoterin in the developing nervous system. *J Biol Chem.* 1995;270(43):25752-25761.

[58] Taguchi A, Blood DC, del Toro G, Canet A, Lee DC, Qu W, Tanji N, Lu Y, Lalla E, Fu C, Hofmann MA, Kislinger T, Ingram M, Lu A, Tanaka H, Hori O, Ogawa S, Stern DM, Schmidt AM. Blockade of RAGE-amphoterin signalling suppresses tumour growth and metastases. *Nature.* 2000;405(6784):354-360.

[59] Yan SD, Chen X, Fu J, Chen M, Zhu H, Roher A, Slattery T, Zhao L, Nagashima M, Morser J, Migheli A, Nawroth P, Stern D, Schmidt AM. RAGE and amyloid-beta peptide neurotoxicity in Alzheimer's disease. *Nature.* 1996;382(6593):685-691.

[60] Sousa MM, Yan SD, Stern D, Saraiva MJ. Interaction of the receptor for advanced glycation end products (RAGE) with transthyretin triggers nuclear transcription factor kB (NF-kB) activation. *Lab Invest.* 2000;80(7):1101-1110.

[61] Chavakis T, Bierhaus A, Al-Fakhri N, Schneider D, Witte S, Linn T, Nagashima M, Morser J, Arnold B, Preissner KT, Nawroth PP. The pattern recognition receptor (RAGE) is a counterreceptor for leukocyte integrins: a novel pathway for inflammatory cell recruitment. *J Exp Med.* 2003;198(10):1507-1515.

[62] Haimoto H, Hosoda S, Kato K. Differential distribution of immunoreactive S100-alpha and S100-beta proteins in normal nonnervous human tissues. *Lab Invest.* 1987;57(5):489-498.

[63] Kato K, Suzuki F, Ogasawara N. Induction of S100 protein in 3T3-L1 cells during differentiation to adipocytes and its liberating by lipolytic hormones. *Eur J Biochem.* 1988;177(2):461-466.

[64] Braga CW, Martinez D, Wofchuk S, Portela LV, Souza DO. S100B and NSE serum levels in obstructive sleep apnea syndrome. *Sleep Med.* 2006;7(5):431-435.

[65] Steiner J, Schiltz K, Walter M, Wunderlich MT, Keilhoff G, Brisch R, Bielau H, Bernstein HG, Bogerts B, Schroeter ML, Westphal S. S100B serum levels are closely correlated with body mass index: an important caveat in neuropsychiatric research. *Psychoneuroendocrinology.* 2010;35(2):321-324.

[66] Steiner J, Walter M, Guest P, Myint AM, Schiltz K, Panteli B, Brauner M, Bernstein HG, Gos T, Herberth M, Schroeter ML, Schwarz MJ, Westphal S, Bahn S, Bogerts B. Elevated S100B levels in schizophrenia are associated with insulin resistance. *Mol Psychiatry.* 2010;15(1):3-4.

[67] Lappalainen T, Kolehmainen M, Schwab U, Pulkkinen L, de Mello VD, Vaittinen M, Laaksonen DE, Poutanen K, Uusitupa M, Gylling H. Gene expression of FTO in human subcutaneous adipose tissue, peripheral blood mononuclear cells and adipocyte cell line. *J Nutrigenet Nutrigenomics.* 2010;3(1):37-45.

[68] Weisberg SP, McCann D, Desai M, Rosenbaum M, Leibel RL, Ferrante AW, Jr. Obesity is associated with macrophage accumulation in adipose tissue. *J Clin Invest.* 2003;112(12):1796-1808.

[69] Xu H, Barnes GT, Yang Q, Tan G, Yang D, Chou CJ, Sole J, Nichols A, Ross JS, Tartaglia LA, Chen H. Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance. *J Clin Invest.* 2003;112(12):1821-1830.

[70] Shoelson SE, Lee J, Goldfine AB. Inflammation and insulin resistance. *J Clin Invest.* 2006;116(7):1793-1801.

[71] Schmidt AM, Vianna M, Gerlach M, Brett J, Ryan J, Kao J, Esposito C, Hegarty H, Hurley W, Clauss M, et al. Isolation and characterization of two binding proteins for advanced glycosylation end products from bovine lung which are present on the endothelial cell surface. *J Biol Chem.* 1992;267(21):14987-14997.

[72] Neeper M, Schmidt AM, Brett J, Yan SD, Wang F, Pan YC, Elliston K, Stern D, Shaw A. Cloning and expression of a cell surface receptor for advanced glycosylation end products of proteins. *J Biol Chem.* 1992;267(21):14998-15004.

[73] Kislinger T, Fu C, Huber B, Qu W, Taguchi A, Du Yan S, Hofmann M, Yan SF, Pischetsrieder M, Stern D, Schmidt AM. N(epsilon)-(carboxymethyl)lysine adducts of proteins are ligands for receptor for advanced glycation end products that activate cell signaling pathways and modulate gene expression. *J Biol Chem.* 1999;274(44):31740-31749.

[74] Reddy S, Bichler J, Wells-Knecht KJ, Thorpe SR, Baynes JW. N epsilon-(carboxymethyl)lysine is a dominant advanced glycation end product (AGE) antigen in tissue proteins. *Biochemistry.* 1995;34(34):10872-10878.

[75] Xie J, Reverdatto S, Frolov A, Hoffmann R, Burz DS, Shekhtman A. Structural basis for pattern recognition by the receptor for advanced glycation end products (RAGE). *J Biol Chem.* 2008;283(40):27255-27269.

[76] Matsumoto S, Yoshida T, Murata H, Harada S, Fujita N, Nakamura S, Yamamoto Y, Watanabe T, Yonekura H, Yamamoto H, Ohkubo T, Kobayashi Y. Solution structure of the variable-type domain of the receptor for advanced glycation end products: new insight into AGE-RAGE interaction. *Biochemistry.* 2008;47(47):12299-12311.

[77] Krieger M, Stern DM. Series introduction: multiligand receptors and human disease. *J Clin Invest.* 2001;108(5):645-647.

[78] Schmidt AM, Yan SD, Yan SF, Stern DM. The multiligand receptor RAGE as a progression factor amplifying immune and inflammatory responses. *J. Clin. Invest.* 2001;108(7):949-955.

[79] Schmidt AM, Hori O, Chen JX, Li JF, Crandall J, Zhang J, Cao R, Yan SD, Brett J, Stern D. Advanced glycation endproducts interacting with their endothelial receptor induce expression of vascular cell adhesion molecule-1 (VCAM-1) in cultured human endothelial cells and in mice. A potential mechanism for the accelerated vasculopathy of diabetes. *J Clin Invest.* 1995;96(3):1395-1403.

[80] Basta G, Lazzerini G, Massaro M, Simoncini T, Tanganelli P, Fu C, Kislinger T, Stern DM, Schmidt AM, De Caterina R. Advanced Glycation End Products Activate Endothelium Through Signal-Transduction Receptor RAGE: A Mechanism for Amplification of Inflammatory Responses. *Circulation.* 2002;105(7):816-822.

[81] Harja E, Bu DX, Hudson BI, Chang JS, Shen X, Hallam K, Kalea AZ, Lu Y, Rosario RH, Oruganti S, Nikolla Z, Belov D, Lalla E, Ramasamy R, Yan SF, Schmidt AM. Vascular and inflammatory stresses mediate atherosclerosis via RAGE and its ligands in apoE-/- mice. *J Clin Invest.* 2008;118(1):183-194.

[82] Hirosumi J, Tuncman G, Chang L, Gorgun CZ, Uysal KT, Maeda K, Karin M, Hotamisligil GS. A central role for JNK in obesity and insulin resistance. *Nature.* 2002;420(6913):333-336.

[83] Tuncman G, Hirosumi J, Solinas G, Chang L, Karin M, Hotamisligil GS. Functional in vivo interactions between JNK1 and JNK2 isoforms in obesity and insulin resistance. *Proc Natl Acad Sci U S A.* 2006;103(28):10741-10746.

[84] Basta G, Lazzerini G, Del Turco S, Ratto GM, Schmidt AM, De Caterina R. At least 2 distinct pathways generating reactive oxygen species mediate vascular cell adhesion molecule-1 induction by advanced glycation end products. *Arterioscler Thromb Vasc Biol.* 2005;25(7):1401-1407.

[85] Yan SD, Schmidt AM, Anderson GM, Zhang J, Brett J, Zou YS, Pinsky D, Stern D. Enhanced cellular oxidant stress by the interaction of advanced glycation end products with their receptors/binding proteins. *J Biol Chem.* 1994;269(13):9889-9897.

[86] Furukawa S, Fujita T, Shimabukuro M, Iwaki M, Yamada Y, Nakajima Y, Nakayama O, Makishima M, Matsuda M, Shimomura I. Increased oxidative stress in obesity and its impact on metabolic syndrome. *J Clin Invest.* 2004;114(12):1752-1761.
[87] Houstis N, Rosen ED, Lander ES. Reactive oxygen species have a causal role in multiple forms of insulin resistance. *Nature.* 2006;440(7086):944-948.
[88] Rodino-Janeiro BK, Salgado-Somoza A, Teijeira-Fernandez E, Gonzalez-Juanatey JR, Alvarez E, Eiras S. Receptor for advanced glycation end-products expression in subcutaneous adipose tissue is related to coronary artery disease. *Eur J Endocrinol.* 2011;164(4):529-537.
[89] Orlova VV, Choi EY, Xie C, Chavakis E, Bierhaus A, Ihanus E, Ballantyne CM, Gahmberg CG, Bianchi ME, Nawroth PP, Chavakis T. A novel pathway of HMGB1-mediated inflammatory cell recruitment that requires Mac-1-integrin. *Embo J.* 2007;26(4):1129-1139.
[90] Schmidt AM, Yan SD, Brett J, Mora R, Nowygrod R, Stern D. Regulation of human mononuclear phagocyte migration by cell surface-binding proteins for advanced glycation end products. *J Clin Invest.* 1993;91(5):2155-2168.
[91] Strissel KJ, Stancheva Z, Miyoshi H, Perfield JW, 2nd, DeFuria J, Jick Z, Greenberg AS, Obin MS. Adipocyte death, adipose tissue remodeling, and obesity complications. *Diabetes.* 2007;56(12):2910-2918.
[92] Uusitupa MI, Niskanen LK, Siitonen O, Voutilainen E, Pyorala K. 5-year incidence of atherosclerotic vascular disease in relation to general risk factors, insulin level, and abnormalities in lipoprotein composition in non-insulin-dependent diabetic and nondiabetic subjects. *Circulation.* 1990;82(1):27-36.
[93] Steinberg HO, Chaker H, Leaming R, Johnson A, Brechtel G, Baron AD. Obesity/insulin resistance is associated with endothelial dysfunction. Implications for the syndrome of insulin resistance. *J Clin Invest.* 1996;97(11):2601-2610.
[94] Gao X, Zhang H, Schmidt AM, Zhang C. AGE/RAGE Produces Endothelial Dysfunction in Coronary Arterioles in Type II Diabetic Mice. *Am J Physiol Heart Circ Physiol.* 2008; 295(2): H491-498.
[95] Malherbe P, Richards JG, Gaillard H, Thompson A, Diener C, Schuler A, Huber G. cDNA cloning of a novel secreted isoform of the human receptor for advanced glycation end products and characterization of cells co-expressing cell-surface scavenger receptors and Swedish mutant amyloid precursor protein. *Brain Res Mol Brain Res.* 1999; 71(2):159-170.
[96] Schlueter C, Hauke S, Flohr AM, Rogalla P, Bullerdiek J. Tissue-specific expression patterns of the RAGE receptor and its soluble forms--a result of regulated alternative splicing? *Biochim Biophys Acta.* 2003;1630(1):1-6.
[97] Park IH, Yeon SI, Youn JH, Choi JE, Sasaki N, Choi IH, Shin JS. Expression of a novel secreted splice variant of the receptor for advanced glycation end products (RAGE) in human brain astrocytes and peripheral blood mononuclear cells. *Mol Immunol.* 2004;40(16):1203-1211.
[98] Ding Q, Keller JN. Splice variants of the receptor for advanced glycosylation end products (RAGE) in human brain. *Neurosci Lett.* 2005;373(1):67-72.
[99] Bierhaus A, Humpert PM, Morcos M, Wendt T, Chavakis T, Arnold B, Stern DM, Nawroth PP. Understanding RAGE, the receptor for advanced glycation end products. *J Mol Med.* 2005;83(11):876-886.
[100] Ohe K, Watanabe T, Harada S, Munesue S, Yamamoto Y, Yonekura H, Yamamoto H. Regulation of alternative splicing of the receptor for advanced glycation

endproducts (RAGE) through G-rich cis-elements and heterogenous nuclear ribonucleoprotein H. *J Biochem.* 2010;147(5):651-659.

[101] Shoji T, Koyama H, Morioka T, Tanaka S, Kizu A, Motoyama K, Mori K, Fukumoto S, Shioi A, Shimogaito N, Takeuchi M, Yamamoto Y, Yonekura H, Yamamoto H, Nishizawa Y. Receptor for advanced glycation end products is involved in impaired angiogenic response in diabetes. *Diabetes.* 2006;55(8):2245-2255.

[102] Hudson BI, Harja E, Moser B, Schmidt AM. Soluble levels of receptor for advanced glycation endproducts (sRAGE) and coronary artery disease: the next C-reactive protein? *Arterioscler Thromb Vasc Biol.* 2005;25(5):879-882.

[103] Hanford LE, Enghild JJ, Valnickova Z, Petersen SV, Schaefer LM, Schaefer TM, Reinhart TA, Oury TD. Purification and characterization of mouse soluble receptor for advanced glycation end products (sRAGE). *J Biol Chem.* 2004;279(48):50019-50024.

[104] Gaens KH, Ferreira I, van der Kallen CJ, van Greevenbroek MM, Blaak EE, Feskens EJ, Dekker JM, Nijpels G, Heine RJ, t Hart LM, de Groot PG, Stehouwer CD, Schalkwijk CG. Association of polymorphism in the receptor for advanced glycation end products (RAGE) gene with circulating RAGE levels. *J Clin Endocrinol Metab.* 2009;94(12):5174-5180.

[105] Pullerits R, Brisslert M, Jonsson IM, Tarkowski A. Soluble receptor for advanced glycation end products triggers a proinflammatory cytokine cascade via beta2 integrin Mac-1. *Arthritis Rheum.* 2006;54(12):3898-3907.

[106] Falcone C, Emanuele E, D'Angelo A, Buzzi MP, Belvito C, Cuccia M, Geroldi D. Plasma levels of soluble receptor for advanced glycation end products and coronary artery disease in nondiabetic men. *Arterioscler Thromb Vasc Biol.* 2005;25(5):1032-1037.

[107] Emanuele E, D'Angelo A, Tomaino C, Binetti G, Ghidoni R, Politi P, Bernardi L, Maletta R, Bruni AC, Geroldi D. Circulating levels of soluble receptor for advanced glycation end products in Alzheimer disease and vascular dementia. *Arch Neurol.* 2005;62(11):1734-1736.

[108] Sakurai S, Yamamoto Y, Tamei H, Matsuki H, Obata K, Hui L, Miura J, Osawa M, Uchigata Y, Iwamoto Y, Watanabe T, Yonekura H, Yamamoto H. Development of an ELISA for esRAGE and its application to type 1 diabetic patients. *Diabetes Res Clin Pract.* 2006;73(2):158-165.

[109] Humpert PM, Djuric Z, Kopf S, Rudofsky G, Morcos M, Nawroth PP, Bierhaus A. Soluble RAGE but not endogenous secretory RAGE is associated with albuminuria in patients with type 2 diabetes. *Cardiovasc Diabetol.* 2007;6:9.

[110] Katakami N, Matsuhisa M, Kaneto H, Matsuoka TA, Sakamoto K, Nakatani Y, Ohtoshi K, Hayaishi-Okano R, Kosugi K, Hori M, Yamasaki Y. Decreased endogenous secretory advanced glycation end product receptor in type 1 diabetic patients: its possible association with diabetic vascular complications. *Diabetes Care.* 2005;28(11):2716-2721.

[111] Katakami N, Matsuhisa M, Kaneto H, Yamasaki Y. Serum endogenous secretory RAGE levels are inversely associated with carotid IMT in type 2 diabetic patients. *Atherosclerosis.* 2007;190(1):22-23.

[112] Katakami N, Matsuhisa M, Kaneto H, Matsuoka TA, Sakamoto K, Yasuda T, Umayahara Y, Kosugi K, Yamasaki Y. Serum endogenous secretory RAGE level is an independent risk factor for the progression of carotid atherosclerosis in type 1 diabetes. *Atherosclerosis.* 2009;204(1):288-292.

[113] Lu L, Pu LJ, Zhang Q, Wang LJ, Kang S, Zhang RY, Chen QJ, Wang JG, De Caterina R, Shen WF. Increased glycated albumin and decreased esRAGE levels are related to angiographic severity and extent of coronary artery disease in patients with type 2 diabetes. *Atherosclerosis.* 2009;206(2):540-545.

[114] Koyama H, Shoji T, Fukumoto S, Shinohara K, Emoto M, Mori K, Tahara H, Ishimura E, Kakiya R, Tabata T, Yamamoto H, Nishizawa Y. Low circulating endogenous secretory receptor for AGEs predicts cardiovascular mortality in patients with end-stage renal disease. *Arterioscler Thromb Vasc Biol.* 2007;27(1):147-153.

[115] Yamagishi S, Adachi H, Nakamura K, Matsui T, Jinnouchi Y, Takenaka K, Takeuchi M, Enomoto M, Furuki K, Hino A, Shigeto Y, Imaizumi T. Positive association between serum levels of advanced glycation end products and the soluble form of receptor for advanced glycation end products in nondiabetic subjects. *Metabolism.* 2006;55(9):1227-1231.

[116] Miura J, Yamamoto Y, Osawa M, Watanabe T, Yonekura H, Uchigata Y, Yamamoto H, Iwamoto Y. Endogenous secretory receptor for advanced glycation endproducts levels are correlated with serum pentosidine and CML in patients with type 1 diabetes. *Arterioscler Thromb Vasc Biol.* 2007;27(1):253-254.

[117] Geroldi D, Falcone C, Emanuele E, D'Angelo A, Calcagnino M, Buzzi MP, Scioli GA, Fogari R. Decreased plasma levels of soluble receptor for advanced glycation end-products in patients with essential hypertension. *J Hypertens.* 2005;23(9):1725-1729.

[118] Yilmaz Y, Ulukaya E, Gul OO, Arabul M, Gul CB, Atug O, Oral AY, Aker S, Dolar E. Decreased plasma levels of soluble receptor for advanced glycation endproducts (sRAGE) in patients with nonalcoholic fatty liver disease. *Clin Biochem.* 2009;42(9):802-807.

[119] D'Adamo E, Giannini C, Chiavaroli V, de Giorgis T, Verrotti A, Chiarelli F, Mohn A. What is the significance of soluble and endogenous secretory receptor for advanced glycation end products in liver steatosis in obese prepubertal children? *Antioxid Redox Signal.* 2011;14(6):1167-1172.

[120] Basta G, Sironi AM, Lazzerini G, Del Turco S, Buzzigoli E, Casolaro A, Natali A, Ferrannini E, Gastaldelli A. Circulating soluble receptor for advanced glycation end products is inversely associated with glycemic control and S100A12 protein. *J Clin Endocrinol Metab.* 2006;91(11):4628-4634.

[121] Devangelio E, Santilli F, Formoso G, Ferroni P, Bucciarelli L, Michetti N, Clissa C, Ciabattoni G, Consoli A, Davi G. Soluble RAGE in type 2 diabetes: association with oxidative stress. *Free Radic Biol Med.* 2007;43(4):511-518.

[122] Challier M, Jacqueminet S, Benabdesselam O, Grimaldi A, Beaudeux JL. Increased serum concentrations of soluble receptor for advanced glycation endproducts in patients with type 1 diabetes. *Clin Chem.* 2005;51(9):1749-1750.

[123] Tan KC, Shiu SW, Chow WS, Leng L, Bucala R, Betteridge DJ. Association between serum levels of soluble receptor for advanced glycation end products and circulating advanced glycation end products in type 2 diabetes. *Diabetologia.* 2006;49(11):2756-2762.

[124] Nakamura K, Yamagishi SI, Adachi H, Kurita-Nakamura Y, Matsui T, Yoshida T, Sato A, Imaizumi T. Elevation of soluble form of receptor for advanced glycation end products (sRAGE) in diabetic subjects with coronary artery disease. *Diabetes Metab Res Rev.* 2007;23(5): 368-371.

[125] Yamamoto Y, Miura J, Sakurai S, Watanabe T, Yonekura H, Tamei H, Matsuki H, Obata KI, Uchigata Y, Iwamoto Y, Koyama H, Yamamoto H. Assaying soluble

forms of receptor for advanced glycation end products. *Arterioscler Thromb Vasc Biol.* 2007;27(6):e33-e34.

[126] Humpert PM, Kopf S, Djuric Z, Wendt T, Morcos M, Nawroth PP, Bierhaus A. Plasma sRAGE is independently associated with urinary albumin excretion in type 2 diabetes. *Diabetes Care.* 2006;29(5):1111-1113.

[127] Hudson BI, Moon YP, Kalea AZ, Khatri M, Marquez C, Schmidt AM, Paik MC, Yoshita M, Sacco RL, Decarli C, Wright CB, Elkind MS. Association of serum soluble Receptor for Advanced Glycation End-products with subclinical cerebrovascular disease: The Northern Manhattan Study (NOMAS). *Atherosclerosis.* 2011 in press.

[128] Hou FF, Ren H, Owen WF, Jr, Guo ZJ, Chen PY, Schmidt AM, Miyata T, Zhang X. Enhanced expression of receptor for advanced glycation end products in chronic kidney disease. *J Am Soc Nephrol.* 2004;15(7):1889-1896.

[129] Kalousova M, Hodkova M, Kazderova M, Fialova J, Tesar V, Dusilova-Sulkova S, Zima T. Soluble receptor for advanced glycation end products in patients with decreased renal function. *Am J Kidney Dis.* 2006;47(3):406-411.

[130] Nin JW, Jorsal A, Ferreira I, Schalkwijk CG, Prins MH, Parving HH, Tarnow L, Rossing P, Stehouwer CD. Higher plasma soluble Receptor for Advanced Glycation End Products (sRAGE) levels are associated with incident cardiovascular disease and all-cause mortality in type 1 diabetes: a 12-year follow-up study. *Diabetes.* 2010;59(8):2027-2032.

[131] Kalousova M, Bartosova K, Zima T, Skibova J, Teplan V, Viklicky O. Pregnancy-associated plasma protein a and soluble receptor for advanced glycation end products after kidney transplantation. *Kidney Blood Press Res.* 2007;30(1):31-37.

[132] Wendt TM, Tanji N, Guo J, Kislinger TR, Qu W, Lu Y, Bucciarelli LG, Rong LL, Moser B, Markowitz GS, Stein G, Bierhaus A, Liliensiek B, Arnold B, Nawroth PP, Stern DM, D'Agati VD, Schmidt AM. RAGE drives the development of glomerulosclerosis and implicates podocyte activation in the pathogenesis of diabetic nephropathy. *Am J Pathol.* 2003;162(4):1123-1137.

[133] Bierhaus A, Haslbeck KM, Humpert PM, Liliensiek B, Dehmer T, Morcos M, Sayed AA, Andrassy M, Schiekofer S, Schneider JG, Schulz JB, Heuss D, Neundorfer B, Dierl S, Huber J, Tritschler H, Schmidt AM, Schwaninger M, Haering HU, Schleicher E, Kasper M, Stern DM, Arnold B, Nawroth PP. Loss of pain perception in diabetes is dependent on a receptor of the immunoglobulin superfamily. *J Clin Invest.* 2004;114(12):1741-1751.

[134] Park L, Raman KG, Lee KJ, Lu Y, Ferran LJ, Jr., Chow WS, Stern D, Schmidt AM. Suppression of accelerated diabetic atherosclerosis by the soluble receptor for advanced glycation endproducts. *Nat Med.* 1998;4(9):1025-1031.

[135] Kislinger T, Tanji N, Wendt T, Qu W, Lu Y, Ferran LJ, Jr., Taguchi A, Olson K, Bucciarelli L, Goova M, Hofmann MA, Cataldegirmen G, D'Agati V, Pischetsrieder M, Stern DM, Schmidt AM. Receptor for advanced glycation end products mediates inflammation and enhanced expression of tissue factor in vasculature of diabetic apolipoprotein E-null mice. *Arterioscler Thromb Vasc Biol.* 2001;21(6):905-910.

[136] Bucciarelli LG, Wendt T, Qu W, Lu Y, Lalla E, Rong LL, Goova MT, Moser B, Kislinger T, Lee DC, Kashyap Y, Stern DM, Schmidt AM. RAGE blockade stabilizes established atherosclerosis in diabetic apolipoprotein E-null mice. *Circulation.* 2002;106(22):2827-2835.

[137] Sakaguchi T, Yan SF, Yan SD, Belov D, Rong LL, Sousa M, Andrassy M, Marso SP, Duda S, Arnold B, Liliensiek B, Nawroth PP, Stern DM, Schmidt AM, Naka Y.

Central role of RAGE-dependent neointimal expansion in arterial restenosis. *J Clin Invest.* 2003;111(7):959-972.

[138] Chen Y, Yan SS, Colgan J, Zhang HP, Luban J, Schmidt AM, Stern D, Herold KC. Blockade of late stages of autoimmune diabetes by inhibition of the receptor for advanced glycation end products. *J Immunol.* 2004;173(2):1399-1405.

[139] Goova MT, Li J, Kislinger T, Qu W, Lu Y, Bucciarelli LG, Nowygrod S, Wolf BM, Caliste X, Yan SF, Stern DM, Schmidt AM. Blockade of receptor for advanced glycation end-products restores effective wound healing in diabetic mice. *Am J Pathol.* 2001;159(2):513-525.

[140] Arancio O, Zhang HP, Chen X, Lin C, Trinchese F, Puzzo D, Liu S, Hegde A, Yan SF, Stern A, Luddy JS, Lue LF, Walker DG, Roher A, Buttini M, Mucke L, Li W, Schmidt AM, Kindy M, Hyslop PA, Stern DM, Du Yan SS. RAGE potentiates Abeta-induced perturbation of neuronal function in transgenic mice. *Embo J.* 2004;23(20):4096-4105.

[141] Cheng C, Tsuneyama K, Kominami R, Shinohara H, Sakurai S, Yonekura H, Watanabe T, Takano Y, Yamamoto H, Yamamoto Y. Expression profiling of endogenous secretory receptor for advanced glycation end products in human organs. *Mod Pathol.* 2005;18(10):1385-1396.

[142] Tanaka N, Yonekura H, Yamagishi S, Fujimori H, Yamamoto Y, Yamamoto H. The receptor for advanced glycation end products is induced by the glycation products themselves and tumor necrosis factor-alpha through nuclear factor-kappa B, and by 17beta-estradiol through Sp-1 in human vascular endothelial cells. *J Biol Chem.* 2000;275(33):25781-25790.

[143] Coughlan MT, Thallas-Bonke V, Pete J, Long DM, Gasser A, Tong DC, Arnstein M, Thorpe SR, Cooper ME, Forbes JM. Combination therapy with the advanced glycation end product cross-link breaker, alagebrium, and angiotensin converting enzyme inhibitors in diabetes: synergy or redundancy? *Endocrinology.* 2007;148(2):886-895.

[144] Zhong Y, Li SH, Liu SM, Szmitko PE, He XQ, Fedak PW, Verma S. C-Reactive protein upregulates receptor for advanced glycation end products expression in human endothelial cells. *Hypertension.* 2006;48(3):504-511.

[145] Geroldi D, Falcone C, Minoretti P, Emanuele E, Arra M, D'Angelo A. High levels of soluble receptor for advanced glycation end products may be a marker of extreme longevity in humans. *J Am Geriatr Soc.* 2006;54(7):1149-1150.

[146] Forbes JM, Thorpe SR, Thallas-Bonke V, Pete J, Thomas MC, Deemer ER, Bassal S, El-Osta A, Long DM, Panagiotopoulos S, Jerums G, Osicka TM, Cooper ME. Modulation of soluble receptor for advanced glycation end products by angiotensin-converting enzyme-1 inhibition in diabetic nephropathy. *J Am Soc Nephrol.* 2005;16(8):2363-2372.

[147] Marx N, Walcher D, Ivanova N, Rautzenberg K, Jung A, Friedl R, Hombach V, de Caterina R, Basta G, Wautier MP, Wautiers JL. Thiazolidinediones reduce endothelial expression of receptors for advanced glycation end products. *Diabetes.* 2004;53(10):2662-2668.

[148] Okamoto T, Yamagishi S, Inagaki Y, Amano S, Koga K, Abe R, Takeuchi M, Ohno S, Yoshimura A, Makita Z. Angiogenesis induced by advanced glycation end products and its prevention by cerivastatin. *Faseb J.* 2002;16(14):1928-1930.

[149] Cuccurullo C, Iezzi A, Fazia ML, De Cesare D, Di Francesco A, Muraro R, Bei R, Ucchino S, Spigonardo F, Chiarelli F, Schmidt AM, Cuccurullo F, Mezzetti A,

Cipollone F. Suppression of RAGE as a basis of simvastatin-dependent plaque stabilization in type 2 diabetes. *Arterioscler Thromb Vasc Biol.* 2006;26(12):2716-2723.

[150] Tam HL, Shiu SW, Wong Y, Chow WS, Betteridge DJ, Tan KC. Effects of atorvastatin on serum soluble receptors for advanced glycation end-products in type 2 diabetes. *Atherosclerosis.* 2010;209(1):173-177.

[151] Tan KC, Chow WS, Tso AW, Xu A, Tse HF, Hoo RL, Betteridge DJ, Lam KS. Thiazolidinedione increases serum soluble receptor for advanced glycation end-products in type 2 diabetes. *Diabetologia.* 2007;50(9):1819-1825.

[152] Oz Gul O, Tuncel E, Yilmaz Y, Ulukaya E, Gul CB, Kiyici S, Oral AY, Guclu M, Ersoy C, Imamoglu S. Comparative effects of pioglitazone and rosiglitazone on plasma levels of soluble receptor for advanced glycation end products in type 2 diabetes mellitus patients. *Metabolism.* 2010;59(1):64-69.

[153] Yan XX, Lu L, Peng WH, Wang LJ, Zhang Q, Zhang RY, Chen QJ, Shen WF. Increased serum HMGB1 level is associated with coronary artery disease in nondiabetic and type 2 diabetic patients. *Atherosclerosis.* 2009;205(2):544-548.

[154] Mahajan N, Malik N, Bahl A, Sharma Y, Dhawan V. Correlation among soluble markers and severity of disease in non-diabetic subjects with pre-mature coronary artery disease. *Mol Cell Biochem.* 2009;330(1-2):201-209.

[155] Basta G, Corciu AI, Vianello A, Del Turco S, Foffa I, Navarra T, Chiappino D, Berti S, Mazzone A. Circulating soluble receptor for advanced glycation end-product levels are decreased in patients with calcific aortic valve stenosis. *Atherosclerosis.* 2011 in press;210(2):614-618.

[156] Basta G, Leonardis D, Mallamaci F, Cutrupi S, Pizzini P, Gaetano L, Tripepi R, Tripepi G, De Caterina R, Zoccali C. Circulating soluble receptor of advanced glycation end product inversely correlates with atherosclerosis in patients with chronic kidney disease. *Kidney Int.* 2010;77(3):225-231.

[157] Yokota C, Minematsu K, Tomii Y, Naganuma M, Ito A, Nagasawa H, Yamaguchi T. Low levels of plasma soluble receptor for advanced glycation end products are associated with severe leukoaraiosis in acute stroke patients. *J Neurol Sci.* 2009;287(1-2):41-44.

[158] Park HY, Yun KH, Park DS. Levels of Soluble Receptor for Advanced Glycation End Products in Acute Ischemic Stroke without a Source of Cardioembolism. *J Clin Neurol.* 2009;5(3):126-132.

[159] Yao L, Li K, Zhang L, Yao S, Piao Z, Song L. Influence of the Pro12Ala polymorphism of PPAR-gamma on age at onset and sRAGE levels in Alzheimer's disease. *Brain Res.* 2009;1291:133-139.

[160] Chiang KH, Huang PH, Huang SS, Wu TC, Chen JW, Lin SJ. Plasma levels of soluble receptor for advanced glycation end products are associated with endothelial function and predict cardiovascular events in nondiabetic patients. *Coron Artery Dis.* 2009;20(4):267-273.

[161] Kalousova M, Jachymova M, Mestek O, Hodkova M, Kazderova M, Tesar V, Zima T. Receptor for advanced glycation end products--soluble form and gene polymorphisms in chronic haemodialysis patients. *Nephrol Dial Transplant.* 2007;22(7):2020-2026.

[162] Semba RD, Ferrucci L, Fink JC, Sun K, Beck J, Dalal M, Guralnik JM, Fried LP. Advanced glycation end products and their circulating receptors and level of kidney function in older community-dwelling women. *Am J Kidney Dis.* 2009;53(1):51-58.

[163] Nakamura K, Yamagishi S, Adachi H, Kurita-Nakamura Y, Matsui T, Yoshida T, Imaizumi T. Serum levels of sRAGE, the soluble form of receptor for advanced glycation end products, are associated with inflammatory markers in patients with type 2 diabetes. *Mol Med.* 2007;13(3-4):185-189.

[164] Nakamura K, Yamagishi SI, Adachi H, Matsui T, Kurita-Nakamura Y, Takeuchi M, Inoue H, Imaizumi T. Circulating advanced glycation end products (AGEs) and soluble form of receptor for AGEs (sRAGE) are independent determinants of serum monocyte chemoattractant protein-1 (MCP-1) levels in patients with type 2 diabetes. *Diabetes Metab Res Rev.* 2007.

[165] Nozaki I, Watanabe T, Kawaguchi M, Akatsu H, Tsuneyama K, Yamamoto Y, Ohe K, Yonekura H, Yamada M, Yamamoto H. Reduced expression of endogenous secretory receptor for advanced glycation endproducts in hippocampal neurons of Alzheimer's disease brains. *Arch Histol Cytol.* 2007;70(5):279-290.

Mitochondria Function in Diabetes – From Health to Pathology – New Perspectives for Treatment of Diabetes-Driven Disorders

Magdalena Labieniec-Watala[1*], Karolina Siewiera[1],
Slawomir Gierszewski[1] and Cezary Watala[2]
[1]*University of Lodz, Department of Thermobiology,*
[2]*Medical University of Lodz, Department of Haemostasis and Haemostatic Disorders,*
Poland

1. Introduction

A few words about mitochondria ….

Mitochondria are active intracellular structures that collide, divide, and fuse with other mitochondria. Mitochondria exist as branched-chain reticulum networks, but can also exist as punctuated structures. Their distribution within the cells is quite diverse, is regulated by the interactions with the cytoskeleton and maintained by the balance between mitochondrial fusion and fission.

Mitochondria play crucial physiological functions that underlie the distortions of fragile balance between health and disease. In recent years, the role of mitochondria has gained much interest in the field of diabetic pathology since mitochondrial abnormalities were found in insulin resistance and in both types of diabetes.

A few words about hyperglycaemia and diabetes….

Hyperglycaemia, resulting from uncontrolled regulation of glucose metabolism, is widely recognized as the causal link between diabetes and diabetic complications. Diabetes mellitus has been classified into two forms. Type 1 diabetes, which accounts for about 10% of all cases of diabetes, is caused by autoimmune destruction of pancreatic β-cells that implies insulin deficiency. Type 2 diabetes, the more prevalent form of diabetes, is considered a heterogeneous disease due to the multiplicity of factors that cause the observed phenotype. It results from the combination of insulin resistance and/or a β-cell secretory defects. The explosive increase in the prevalence of type 2 diabetes is predicted in the nearest future. It is considered that about 220 millions people all over the world may suffer from the condition, and an equal number is thought to be "prediabetic", having early symptoms and not yet full manifestation of the disease.

A few words about impact of hyperglycaemia on mitochondria …..

Hyperglycaemia has been indicated as one of the main causes of altered mitochondrial function in diabetic individuals (animals and humans). Rather huge variation in the extent of damage has been observed and attributed to both the type and duration of diabetes studied. General consensus points out to the formation of glycation products as a crucial mechanism, by which hyperglycaemia affects mitochondrial and cellular function in diabetes.

Hyperglycaemia elicits an increased ROS production, presumably to the major extent originating from mitochondrial respiratory chain. ROS play a central role in mediating various metabolic defects associated with a diabetic state. Therefore, the inhibition of ROS production and/or enhancement of ROS scavenging might prove to be beneficial therapies. Alterations in metabolic regulators and glucose-stimulated insulin secretion are also associated with mitochondrial dysfunction in diabetes.

Impairments in mitochondrial function are intrinsically related to diabetes. The prevailing hypothesis is that hyperglycaemia-induced increase in electron transfer donors (NADH and $FADH_2$) may increase electron flux through the mitochondrial electron transport chain. Consequently, the ATP/ADP ratio and hyperpolarisation of the mitochondrial membrane (electrochemical potential difference) also become increased. This high electrochemical potential difference generated by the proton gradient leads to partial inhibition of the electron transport in the complex III, resulting in the augmented electron flow towards coenzyme Q. In turn, this drives partial reduction of O_2 to generate the free radical anion superoxide. It is accelerated reduction of coenzyme Q and generation of ROS that are believed to constitute the fundamental source for mitochondrial dysfunction that plays a critical role in diabetes-related metabolic disorders and tissue histopathology.

A few words about oxidative stress…..

Oxidative stress has been implicated as a major contributor to both the onset and the progression of diabetes and its associated complications. Some of the consequences of an oxidative environment may be the development of insulin resistance, β-cell dysfunction, impaired glucose tolerance, and mitochondrial dysfunction, which can ultimately lead to the diabetic disease state. Experimental and clinical data suggest an inverse association between insulin sensitivity and ROS levels. Oxidative stress can arise from a number of different sources, like the disease state or lifestyle, including episodes of ketosis, sleep restriction, and excessive nutrient intake. Oxidative stress can be reduced by controlling calorie intake, hyperglycaemia and mitochondrial metabolism.

It is now established that 90% of intracellular ROS are generated by mitochondria. The mitochondrial respiratory chain is the principal source of cellular oxygen radicals (ROS), such as superoxide anion radicals and hydroxyl radicals. The primary factor governing mitochondrial ROS generation is the redox state of the respiratory chain. If the membrane potential across the inner mitochondrial membrane rises above a certain threshold value, a massive stimulation of ROS generation occurs. Electrons leak mainly from the complexes I and III of the electron transport chain (ETC) and thereby generate incompletely reduced forms of oxygen. The rise in a membrane potential may occur as a consequence of augmented delivery of electrons to the respiratory chain, which results from either increased glucose or fatty acid oxidation or as a result of altered ETC stoichiometry. Consequently, an increased reverse electron flow occurs. Also, there is an evidence that increased cytosolic generation of ROS might precipitate increased mitochondrial ROS. The balance between the genesis of physiological mitochondrial ROS and antioxidant defenses may thus become disturbed, causing numerous pathological events, and finally leading to cell death.

A few words about an interesting association between mitochondrial ATP production and ATP deficiency in pathology….

Mitochondria are the primary source of ATP production in every cell. Therefore, disruption of mitochondrial respiratory function is regarded as a key event in the development of pathologic complications due to ATP depletion in different tissues, like for instance in heart tissue in diabetic patients. However, no general consensus has been raised about the occurrence of mitochondrial defects under diabetic conditions. Evidence associating diabetes with impaired mitochondrial respiratory function in the liver, heart and kidney of diabetic animals dates back more than 45 years. Despite this long history of research, we have still no comprehensive knowledge on the nature and extent of mitochondrial dysfunction in diabetics, as well as about the mechanisms linking this secondary metabolic abnormality with the primary metabolic defect in insulin and hyperglycaemia. The relationship between mitochondrial dysfunction and diabetic pathology has also not yet been defined and elucidated.

In summary…..

In this chapter we discuss how to possibly modulate the "vicious circle" established between mitochondria, oxidative stress and hyperglycaemia. The potential application of some existing and some new agents possessing promising anti-glycation properties to reduce glycation phenomenon and to increase the antioxidant defense system by targeting mitochondria is discussed. Moreover, this chapter outlines various mechanisms present in mitochondria that may lead to the development of diabetes. Intervention and therapy that alter or disrupt these mechanisms may serve to reduce the risk of development of this pathology.

2. Impact of hyperglycaemia on cellular biochemistry/metabolism – overview of the recent achievements in the field

Glycation and oxidative stress are two important processes known to play a key role in the etiopathology of complications in numerous disease processes. Oxidative stress, either via increasing reactive oxygen species (ROS), or by depleting the antioxidants, may modulate the genesis of glycated proteins *in vitro,* as well as *in vivo.*

2.1 Hyperglycaemia – the basic knowledge on Louis Maillard's discovery

Glycation (non-enzymatic N-glycosylation) is an endogenous process that contributes to the post-translational modification of proteins. It is slow under normal physiological conditions, giving rise to the presence of lysine- and arginine-derived glycation adducts in cellular and extracellular proteins. Inside cells, the impact of glycation is countered by high turnover and short half-life of numerous cellular proteins. In long-lived extracellular proteins, however, glycation adducts accumulate with age (Sell et al., 1996). Then, some of these adducts may be removed by enzymatic repair mechanisms, whilst all are removed by degradation of the glycated proteins. Degradation of extracellular glycated proteins requires specific recognition by receptors, internalisation and proteolytic processing. There are specific receptors, AGE receptors, which fulfill this role (Thornalley, 1998).

The Maillard reaction is named after Louis Maillard, who discovered over 80 years ago that some amines and reducing carbohydrates react to produce brown pigments (Ellis, 1959). The Maillard reaction proceeds via three major stages (early, advanced and final stage) and is dependent upon factors such as pH, time, temperature, as well as type and concentrations of reactants. Maillard reactions occur both *in vivo* and *in vitro*, and are associated with the chronic complications of diabetes, aging and age-related diseases (Edeas et al., 2010). The first step of this reaction typically involves the nucleophilic addition of a reducing sugar to a primary amine group (e.g. as found on a lysine or at the N-terminus of a protein). In this stage a reversible Schiff base is formed, which can undergo a slow irreversible rearrangement to form more stable Amadori product that accumulates over time (Fig. 1). The total amount of such accumulated products is known to be dependent on the type of sugar that is causing the glycation, the incubation time and sugar concentration, as well as the type protein that is being modified (Barnaby et al., 2011).

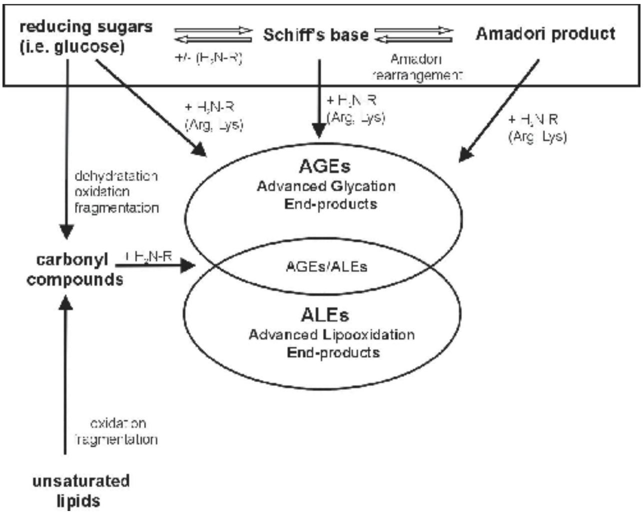

reducing sugars (i.e. glucose)
Schiff's base
Amadori product
+/- (H_2N-R)
Amadori rearrangement
+ H_2N-R (Arg, Lys)
+ H_2N-R (Arg, Lys)
+ H_2N-R (Arg, Lys)
dehydratation oxidation fragmentation
AGEs Advanced Glycation End-products
carbonyl compounds
+ H_2N-R
AGEs/ALEs
ALEs Advanced Lipooxidation End-products
oxidation fragmentation
unsaturated lipids

Fig. 1. General scheme of Maillard pathways in diabetic organism.

The first product of Maillard reaction is a simple glycosylamine, which readily undergoes the Amadori rearrangement to produce 1-amino-1-deoxy-2-ketoses. The large body of literature on these reactions is due to the multitude of possible reaction pathways and products, including fragmentations of the carbohydrates and formation of aromatic compounds from cyclisation/dehydration processes (Hodge, 1955). Reducing disaccharides also undergo this reaction, and it is a well-documented process for the degradation of lactose during the heating of milk. Reducing carbohydrates such as glucose, maltose, and lactose are tautomers and are in equilibrium with their more reactive aldehyde forms; nonreducing carbohydrates, such as mannitol, sucrose and trehalose, are not subject to Maillard reactions. Although early scientists believed that only primary aromatic amines were capable to become glycated, subsequent research has shown that nearly all primary and secondary amines, both aromatic and aliphatic, are capable of this reaction (Wirth et al., 1998).

Fragmentation of glucose adducts in early glycation processes establishes many parallel glycation pathways that lead to the subsequent formation of the so-called **A**dvanced **G**lycation **E**nd-products (AGEs). Analogous oxidation and dehydration reactions have been found in glycation by other hexose and pentose derivatives. Glyoxal, methylglyoxal and 3-deoxyglucosone (3-DG)-derived AGEs may be present in proteins glycated by glucose. Methylglyoxal-derived AGEs are common to proteins modified by glucose and by the authentic α-oxoaldehyde. Indeed, similar binding to AGE receptors has been found for these proteins. The formation of α-oxoaldehydes from monosaccharides, Schiff's bases and fructosamines suggests that AGEs may be formed at all stages of glycation (Westwood et al., 1999). AGEs alter structure and functions of proteins. It has been shown that the formation of AGEs *in vivo* contributes to several pathophysiological impairments associated with aging and diabetes mellitus, such as chronic renal insufficiency, Alzheimer's disease, nephropathy, neuropathy and cataract (Ravelojaona et al., 2007).

2.2 Impact of AGEs on human organism – undesirable effects on our health

Glycation of amino-groups on small or large cell constituents induces a number of undesirable effects in a plethora of age-related pathologies, overall referred to as glycation-induced health hazards and including cardiovascular diseases, kidney insufficiencies, retinopathy, or effects of AGEs on embryonic development, as observed in diabetes-related gravidities. As such reactions proceed with a speed proportional to the concentrations of the interacting substances, hyperglycaemia is an important factor for its acceleration (Urios et al., 2007).

It is very important to remember that Maillard products derive also from ''ready made'' ingested food. Table 1 shows the contents of Maillard products in some foods (expressed as N^{ε}-carboxylysine), as selected from data published by Goldberg et al. (2004).

Name of the selected food	**AGE content [U/g]***
Bread Whole wheat, crust, toasted	1.39
Corn flakes	2.32
Peanut Butter Chocolate	32
Popcorn, microwave	336
Butter	265
Fruits:	
Apple	127
Apple baked	445
Banana	87
Vegetables:	
Broccoli, carrots, celery	2.26
Carrots, canned	103
Pepper, mushrooms	2.66
Tomato, raw	234
Liquids:	
Milk, whole	48
Formula, infant	486
Human milk, fresh	52
Apple juice	20
Orange juice, carton	56
Beverages:	
Coffee, instatnt	53
Tea	19
Cola	65
Condiments:	
Ketchup	103
Mustard	29
Vinegar sauce, white	377
Cheese:	
Feta	84
Mozzarella	17
Parmesan	169
Hamburger, fast food	54

*AGE denotes N-carboxymethyllysine (CML)-like immunoreactivity, assessed by enzyme-linked immunosorbent assay using monoclonal antibody 4G9.

Table 1. The content of Maillard products in the selected victuals.

The Maillard reaction between reducing sugars and amino acids is a common reaction in foods, which undergo thermal processing. Desired consequences, like the formation of flavor and brown color of some cooked foods, but also the destruction of essential amino acids and the production of anti-nutritive compounds, require to consider the relevant mechanisms for controlling of Maillard reaction intermediates and final products. Processes such as roasting, baking or frying rely on favorable effects of the Maillard reaction, such as color and flavor formation, whereas during drying, pasteurisation and sterilisation the occurrence of the Maillard reaction is unfavorable. Nutritional losses of essential amino acids that are involved in the reaction, as well as the formation of reaction products are among those unwanted effects (Jaeger et al., 2010).

There is a limited number of studies that have been used to investigate the health effects of dietary Maillard neoformed compounds in humans. Some observational studies have been carried out to address the question of absorption, biodistribution and elimination of dietary Maillard Reaction Products (MRP), and to observe the associations between food exposure to MRPs and their *in vivo* levels.

Some reports have shown that in tobacco leaves, which are dried in the presence of sugars, the Maillard reaction cascade leads to a formation of glycated and oxidative derivatives. These compounds are inhaled during the smoking, after that they are absorbed by lungs and conjugated with serum proteins. It was evidenced that total serum AGE level in cigarette smokers is significantly higher in comparison with non-smokers. However, the highest level of AGEs was detected in the arteries and ocular lenses in diabetic smokers (Vlassara & Palace, 2002).

Furthermore, high AGE levels were observed in industrially preprocessed foods from animal products, like frankfurters, bacon, and powdered egg whites, compared with the unprocessed forms. Across all categories, exposure to higher temperature most of all raised the AGE content (for equal food weights). The temperature level appeared to be more critical than the duration. Also, microwaving increased AGE content more rapidly compared with conventional cooking methods (Peppa et al., 2002). Based on the above data, it is well evidenced that dietary glycoxidation products may constitute an important link between the increased consumption of animal fat and meat and the subsequent development of diabetic complications. However, the problem of AGEs' presence in food is well known, and therefore presently scientists call to use diets containing low contents of these compounds undesirable for out health.

Paradoxically, because of the metabolic demands of the brain, the human body has an obligatory requirement for glucose, approaching 200 g/day. The blood glucose concentration is tightly regulated by homeostatic regulatory systems and maintained between 40 mg/dl (2.2 mmol/l) and 180 mg/dl (10.0 mmol/l). Hypoglycaemia below the lower limit may result in coma, seizures, or even death. Hyperglycaemia, exceeding the upper limit, is associated with immediate glycosuria and caloric loss, as well as long-term consequences, like retinopathy, atherosclerosis, renal failure, etc. Under normal physiological conditions hyperglycaemia stimulates insulin secretion, promoting uptake of glucose by muscles and adipose tissue (Chiu & Taylor, 2011).

Nevertheless, several studies suggest that some MRPs present in foods could have beneficial effects on human health. For instance, the melanoidins are brown Maillard polymers, which seem to have functional properties in food products and are also capable of inhibiting growth of a tumour cell line in culture (Marko et al., 2003). In addition, it was also found recently that a selection of foods rich in MRPs could inhibit the oxidation of LDL *in vitro*.

The high diversity of the MRPs formed in the very diverse food matrices makes it impossible to classify all of them as glycotoxins. It is admitted that they have different beneficial or detrimental biological activities. Thus, more well-controlled clinical experiments are needed to establish the role of the ingested MRPs, pure or added to food matrices, following acute or chronic exposures (Tessier & Birlouez-Aragon, 2010).

2.3 The role of Reactive Oxygen Species (ROS) in glycation process

Free radicals in biological materials were discovered less than 60 years ago. Soon thereafter, Denham Harman hypothesized that reactive oxygen radicals may be formed as by-products of enzymatic reactions *in vivo.* In 1956 he described free radicals as a Pandora's box of evils that may account for gross cellular damage, mutagenesis, cancer, and, last but not the least, the degenerative process of biological aging (Harman, 1956). Presently, the list of cell and tissue disorders caused by free radicals is very long and the diseases, such as diabetes and/or impairments in mitochondria functions, also belong to the "victims" of ROS attack.

Glycation and oxidative stress are closely linked, and both phenomena coincide in a vicious process referred to as ''glycoxidation''. In all steps of glycoxidation there is a massive generation of oxygen-free radicals, some of them being common with lipidic peroxidation pathways. Besides, glycated proteins, and especially their advanced adducts, activate membrane receptors, such as RAGE, and induce an intracellular oxidative stress and a pro-inflammatory status. Glycated proteins may modulate functions of cells involved in oxidative metabolism and induce inappropriate responses. Finally, some oxidative products (reactive aldehydes such as methylglyoxal) or lipid peroxidation products (malondialdehyde) may bind to proteins and amplify glycoxidation generated lesions (Hunt et al., 1998).

Recently, oxygen free radicals, antioxidant defences and the cellular redox status have been considered as central players in pathogenesis of diabetes. The role of glycaemic control on the pro-oxidant/antioxidant balance deserves special attention. Metabolic disturbances and oxidative stress seem to be closely related, improved glycaemic control being associated with a lowered pro-oxidant status (Wierusz-Wysocka et al., 1995).

It was also evidenced that there is a relationship between oxidative stress and insulin resistance observed in diabetes. Hyperinsulinaemia increases the concentrations of ROS, which, in turn, may be responsible for the impaired intracellular insulin actions. Amongst ROS, hydrogen peroxide has been shown to contribute to insulin receptor signaling, and may play a key role in the modulation of the signalling transduction pathways regulated by insulin through coupled receptors.

Consequently, the inactivation of hydrogen peroxide by catalase could represent a critical step for the removal of intracellular ROS in insulin-producing cells. On the other side, the inhibition of catalase under conditions of insulin resistance could also represent an adaptive response to maintain the homeostasis of intracellular hydrogen peroxide as an intermediate of the insulin-activated physiological processes. Overall, relationships between ROS and diabetes seem extremely complex (Bonnefont-Rousselot, 2002).

A second source of ROS formation is an excessive production of AGEs, especially due to a hyperglycaemia-induced overproduction of methylglyoxal. AGEs are also able to produce oxygenated free radicals via complex biochemical mechanisms. AGEs have been shown to interact with their specific receptors (RAGE) and thus they induce oxidative stress, enhance vascular cell adhesion molecule type 1 (VCAM-1) expression, and increase endothelial adhesiveness for monocytes. This overproduction of AGEs appears to play a key role in the pathogenesis of diabetic complications. In particular, the accumulation of two AGEs biomarkers, namely carboxymethyllysine and pentosidine, has been related to the severity of diabetic nephropathy and the so-called 'carbonyl stress'. The toxic effects of AGEs result from structural and functional alterations in proteins, especially the cross-linking of proteins, and from their interactions with RAGEs leading to the enhanced formation of oxygen free radicals (Miyata et al., 2001; Singh et al., 2001).

2.4 Diabetes – a frequent disease or an epidemic?

In 1993 the World Health Organisation (WHO) Ad Hoc Diabetes Reporting Group published standardized global estimates for the prevalence of diabetes and impaired glucose tolerance in adults, based on data from 75 communities in 32 countries. These estimates provided, for the first time, comparable information on the prevalence of abnormal glucose tolerance from many populations worldwide. However, they did not meet the needs of those who frequently refer to the WHO diabetes program for information on the number of people with diabetes in a particular country/community, nor did they take account of future trends in the burden of diabetes (King & Rewers, 1993). Therefore, a further study has been undertaken that links data from the global database collected by WHO with demographic estimates and projections issued by the United Nations to estimate the number of people with diabetes in all countries of the world at three points in time, i.e., the years 1995, 2000, and 2025. The results of this study suggest that for the world as a whole, between the years 1995 and 2025, the adult population will increase by 72%, prevalence of diabetes in adults will increase by 35%, and the number of people with diabetes will increase by 122% (Fig. 2). For the developed countries, there will be an 11% increase in the adult population, a 27% increase in the prevalence of adult diabetes, and a 42% increase in the number of people with diabetes. For the developing countries, there will be an 82% increase in the adult population, a 48% increase in the prevalence of adult diabetes, and a 170% increase in the number of people with diabetes (King et al., 1998).

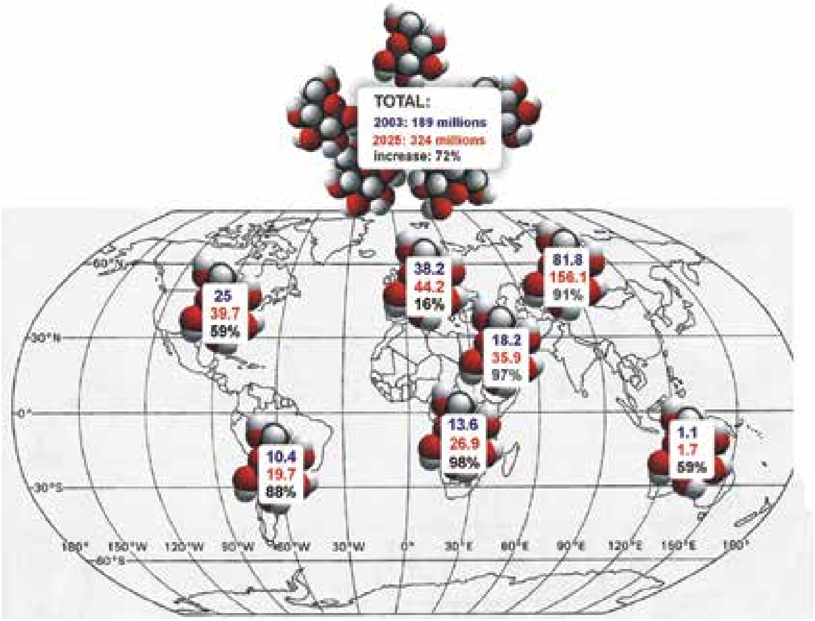


Fig. 2. Global projections for the diabetes epidemic in years 2003-2025.

It is well known that diabetes is one of the most costly and burdensome chronic diseases of our time and is a condition that is increasing in epidemic proportions throughout the world. The prevalence of abnormal glucose tolerance in any population is of public health concern, since diabetes may increase disability burden and health care utilisation.

The relationship between blood glucose concentration in diabetes and the incidence of disease complications was demonstrated in large epidemiological studies. Accurate metabolic control in diabetes is not always feasible, and therefore the issue of molecular mechanisms underlying the damaging effects of hyperglycaemia on body cells and tissues, as well as the possibilities of their pharmacological inhibition, are of the utmost importance in a diabetological practice and anti-diabetic treatment.

The complications resulting from the disease are a significant cause of morbidity and mortality and are associated with the damage or failure of various organs such as eyes, kidneys, and nerves. Although the treatment of diabetes has become increasingly sophisticated, with over a dozen pharmacological agents available to lower blood glucose, a multitude of ancillary supplies and equipment available, and a clear recognition by health care professionals and patients that diabetes is a serious disease, the normalisation of blood glucose for any appreciable period of time is seldom achieved. In addition, in well-controlled so called "intensively" treated patients, serious complications still occur, and the economic and personal burden of diabetes remains (Turner et al., 1999).

Nowadays, diabetes is treated not only as a disease, but as an epidemic. However, as a discipline, diabetes epidemiology is relatively young. The first significant gathering of researches interested in diabetes epidemiology took place just in 1978. Then, in the relatively short span of 2 decades, epidemiology studies have had a profound impact on diabetes research, care and prevention. This explosion of interest and activity in the epidemiology of diabetes should contribute to an effective reduction in the number of patients with this disease.

3. Mitochondria – the relationship between the structure and the function

Mitochondria are multifunction organelles, which play a key role in both the proper functioning of the cell and normal cell death scenario (Kuznetsov & Margreiter, 2009, as cited in McBride et al., 2006). Their main role is the production of adenosine triphosphate (ATP) through metabolic processes involving tricarboxylic acid cycle (TCA) and the electron transport chain (ETC). Most cellular ATP is generated in the process of oxidative phosphorylation, which is possible thanks to the 'sophisticated machinery' located in the inner mitochondrial membrane. Mitochondria participate in the regulation of redox state and calcium homeostasis in cell. Cations of calcium regulate some mitochondrial processes, such as enzyme activity, i.e. pyruvate dehydrogenase, or metabolic rate. These organelles participate in biosynthesis of amino acids, vitamin cofactors, fatty acids and neurotransmitters (Waldbaum & Patel, 2009). Many other biochemical reactions are associated with the functioning of these structures, including synthesis of heme group and some steps of steroid synthesis. Also, a part of the processes occurring in the urea cycle take place there (Pinti et al. 2010). Mitochondria have critical function in the control of apoptotic and necrotic cell death and in most types of cells they are also a major site of reactive oxygen species (ROS) generation (Duchen, 2004). ROS are involved in many signaling pathways. Most of them are second messengers that trigger different cellular events, such as cytokine secretion or activation of transcription factors, but in excess they can contribute to the formation of defects in mitochondria, as well as in a whole cell (Edeas et al., 2010a).

3.1 Mitochondrial structure and biogenesis

Mitochondria are encapsulated by two membranes, each with different structure and function, separated by intermembrane space, in which some important proteins involved in the mitochondrial bioenergetics and/or cell death are located (Fig. 3) (Duchen, 2004, Borutaite, 2010). The outer membrane contains porins, which make it permeable to molecules smaller than 5-6 kDa (Waldbaum & Patel, 2009). Compounds such as water, O_2, CO_2, and NH_3 easily pass through the membrane, but hydrophilic metabolites and all inorganic ions in order to get over this membrane require the participation of specific channels and carrier proteins. Such a transport is generally based on the exchange of molecules, i.e. ADP is exchanged for ATP and P_i (inorganic phosphate) for OH^- (Szewczyk & Wojtczak, 2002). The mitochondrial inner membrane contains enzymes facilitating an oxidative phosphorylation (OXPHOS). This complex of enzymes consists of four oxidoreductases involved in respiratory electron transport (Complexes I - IV) and the ATP synthase complex (Complex V).

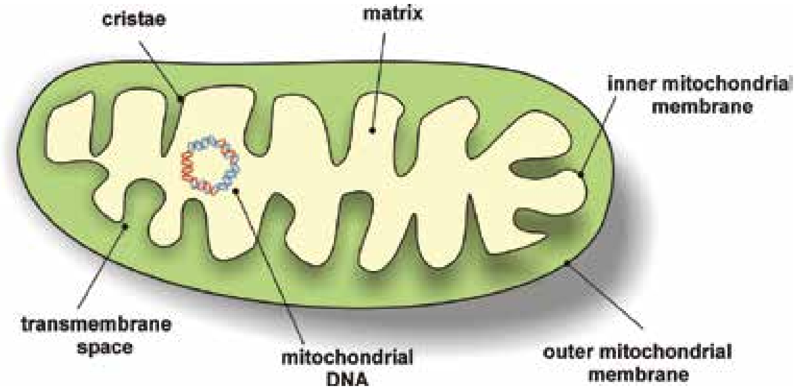


Fig. 3. Mitochondria structure and components.

Until recently, the inner membrane has been described as a multiple infolded structure forming cristae and containing numerous mitochondrial proteins (Duchen, 2004). However, electron tomographic analyses of a variety of mitochondria (both isolated and observed *in situ* in various cell types) have provided overwhelming evidence showing the need of some changes in the perception of the structure of these organelles. These infoldings or rather invaginations are not randomly spaced in the membrane, as often considered, but resemble microcompartments, which face each other in the peripheral region of the membrane. The narrow junctions are wide enough to pass metabolites and many soluble proteins (Mannella, 2008). However, the number of cristae junctions and the morphology of the intercristal space depends on the metabolic state of mitochondria (Logan, 2006). Isolated mitochondria usually occur in one of two morphologic states, condensed or orthodox. The first one is characterized by contracted, very dense matrix and wide cristae. In the second state matrix is expanded and cristae compartments are more compact. Osmotic and metabolic changes in mitochondria are responsible for these alterations. It is believed that mitochondrial inner membrane topology is regulated by the cell to improve mitochondria capacity in their response to stimuli (Mannella, 2008).

Tissue cells contain from a few dozen to several thousands of mitochondria and their number is associated with cell energy demands. Organs such as heart, muscles or brain contain the largest number of mitochondria. Mitochondria are very dynamic structures, which can divide, undergo fusion and can take the form of the network of elongated and interconnected filaments. The phenomena of fission and fusion have an impact on mitochondrial shape, size and number (Logan, 2006). The division and replication of mitochondria is under control of the nucleus and is somehow associated with division and replication of nuclear DNA. Replication of mitochondria requires coordination between the process of mtDNA replication and synthesis of proteins encoded in both genomes (nucleus and mitochondrial). The production of both types of proteins must be synchronized to preserve their functionality.

3.2 Electron transport chain and ATP synthesis

Mitochondrial ATP production involves three main steps: a) the enzymatic “combustion” of acetyl in tricarbolxylic acid cycle (TCA), b) the electron transport chain activity and c) ATP synthase action. Energy released during this cycle is used to reduce the electron carriers NAD^+ to NADH and FAD^{2+} to FADH (Duchen, 2004). Electrons from NADH and $FADH_2$ are transferred to the respiratory chain - a coupled enzyme systems composed of four complexes (Complex I - IV). Complex I (NADH dehydrogenase) is the major entrance point of electrons to respiratory chain and is composed of two domains. One domain, localized in withe membrane, is involved in proton translocation across the bilayer, and the other, matrix-exposed domain, is responsible for oxidation of NADH. $FADH_2$ is the donor of electrons to succinate dehydrogenase (Complex II) which is the second entrance point of electrons to the ETC. Electrons from both complexes are transferred on mobile intermediate – ubiquinone, which is converted to reduced form - ubiquinol. The flow of electrons from ubiquinol is directed through the Complex III, also known as ubiquinol-cytochrome c reductase, to another carrier - cytochrome c, which transfers electrons to Complex IV - cytochrome c oxidase. Finally, at the very end of the respiratory chain, Complex IV reduces the oxygen to water in sequential four-electron transfer (Adam-Vizi & Chinopoulos, 2006). The oxidation of NADH and $FADH_2$ provides the energy to transport protons from mitochondrial matrix into the intermembrane space by the proton pumps (Complexes I, III, IV). The difference in the proton concentration, and thus the difference in the electric charge across the inner mitochondrial membrane creates the electrochemical potential gradient, also called an electrochemical proton gradient or a ′proton-motive force′, which is mainly expressed as a mitochondrial transmembrane potential (Nazaret, 2008). The structure of mitochondrial electron transport and the scheme showing ATP production by mitochondria was introduced in Fig. 4.

Energy needed to phosphorylate ADP by ATP synthase comes from the entry of protons back into the matrix through the proton channel of this complex. This process is called oxidative phosphorylation (Frey & Mannella, 2000). ATP is then transported to the cytoplasm by the adenine nucleotide translocase (ANT). However, there are several mechanisms that may lead to the loss of mitochondrial potential, including an inhibition of respiration, failure in substrate supply and uncoupling mechanisms that cause proton leak across the membrane.

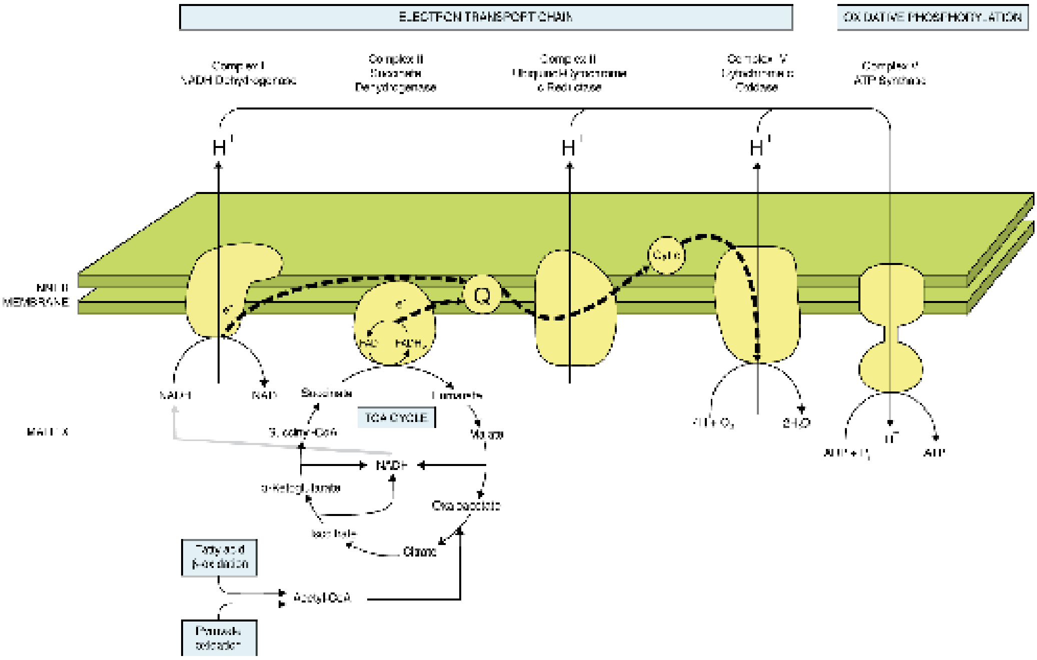

ELECTRON TRANSPORT CHAIN
OXIDATIVE PHOSPHORYLATION
Complex I NADH Dehydrogenase
Complex II Succinate Dehydrogenase
Complex III Ubiquinol Cytochrome c Reductase
Complex IV Cytochrome c Oxidase
Complex V ATP Synthase
H+
H+
H+
INNER MEMBRANE
e-
e-
Q
Cyt c
FAD
FADH2
MATRIX
NADH
NAD
Succinate
Fumarate
TCA CYCLE
Succinyl-CoA
Malate
NADH
α-Ketoglutarate
Oxaloacetate
Isocitrate
Citrate
Fatty acid β-oxidation
Acetyl CoA
Pyruvate oxidation
4H + O2
2H2O
ADP + Pi
H+
ATP

Fig. 4. The general mechanism leading to oxidative phosphorylation is as follows: high-energy electrons (marked as e^-) derived from NADH and $FADH_2$, are moving along the respiratory chain composed of four protein complexes (Complex I - IV) and two additional electrons carriers: ubiquinone (coenzyme Q, **Q**), a small molecule freely moving in the inner mitochondrial membrane layer, and cytochrome c (**Cyt c**), localized in the intermembrane space attached to the inner membrane. Part of the energy released in this process is used up in the action of proton pumps transporting protons (H^+) from matrix to the intermembrane space. Across the inner membrane electrochemical gradient of protons is formed. Protons tend to return to the mitochondrial matrix and restore alignment of H^+ concentration on both sides of the membrane. When they pass back through transmembrane protein complex – ATP synthase - the energy of their movement is used for the synthesis of ATP from ADP and inorganic phosphate (**P_i**).

3.3 Free radical generation by mitochondria

Oxidative metabolism and ATP synthesis are closely associated with ROS generation in mitochondria. These organelles consume 80–90% of cell's oxygen during oxidative phosphorylation. The electron transport chain is the main source of ROS in functioning mitochondria. Approximately 0.2–2% of the oxygen taken up by a cell is converted by mitochondria to ROS. Superoxide ($O2^{\bullet-}$) is the main product of these transformations, and it is then converted to hydrogen peroxide (H_2O_2) by spontaneous dismutation or by superoxide dismutase (SOD). Glutathione peroxidase or catalase, in turn, convert hydrogen peroxide into water. If this change does not occur, in the presence of divalent cations H_2O_2 can undergo Fenton's reaction to produce even more harmful hydroxyl radical (•OH). Oxygen can be reduced to superoxide in one-electron step, theoretically, at each step of the respiratory chain, but in reality two major sites of superoxide generation are Complex I and Complex III (Paradies et al., 2010, as cited in Murphy, 2009). There is a considerable experimental support for two mechanisms of ROS production by complex I. The first one is the production of ROS as a consequence of so-called reverse electron transfer (RET) in the mitochondrial respiratory chain. RET is a set of redox reactions in the mitochondrial ETC that allows electrons to flow from coenzyme Q to NAD^+ instead to oxygen. The other one takes place under normal conditions, whereas most of the energy from the creation of mitochondrial potential difference is used to generate ATP through ATP synthase. This process causes collapse of the proton gradient. The amplitude of the electrochemical proton gradient regulates the flow of electrons through the ETC. When the electrochemical potential gradient is high, for instance under conditions of high glucose concentrations, the life of electron transport intermediates that are involved in superoxide formation, such as ubisemiquinone, is prolonged. The reason of such condition is that the activities of of ETC proton pumps depend on the proton gradient across the inner membrane and the membrane itself – two components of proton-motive force (Duchen, 2004).

3.4 Free radical targets and the oxidative vicious circle

Mitochondria are continuously exposed to action of reactive oxygen species so they need to have a system that will prevent them against destructive effect of oxidative damage. In fact, mitochondria are equipped in complicated multi-leveled ROS defense network consisting of enzymes and non-enzymatic antioxidants. They contain a high concentration of glutathione, α-tocopherol and manganese-containing superoxide dismutase (MnSOD). The role of MnSOD is the dismutation process of superoxide radical to H_2O_2. The product of MnSOD reaction is detoxified by other enzymes, i.e. catalase, which converts H_2O_2 into O_2 and H_2O. Mitochondria possess also another system capable of efficient superoxide removal - the cytochrome c, which is than regenerated (oxidized) by its natural electron acceptor, cytochrome c oxidase (Complex IV). In intact mitochondria, superoxide may be efficiently scavenged by intramitochondrial antioxidant defences (Duchen, 2004). An imbalance between oxidants and antioxidants induces oxidative stress responsible for alteration of biomolecules and intracellular signaling pathways present in every cell (Edeas et al., 2010a). Mitochondria are a major source of ROS generation, but what is important, they are also its major target (Duchen, 2004). Mitochondrial membrane lipids, mainly long-chain polyunsaturated fatty acids (PUFAs), are also susceptible to oxidative stress. PUFAs are basal components of mitochondrial phospholipids. The sensitivity of PUFAs to oxidation increase with the increasing number of double bonds per fatty acid molecule. Peroxidation of membrane phospholipids causes alterations in their structure and consequently may disrupt organisation of the lipid bilayer. It contributes also to changes in membrane fluidity and/or permeability, and causes changes in the mitochondrial membrane potential, in respiratory capacity and in oxidative phosphorylation. ROS are also responsible for alterations in proteins, which may manifest by changes in their structure, proteolytic susceptibility and spontaneous fragmentation. Oxidative damage especially affects the mitochondrial electron transport chain and, when the ETC enzymes stop working properly, the ROS production increases. This may result in the incomplete oxygen consumption, reduced production of ATP, and finally overproduction of ROS (Waldbaum & Patel, 2009).

4. Mitochondrial physiology in diabetes

Mitochondria are provided with a variety of bioenergetic functions mandatory for the regulation of intracellular energy production. Alteration of bioenergetic activities may have drastic consequences on cellular function through the perturbation of energetic charge and balance of the cell (Fig. 5).

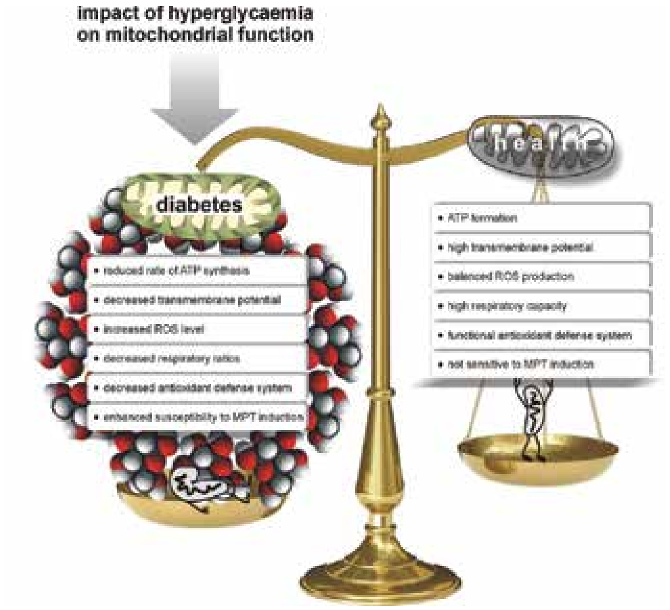

impact of hyperglycaemia on mitochondrial function
diabetes
• reduced rate of ATP synthesis
• decreased transmembrane potential
• increased ROS level
• decreased respiratory ratios
• decreased antioxidant defense system
• enhanced susceptibility to MPT induction
health
• ATP formation
• high transmembrane potential
• balanced ROS production
• high respiratory capacity
• functional antioxidant defense system
• not sensitive to MPT induction

Fig. 5. The overall impact of the burden of hypergycaemia in diabetes on functioning of mitochondria.

Abnormalities of mitochondrial metabolism causing human disease have been recognised for more than 40 years. Numerous reports clearly indicate the association between mitochondrial dysfunction and diabetes. Nevertheless, some mechanisms of mitochondrial role in this pathology still requires further elucidation. Therefore different animal model studies are involved in the investigation of explaining these unknown mechanisms. There are several models of experimental diabetes that mimic two common types of diabetes. Streptozotocin-induced diabetes is a widely accepted animal model for type 1 diabetes, resulting from the inability of the pancreatic beta cells to produce insulin. For the research on type 2 or the insulin resistant state, resulting from the inefficient use of insulin by the tissues to regulate blood glucose concentration, some genetically manipulated animal models (e.g. Zucker fatty rats (ZFR), *ob/ob* (obese) mice, CP (corpulent) rats, GK (Goto-Kakizaki) rats, Akita mice) may be utilized (Srinivasan & Ramarao, 2007).

4.1 Mitochondrial dysfunction and diabetes type 1

Streptozotocin (STZ) is a naturally occurring chemical that is particularly toxic to the insulin-producing β cells of the pancreas in mammals and is used to generate Type 1 diabetes in the experimental model. Animals with diabetes induced by STZ exhibit increased mitochondrial oxidative stress and dysfunction. Other agent, alloxan, a toxic glucose analogue, is also used in order to generate type 1 of diabetes. Alloxan selectively destroys insulin-producing cells in the pancreas when administered to rodents and many other animal species, and has been shown to cause also mitochondrial dysfunction. Alloxan-treated severe diabetic rats were shown to exhibit impaired mitochondrial phosphorylative activities and low mitochondrial oxidation–reduction states (Yamamoto et al., 1981). In one month old alloxan-diabetic animals the enzyme activity of the mitochondrial membrane marker, F_0F_1-ATPase, was found to be decreased. Insulin treatment caused hyperstimulation of the activity, whereas in late-stage diabetes the catalytic efficiency of the enzyme was increased and became decreased upon insulin treatment (Patel & Katyare, 2006).

Mitochondrial dysfunction in diabetic rats can be succinctly summarized into: decreased mitochondrial 3'-AMP forming enzyme activity, increased oxidative and nitrosative stress, decreased oxygen consumption, loss in mitochondrial transcriptional capacity, increased HMG-CoA synthase, increased levels of pyruvate and dicarboxylate transporters, increased degradation of ATPase, changes in phospholipid composition, increased pyruvate carboxylase activity, increased fatty acid beta oxidation and ultrastructure alterations.

4.2 Mitochondrial dysfunction and diabetes type 2

Type 2 diabetes is the most common metabolic disease in the world, and its prevalence much exceeds the prevalence of type 1 diabetes. Among different causes leading to diabetes the role of mitochondria is considered substantial. Disorders of the mitochondrial electron transport chain, overproduction of ROS and lipoperoxides or impairments in antioxidant defenses are encountered in type 2 diabetes. Increased ROS levels lead to generalized oxidative damage to all mitochondrial components. Moreover, it is well established that mitochondrial function is required for normal glucose-stimulated insulin secretion from pancreatic β cells. However, the studies in humans suggest that more subtle defects in mitochondrial function may also play a role in the pathogenesis of insulin resistance and type 2 diabetes (Fig. 6) (Luft, 1994).

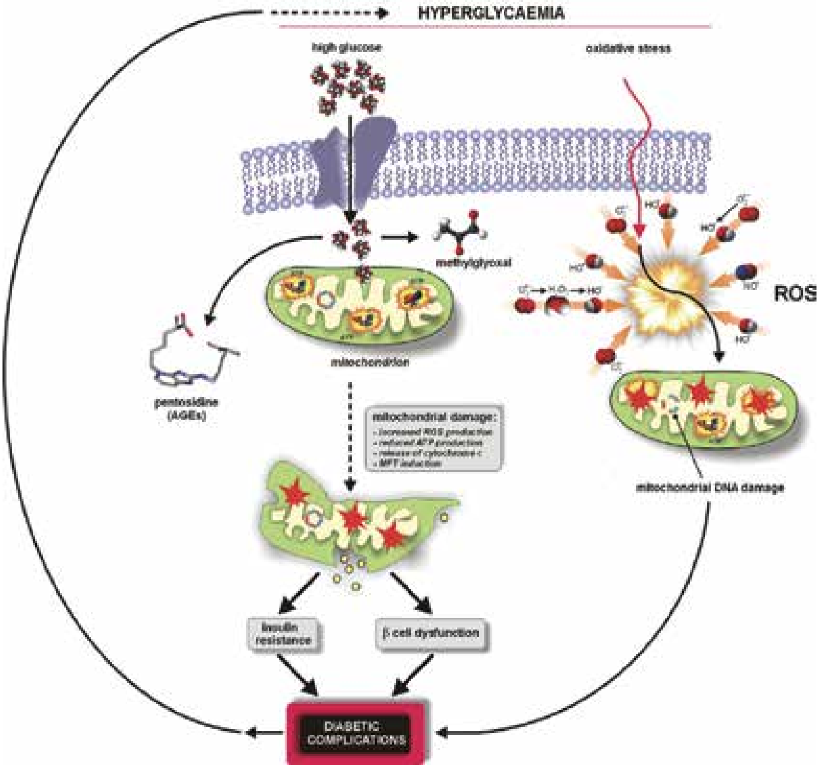

HYPERGLYCAEMIA
high glucose
oxidative stress
methylglyoxal
mitochondrion
pentosidine
(AGEs)
ROS
mitochondrial damage:
- increased ROS production
- reduced ATP production
- release of cytochrome c
- MPT induction
mitochondrial DNA damage
insulin
resistance
β cell dysfunction
DIABETIC
COMPLICATIONS

Fig. 6. Relationships among hyperglycaemia, mitochondrial damage, oxidative burst and diabetic complications.

Some data support the hypothesis that insulin resistance in humans arises from defects in mitochondrial fatty acid oxidation, which in turn leads to increased intracellular fatty acid metabolites that disrupt insulin signaling (Petersen et al., 2003). Alternatively, the reduction in mitochondrial oxidative phosphorylation activity in insulin-resistant individuals could be due not to mitochondrial loss, but rather to a defect in mitochondrial function. This hypothesis is supported by muscle biopsy studies. In one such study, the activity of mitochondrial oxidative enzymes was found to be lower in type 2 diabetic subjects, and in another, the activity of mitochondrial rotenone-sensitive nicotinamide adenine dinucleotide oxidoreductase [NADH:O(2)] was found to be lower (Lowell & Shulman, 2005).

4.3 Alterations in cardiac mitochondria observed in diabetes

Cardiovascular diseases are the predominant cause of death in patients with diabetes mellitus. Underlying mechanism for the susceptibility of diabetic patients to cardiovascular diseases still remains unclear. Elevated oxidative stress was detected in diabetic patients and in animal models of diabetes. Hyperglycaemia, oxidatively modified atherogenic lipoproteins, and advanced glycation end products act in a concerted action together with oxidative stress, and cumulatively contribute to progression of late diabetic complications. Mitochondrial dysfunction increases electron leak and the generation of ROS from the mitochondrial respiratory chain (MRC). High levels of glucose and lipids impair the activities of MRC complex enzymes. Furthermore, increased activity of NADPH oxidase (NOX), which generates superoxide from NADPH in cells, was detected in diabetic patients (Shen, 2010).

Because mitochondria constitute 20–30% of the cardiac myocytes, one of the potent causes for heart malfunctioning in diabetes is the impaired mitochondrial function and consequently the decreased ATP generation (Rolo & Palmeira 2006).

Many reports evidenced that diabetic hearts show impaired mitochondrial function, decreased ATP generation, decreased oxidative capacity, increased ROS, abnormal morphology, increased UCP-3 level, decreased mitochondrial calcium uptake and increased susceptibility to MPT induction (Fig. 5).

Distortions in cardiac mitochondrial bioenergetics are known to occur in both human types of diabetes and in models of diabetes in animals. Reduced mitochondrial calcium uptake was observed in heart mitochondria from STZ-treated rats. This was related to enhanced susceptibility to MPT induction rather than damage to the calcium uptake machinery. Interestingly, heart mitochondria from GK rats were less susceptible to the induction of MPT, showing larger calcium accumulation before the overall loss of mitochondrial impermeability. Different approaches of antioxidant administration in GK rats (vitamin E or coenzyme Q_{10}) showed no success in reversing the diabetic phenotype (Oliveira et al., 2003).

Diabetic heart failure may be causally associated with alterations in cardiac energy metabolism. Fuel selection and capacity for ATP production in the normal and failing heart are dictated by several metabolic regulatory events at the level of gene expression. Decline in the capacity for ATP, as caused by progressive impairment of mitochondrial function, is a gradual step in the progression to heart failure of any cause. Fetal heart depends on glucose and the adult heart on glucose and fatty acids. The switch between fatty acid oxidation and glucose in the adult heart leads to a healthy metabolic situation (Huss & Kelly, 2005). In the insulin-resistant and diabetic heart, fatty acid oxidation is increased and glucose utilisation is diminished. Long-term consequence of fatty acid oxidation is mitochondrial dysfunction. A number of mechanisms may be responsible for enhanced fatty acid utilisation in type 2 diabetic hearts, such as increased fatty acid uptake into the cell and mitochondria, increased UCP-3 expression, and stimulation of peroxisome proliferator-activated receptor-α (PPARα) (Fig. 7) (Rolo & Palmeira, 2006).

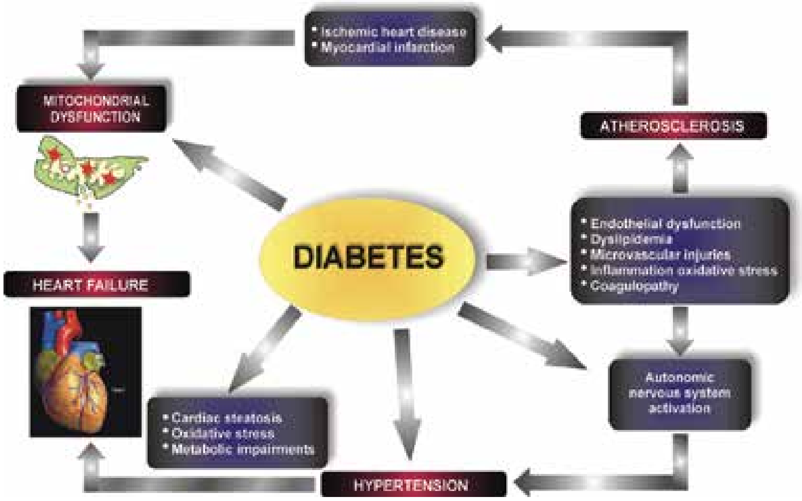

• Ischemic heart disease
• Myocardial infarction
MITOCHONDRIAL DYSFUNCTION
ATHEROSCLEROSIS
DIABETES
• Endothelial dysfunction
• Dyslipidemia
• Microvascular injuries
• Inflammation oxidative stress
• Coagulopathy
HEART FAILURE
Autonomic nervous system activation
• Cardiac steatosis
• Oxidative stress
• Metabolic impairments
HYPERTENSION

Fig. 7. The role of cardiac mitochondria in the development of heart failure in a course of diabetes.

Diabetes-associated metabolic disorders may cause the mitochondrial dysfuntion and upregulation of NOX in the cardiovascular system, which lead to increased ROS production and oxidative stress in vasculature and blood circulation. ROS may directly oxidize or indirectly regulate molecules related to atherosclerosis and thrombosis. Mitochondrial NOX, or its regulators may be considered as potential drug targets for the prevention and/or treatment of diabetic cardiovascular complications.

5. Therapeutic approaches to reduce diabetic complications

Patients with diabetes mellitus are usually treated with a combination of pharmacological agents and habitual approach, i.e. their lifestyle modification. The development of new antidiabetic agents, such as insulin analogs and incretin-based therapies, has led to treatment strategies that enable numerous patients with diabetes to improve their lifestyles.

5.1 Use of insulin in a common diabetes therapy – its advantages and disadvantages

Type 1 diabetes mellitus is underlied by the shortage of insulin, which plays a crucial role of carbohydrate and fat metabolism. Its absence causes rather a complex array of serious impairments in patients' health, e.g. hiperglycaemia, ketoacidosis, coma and even death. Hence, the injection of exogenous insulin to diabetic individual is essential to: (a) maintain a normal glucose concentration and (b) avoid the advanced microvascular complications, such as retinopathy, nephropathy or neuropathy.

Insulin was discovered by Banting and Best and this event was a milestone in the treatment of patients with diabetes (Bliss, 1982). Initially, bovine, porcin and even some fish analogues were applied to avoid diabetic ketoacidosis. However, gradually the scientists were challenged by the uprising problems in patients injected with animal insulins (mainly because of their rather high impurity and immunogenicity) and were forced to develop a new class of insulins. Improved techniques used for insulin purification combined with other compounds like protamine and zinc, enabled to manufacture protamine insulin with the prolonged time of activity and later, the more stable protamine zinc insulin (Hagedorn et al., 1936). Moreover, further scientific discoveries shed light on better understanding of insulin structure and activity and initiated a new avenue to design human insulin analogues characterized by the properties of prolonged hormone activity in a bloodstream.

At present, there are few types of short-acting insulin analogues used in anti-diabetes therapy, and among them:

- **insulin lispro** (Humalog manufactured by Elli Lilly and Company), which was approved and launched into the market in 1996. Its modified amino acid sequence provides faster absorption, which is essential to ameliorate postprandial glucose level. The studies revealed that the activity peak of lispro appears in 1 hour after using and lasts for next 3-4 hours (Howey et al., 1994)
- **insulin glulisine** (Apidra manufactured by Sanofi-Aventis), which possesses asparagine at position B3 and glycine at position B29 in amino acid chain (Garg et al., 2005).
- **insulin aspart** (NovoRapid manufactured by Novo Nordisk), in which proline is replaced with aspartic acid what facilitates its faster absorption (Mudaliar et al., 1999).

Long-acting insulin analogues are crucial to mimic the endogenous insulin secretion. Thereby a specific modification of insulin structure was essential to obtain longer acting analogues. There are two approaches leading to diminish absorption: the first one is to change the isoelectric point of insulin and the second one is to acetylate a hydrophobic residue with fatty acid.

- **insulin glargine** (Lantus, Sanofi-Aventis) exemplifies a long-acting insulin analogue, in which asparagine is substituted by glycine at position A21 and the position B30 is enriched with two molecules of arginine at B31 and B32 (Bolli & Owens, 2000). These alterations have an impact on the structure and an isoelectric point of insulin, contributing to a decrease in its solubility after injection. Finally, the result of these changes is the product, which acts about 20 hours (Heise et al., 2002).
- **insulin detemir** (Levemir, Novo Nordisk) is characterized by long acting properties (17-20 hours) obtained as a result of acetylation with fatty acid at the position B29 and by removal of threonine at B30 (Havelund et al., 2004).

After long years of experience and observations, nowadays, multiple daily injection program is believed to be the most reasonable approach in modern diabetic treatment in order to mimic the physiological insulin release. However, it implies that patients have to undertake more inconvenient therapy resulting from the scheduled injections of both short- and long-acting insulins. To deal with this problem, clinicians and patients may choose an alternative method, which requires biphasic insulin analogues administration.

Patients with pre-diagnosed type 2 diabetes should take seriously into account the radical change of their lifestyle in order to cause a delay of possible medical intervention. Under conditions when diet or changing a lifestyle may not be sufficient enough, additional pharmacological treatment is required to avoid a severe consequence of this disease.

Nowadays, medicine is focused on delivery a large number of drugs. From pharmacological point of view, these compounds should be effective in improving insulin efficiency or effective in enhancing its secretion from pancreas. Among those substances are: sulfonylureas, biguanides, thiazolidinediones, meglitinides, α-glucosidase inhibitors, amylin analogues, incretin hormone mimetics and dipeptidyl peptidase 4 inhibitors.

- **sulfonylureas** belong to the drugs most frequently used in diabetes treatment. It is known that these oral hypoglycaemic agents interact with β-cell pancreas cells causing insulin secretion, insulin sensitivity amelioration, as well as glucose synthesis reduction. However, in order to treat patients using sulfonylureas, an endogenous secretion of insulin at the same/similar level and a balanced diet are required (Gerich, 1989)
- **metformin**, representing a class of biguanides, is a commonly used oral hypoglycemic agent for the treatment of type 2 diabetes. Metformin is also known as an inhibitor of high glucose- or AGEs-induced ROS generation (Bellin et al., 2006)
- **thiazolidinediones (glitazones, TZD)**, oral anti-diabetic drugs, are based on the improving of adipose tissue and muscle sensitivity to insulin treatment (Day, 1999). However, troglitazone, one of the class of thiazolidinediones, was taken off the market since the hepatotocity has been noticed (Watkins & Whitcomb, 1998).

Development of the obesity associated with diabetes requires using a novel combination treatment, which aims at retarding the microvascular and macrovascular complications occurring in diabetes and obesity. Therefore, some anti-diabetic agents have been indicated to maintain an adequate glucose level in an organism suffering from diabetes. The number of anti-diabetic drugs delivered by subcutaneous injection increases constantly and to date there are numerous agents already launched into market, as well as others, tested in the clinical trials (Fig. 8).

- **glucagon-like-peptide-1 agonist (GLP-1)**, one of the first among subcutaneous drugs used in medicine. Its beneficial effect was achieved by suppression of glucagon secretion and weight loss. Unfortunately, it soon appeared that in the organism GLP-1 was active only 2 min. It was the main reason why scientists started to work on improved GLP-1 analogues.
- **exenatide and liraglutide** were introduced into market in the year 2005 and 2009, respectively, as novel GLP-1 analogues. Each of them decreases HbA_{1c} level and provides the weight loss. The main difference between these agents is their half-lives in a circulation. Exenatide yields therapeutic effect in 4-6 hours, whereas half-life of liraglutide was enhanced to 12-15 hours (Gentilella et al., 2009).
- **bromocriptin** belongs to drugs used in the treatment of Parkinson disease. However, recently it has been shown that this therapeutic agent ameliorates insulin sensitivity, leading to enhancing glucose control and lowering the incidence of hypoglycaemia (Pijl et al., 2000). Although bromocriptin provides weight loss and diminishes concentration of plasma triglyceride and free fatty acids, it was evidenced that after using bromocriptin the patients suffer from several side effects, like nausea, hypotension and psychiatric disturbances.

Diabetes mellitus has been associated with the increased mortality risk due to non-diabetic factors, like several types of solid tumours, including the cancers of colon, breast and pancreas. Similar associations have been noted for central obesity and other conditions associated with increased levels of circulating insulin. These observations have given rise to the hypothesis that growth of these tumours, which are characterised by abnormal expression and function of the insulin–IGF-1 series of receptors, may be promoted by the trophic action of insulin interacting with these receptors. The cancer risk associated with diabetes may also be influenced by therapy in a given diabetic individual: for example, the risk of colon cancer is higher in individuals on insulin, patients on metformin are less likely to be diagnosed with cancer, and the risk of mortality from solid tumours is lower for metformin than for exogenous insulin or sulfonylureas. As a recognition dawns that cancer should be numbered among the complications of diabetes, the possibility that therapies for diabetes may influence tumour progression is likely to attract the increasing interest and concern. Furthermore, the observation that both endogenous insulin and exogenous insulin therapy are associated with tumour progression raises the questions as to the safety of the insulin analogues, which have subtly modified receptor binding properties and may accelerate the growth and proliferation of both healthy and tumour cell lines in culture.

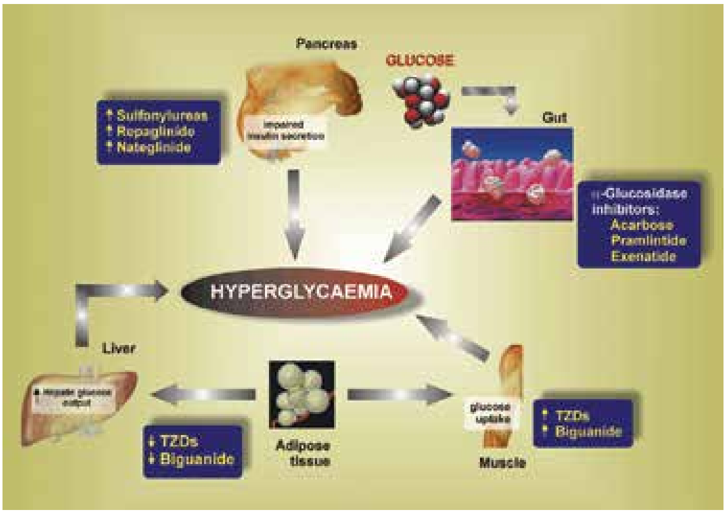


Fig. 8. Sites of action of major oral therapeutical agents used in the treatment of type 2 diabetes. Pharmacological therapies aimed at: inhibiting carbohydrate breakdown in the gut (α-glucosidase inhibitors), stimulating insulin secretion (sulfonylureas, repaglinide, nateglinide), suppressing hepatic gluconeogenesis (thiazolidinediones, biguanide), or accelerating skeletal muscle glucose metabolism (thiazolidinediones, biguanide) exhibit beneficial effects on fasting and/or postprandial plasma glucose, and consequently concord overall metabolic control in type 2 diabetic patients. Thus, a need for concerted combination therapy to successfully control a burden of hyperglycaemia in a majority of DM2 patients becomes a growing expectation by physicians and other health caregivers.

5.2 New agents – new hopes and new perspectives

The growing number of people with diabetes still requires novel combined treatments aimed at retardation of microvascular and macrovascular complications in the future. Therefore, anti-diabetic complications agents are sought to maintain an adequate glucose level in an organism. The number of anti-hyperglycaemic drugs delivered by subcutaneous injection increases constantly and to date there are numerous agents launched into market and examined in clinical trials. However, no such compounds/drugs that could be successfully applied in the treatment of diabetes have emerged hitherto, as the validated outcomes of clinical trials. Intensive studies are continued in order to develop a modern formula for effective amelioration of a burden associated with diabetes and late diabetic complications in diabetic patients in the future.

Studies on the formation of AGEs have been conducted with the goal to find promising pharmacological agents used in prevention or curing diabetic complications. The main target for these agents is to retard the formation of AGEs or to brake the AGE cross-links formed during Maillard reaction.

In order to prevent the AGEs formation the following therapeutic inhibitors have been developed and studied:

- **pyridoxamine (PM)**, a form of vitamin B_6, is thought to inhibit AGEs structures by capturing redox metal ions (Voziyan et al., 2003). In order to examine the anti-diabetic properties of PM under *in vivo* conditions, the model of diabetes mellitus induced by streptozotocin (STZ) was applied. The results of PM administration demonstrated the reduction in hyperglycaemia level and improvement in the plasma lactate/puryvate ratio. The decreased amounts of AGEs have also been noted (Degenhard et al., 2002). Moreover, it was also revealed that PM can act as inhibitor of protenuria and hyperlipidaemia in diabetes mellitus type 1 (Voziyan, 2005).
- **benfotiamine** is a derivative of vitamin B_1, which is involved in a limitation of methylglyoxal formation and lowering of AGEs accumulation (Gadau et al., 2006).
- **ALT-711** is a new stable analogue of N-phenal thiazolinium bromide (PTB). Its mechanism of action is based on preventing metal-catalyzed glycation (Price et al., 2001).
- **Acetylsalicylic acid (Aspirin®), salicylates and ibuprofen** possess an anti-inflammatory properties, which are important in the decreasing of the risk of cataract in people suffering from diabetes. As a radical scavengers, they may reduce the level of free radicals and/or chelate metal ions (Dinis et al., 1994).
- **chromium** deficiency is associated with the blood sugar irregularities of diabetes. Recent studies have demonstrated that chromium is effective in treating various types of diabetes, including types 1 and 2, gestational, and steroid-induced diabetes. Treatment of type 2 diabetes with chromium has led to improvement in blood glucose, insulin, and haemoglobin A_{1c} (HbA_{1c}) levels. The use of organic chromium complexes has been found to give superior results when compared to inorganic salts. Chromium has been found to be effective in reversing diabetes caused by the therapeutic use of glucocorticoids. Chromium picolinate (600 μg/day) was effective in lowering blood glucose almost twice (from 13.9 mM/L to 8.3 mM/L) in 47 of 50 patients. This therapy in patients was also able to reduce the doses of insulin and/or hypoglycaemic medications by half within one week from the beginning of chromium supplementation (Lamson & Plaza, 2002)

Among the newest agents tested in both *in vitro* and *in vivo* studies are poly(amido)amine PAMAM dendrimers and β-resorcylidene aminoguanidine (RAG), a derivative of aminoguanodine.

- **PAMAM dendrimers** are widely studied all over the world in almost every field of science. Their unique structure with nucleophilic character provided by surface amino groups may play an important role in the prevention or amelioration of hyperglycaemia. The ability of conjugation of PAMAM dendrimers to certain biologically relevant molecules makes them promising agents for using in biomedicine area, either as drugs or drug delivery systems. Experimental *in vitro* studies revealed that dendrimers appear very effective in scavenging glucose and reducing protein glycation (Fig. 9) (Labieniec & Watala, 2010). It was also evidenced that PAMAM dendrimer G4 administrated to rats with streptozotocin-diabetes acted as glucose scavenger and suppressed the accumulation of AGEs products, as well as some other markers of oxidative and carbonyl stress (Labieniec et al., 2008).

- **aminoguanidine** was the most promising oral antihyperglicaemic agent based on antioxidant capability and reduction of carbonyl reactive intermediates (Brownlee et al., 1986). Nevertheless, the B_6 vitamin depletion and oxidative stress production after using of aminoguanidine by patients with diabetes has been recorded, and therefore this compound was removed from clinical trials. Since then, new analogues of aminoguadine were synthesized in order to avoid the undesirable side effects of aminoguanidine itself. **β-resorcylidene aminoguanidine (RAG)** is one of these analogues, which seems to be the most promising and effective antioxidative and anti-diabetic agent amongst the others tested hitherto. The majority of studies have demonstrated that RAG is able to limit diabetes-associated long-term complications (protein glycation, AGEs formation, ROS level). Scientists suggest that RAG acts not only as antioxidative and/or anti-diabetic agent (Waczulíková et al., 2000, Vojtaśśak et al., 2008), but has also been shown to act as antithrombotic compound, independently of its anti-glycation activities (Watala et al., 2009).

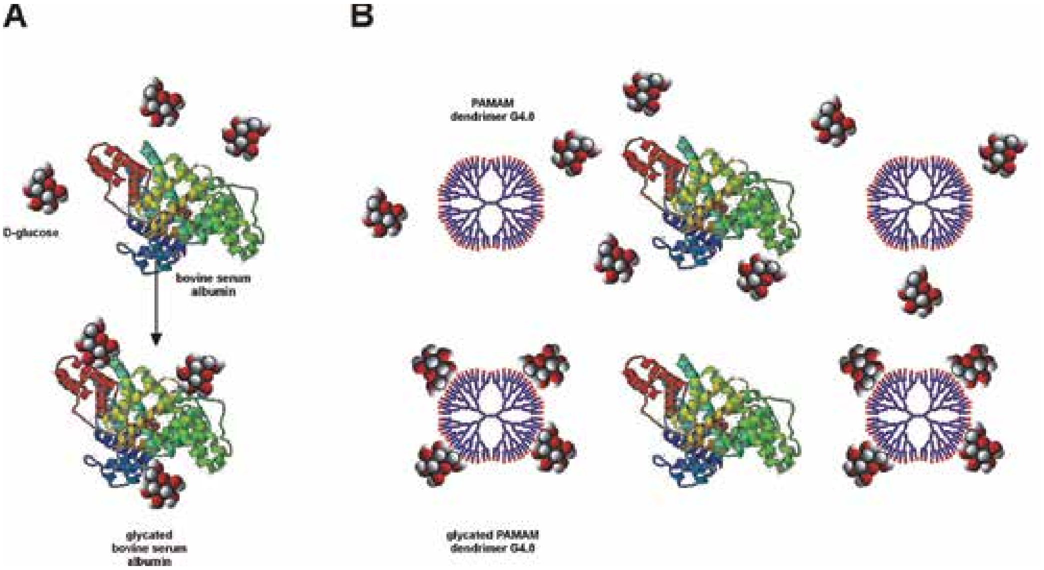


Fig. 9. The proposed mechanism of anti-glycation action of poly(amido)amine dndrimers (PAMAM).

Under conditions of excessive glucose the model protein (bovine serum albumin, BSA) undergoes extensive glycation (A), which becomes retarded and reduced to a large extent in the presence of poly(amido)amine dndrimers, generation 4.0 (PAMAM G4.0).

6. Conclusion

Diabetes is not merely a disease of impaired insulin sensitivity or insulin release, but may be a global metabolic dysfunction, including, among others, the collapse of the mitochondrial energy system. The role of the mitochondria in the metabolism associated with the pathophysiology of diabetes seems unique, mainly because a generation of ROS (which seem a natural part of mitochondrial physiology) constitutes a major treat in the development of diabetic sequelae. ROS play the central role in mediating various metabolic defects associated with the diabetic state. Therefore, inhibition of ROS production and/or enhancement of ROS scavenging will prove to be beneficial therapies. Hyperglycaemia elicits an increased ROS production, presumably from the mitochondrial respiratory chain. An important challenge for future research is to determine whether strategies aimed to improving mitochondrial functionality by using agents with anti-diabetic properties might have therapeutic potential in the treatment of diabetes. On the other hand, the better understanding of mitochondrial biology is still needed to facilitate the judicious selection and development of compounds/agents, which could be used as "mitochondrial drugs". Further studies are certainly required to better understand how these novel compounds and mitochondria may interact with each other, and how our understanding of such interaction might be utilized for the impaired mitochondrial functioning in the presence of diabetes. These investigations should also determine, which genetic, environmental, pharmacological and nutritional factors are possibly involved in an individual patient's susceptibility and, which treatments can be used safely in those patients, who suffer from heavy diabetes and are crushed by the burden of advanced long-term complications.

7. References

Adam-Vizi, V. & Chinopoulos, C. (2006). Bioenergetics and the formation of mitochondrial reactive oxygen species. *Trends Pharmacol Sci*, Vol. 27, No. 12, n.d., pp. 639-645

Barnaby, O.S.; Cerny, R.L., Clarke, W. & Hage, D.S.(2011). Comparison of modification sites formed on human serum albumin at various stages of glycation. *Clin Chim Acta,* Vol. 412, No. 3-4, (January 2011), pp. 277-285

Bellin, C.; de Wiza, D.H.; Wiernsperger, N.F. & Rosen, P. (2006). Generation of reactive oxygen species by endothelial and smooth muscle cells: influence of hyperglycemia and metformin. *Horm Metab Res,* Vol. 38, No. 11, (November 2006), pp. 732-739

Bliss, M. (1982). Banting's, Best's, and Collip's accounts of the discovery of insulin. *Bull Hist Med,* Vol. 56, No. 4, n.d., pp. 554-568

Bolli, G.B. & Owens, D.R. (2000). Insulin glargine. *Lancet,* Vol. 356, No. 9228, (August 2008), pp. 443-445

Bonnefont-Rousselot, D. (2002). Glucose and reactive oxygen species. *Curr Opin Clin Nutr Metab Care,* Vol. 5, No. 5, (September 2002), pp. 561-568

Borutaite V. (2010). Mitochondria as decision-makers in cell death. *Environ Mol Mutagen,* Vol. 51, No. 5, (March 2010), pp. 406-416

Brownlee, M.; Vlassara, H.; Kooney, A.; Ulrich, P. & Cerami, A. (1986). Aminoguanidine prevents diabetes-induced arterial wall protein cross-linking. *Science,* Vol. 232, No. 4758, (June 1986), pp. 1629-1632

Chiu, C.J. & Taylor. A. (2011). Dietary hyperglycemia, glycemic index and metabolic retinal diseases. *Prog Retin Eye Res.* Vol. 30, No. 1, (January, 2011), pp.18-53

Dinis, T.C.; Maderia, V.M. & Almeida, L.M. (1994). Action of phenolic derivatives (acetaminophen, salicylate, and 5-aminosalicylate) as inhibitors of membrane lipid peroxidation and as peroxyl radical scavengers. *Arch Biochem Biophys,* Vol. 315, No. 1, (November 1994), pp.161-169

Duchen, M.R. (2004). Mitochondria in health and disease: perspectives on a new mitochondrial biology. *Mol Aspects Med*, Vol. 25, No. 4, (August 2004), pp. 365-451

Edeas, M.; Attaf, D.; Mailfert, A.S; Nasu, M. & Joubet, R. (2010a) Maillard Reaction, mitochondria and oxidative stress: Potential role of antioxidants. *Pathologie Biologie,* Vol. 58, No. 3,(June 2010), pp. 220–225

Edeas, M. & Robert. R. (2010b). The Maillard reaction, its nutritional and physiopathological aspects. Introduction. *Pathol Biol, Vol.* 58, No. 3, (June 2010), pp.199

Ellis, G.P. (1959). The Maillard reaction. *Adv Carbohydr Chem, Vol.* 14, n.d., pp. 63-134

Frey, T.G. & Mannella C.A. (2000). The internal structure of mitochondria. Trends *Biochem Sci*, Vol. 25, No. 7, (July 2000), pp. 319-324

Gadau, S.; Emanueli, C.; Van Linthout, S.; Graiani, G.; Todaro, M.; Meloni, M.; Campesi, I.; Invernici, G.; Spillmann, F.; Ward, K. & Madeddu, P. (2006). Benfotiamine accelerates the healing of ischaemic diabetic limbs in mice through protein kinase B/Akt-mediated potentiation of angiogenesis and inhibition of apoptosis. *Diabetologia,* Vol. 49, No. 2, (February 2006), pp. 405-420

Garg, S.K.; Ellis, S.L. & Ulrich, H. (2005). Insulin glulisine: a new rapid-acting insulin analogue for the treatment of diabetes. *Expert Opin Pharmacother.* Vol. 6, No.4,(April 2005), pp. 643-651

Gerich, J.E. (1989). Oral hypoglycemic agents. *N Engl J Med,* Vol. 321, No. 18, (November 1989), pp.1231-1245

Goldberg, T.; Cai, W.; Peppa, M.; Dardaine, V.; Baliga, B.S.; Uribarri, J. & Vlassara. H. (2004). Advanced glycoxidation end products in commonly consumed foods. *J Am Diet Assoc,* Vol. 104, No. 8, (August 2004), pp.1287-1291

Heise, T.;. Bott, S.; Rave, K.; Dressler, A.; Rosskamp, R. & Heinemann, L. (2002). No evidence for accumulation of insulin glargine (LANTUS): a multiple injection study in patients with Type 1 diabetes. *Diabet Med,* Vol. 19, No. 6, (June 2002), pp. 490-495

Hagedorn, H.C. (1937). Protamine Insulinate: (Section of Therapeutics and Pharmacology). *Proc R Soc Med,* Vol. 30, No. 6, (April 1937), pp. 805-814

Harman, D. (1956). Aging: a theory based on free radical and radiation chemistry. *J Gerontol,* Vol. 11, No. 3, (July, 1956), pp. 298-300

Havelund, S.; Plum, A.; Ribel, U.; Jonassen, I.; Volund, A.; Markussen, J. & Kurtzhals, P. (2004). The mechanism of protraction of insulin detemir, a long-acting, acylated analog of human insulin. *Pharm Res,* Vol. 21, No. 8, (August 2004), pp. 1498-1504

Hodge, J. E. (1955). The Amadori rearrangement. *Adv Carbohydr Chem,* Vol.10, n.d., pp.169-205

Howey, D.C.; Bowsher, R.R;. Brunelle, R.L. & Woodworth, J.R. (1994). [Lys(B28), Pro(B29)]-human insulin. A rapidly absorbed analogue of human insulin. *Diabetes,* Vol. 43, No. 3, (March 1994), pp. 396-402

Hunt, J.V.; Dean, R.T. & Wolff, S.P. (1988). Hydroxyl radical production and autoxidative glycosylation. Glucose autoxidation as the cause of protein damage in the experimental glycation model of diabetes mellitus and ageing. *Biochem J,* Vol. 256, No. 1, (November, 1988), pp. 205-212

Huss, J. M. & Kelly, D.P. (2005). Mitochondrial energy metabolism in heart failure: a question of balance. *J Clin Invest,* Vol. 115, No. 3, (March, 2005), pp. 547-555

Jaeger, H.; Janositz, A. & Knorr, D. (2010). The Maillard reaction and its control during food processing. The potential of emerging technologies. *Pathol Biol.* Vol. 58, No. 3, (June 2010), pp. 207-213

King, H. & Rewers, M. (1993). Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults. WHO Ad Hoc Diabetes Reporting Group. *Diabetes Care,* Vol. 16, No. 1, (January, 1993), pp.157-177

King, H.; Aubert, R.E. & Herman, W.H. (1998). Global burden of diabetes, 1995-2025:prevalence, numerical estimates, and projections. *Diabetes Care,* Vol. 21, No. 9, (September 1998), pp. 1414-1431

Kuznetsov, A. V. & Margreiter, R. (2009). Heterogeneity of mitochondria and mitochondrial function within cells as another level of mitochondrial complexity. *Int J Mol Sci,* Vol. 10, No. 4, April 2009, pp. 1911-1929, ISSN 1422-0067

Labieniec, M.; Ulicna, O.; Vancova, O.; Glowacki, R.; Sebekova, K.; Bald, E.; Gabryelak, T. & Watala, C. (2008). PAMAM G4 dendrimers lower high glucose but do not improve reduced survival in diabetic rats. *Int J Pharm,* Vol. 364, No. 1, (November 2008), pp. 142-149.

Labieniec, M. & Watala C. (2010). Use of poly(amido)amine dendrimers in prevention of early non-enzymatic modifications of biomacromolecules. *Biochimie,* Vol. 92, No. 10, (October 2010), pp. 1296-1305

Lamson, D.W. & Plaza, S.M. (2002). The safety and efficacy of high-dose chromium. *Altern Med Rev,* Vol. 7, No. 3, (June 2002), pp. 218-235

Logan, D.C. (2006). The mitochondrial compartment. *J Exp Bot*, Vol. 5, No. 6, March 2006, pp. 1225-1243

Lowell, B.B. & Shulman, G.I. (2005). Mitochondrial dysfunction and type 2 diabetes. *Science*, Vol. 307, No. 5708, (January 2005), pp. 384-387

Luft, R. (1994). The development of mitochondrial medicine. *Proc Natl Acad Sci U.S.A,* Vol. 91, No. 19, (September 1994), pp. 8731-8738

Mannella, C.A. (2008). Structural Diversity of Mitochondria: Functional Implications. *Ann NY Acad Sci,* Vol. 1147, pp. 171–179

Marko, D.; Habermeyer, M.; Kemeny, M.; Weyand, U.; Niederberger, E., Frank, O. & Hofmann, T. (2003). Maillard reaction products modulating the growth of human tumor cells in vitro. *Chem Res Toxicol.* Vol. 16, No. 1, (January, 2003), pp. 48-55

Miyata, T.; Sugiyama, S.; Saito, A. & Kurokawa, K. (2001). Reactive carbonyl compounds related uremic toxicity ("carbonyl stress"). *Kidney Int Suppl,*Vol. 78, (February 2001), pp. S25-S31

Mudaliar, S R.; Lindberg, F.A.; Joyce, Beerdsen, M.P.; Strange, P.; Lin, A. & Henry, R.R. (1999). Insulin aspart (B28 asp-insulin): a fast-acting analog of human insulin: absorption kinetics and action profile compared with regular human insulin in healthy nondiabetic subjects. *Diabetes Care,* Vol. 22, No. 9, (September 1999), pp. 1501-1506

Nazaret, C.; Heiske, M.; Thurley, K. & Mazat, J.P. (2008). Mitochondrial energetic metabolism: A simplified model of TCA cycle with ATP production. *J Theor Biol,* Vol. 258, No. 3, (June 2009), pp. 455-464

Oliveira, P.J.; Rolo, A.P.; Seica, R.; Santos, M.S.; Palmeira, C. M. & Moreno, A. J. (2003). Reduction in cardiac mitochondrial calcium loading capacity is observable during alpha-naphthylisothiocyanate-induced acute cholestasis: a clue for hepatic-derived cardiomyopathies? *Biochim Biophys Acta,* Vol. 1637, No. 1, (January 2003), pp. 39-45

Paradies, G.; Petrosillo, G.; Paradies, V. & Ruggiero, F.M. (2010). Oxidative stress, mitochondrial bioenergetics, and cardiolipin in aging. *Free Radic Biol Med*, Vol. 48, No. 10, (May 2010), pp. 1286–1295

Peppa, M.; Goldberg, T.; Cai, W.; Rayfield, E. & Vlassara, H. (2002). Glycotoxins: a missing link in the "relationship of dietary fat and meat intake in relation to risk of type 2 diabetes in men". *Diabetes Care,* Vol. 25, No. 10, (November, 2002), pp.1898-1899

Patel, S.P. & Katyare S.S. (2006). Insulin-status-dependent modulation of FoF1-ATPase activity in rat liver mitochondria. *Lipids,* Vol. 41, No. 7, (July 2007), pp. 695-703

Petersen, K.F.; Befroy, D.; Dufour, S.; Dziura, J.; Ariyan, C.; Rothman, D.L.; DiPietro, L.; Cline, G.W. & Shulman, G.I. (2003). Mitochondrial dysfunction in the elderly: possible role in insulin resistance. *Science,* Vol. 300, No. 5622, (May 2003), pp. 1140-1142

Pinti, M.; Nasi, M.; Gibellini, L.; Roat, E.; De Biasi, S.; Bertoncelli, L. & Cossarizza A. (2010). The role of mitochondria in HIV infection and its treatment. *J Exp Clin Med*, Vol. 2, No. 4, (August 2010), pp. 145–155

Pijl, H.; Ohashi, S.; Matsuda, M.; Miyazaki, Y.; Mahankali, A.; Kumar, V.; Pipek, R.; Iozzo, P.; Lancaster, J.L.; Cincotta, A.H. & DeFronzo, R.A. (2000). Bromocriptine: a novel approach to the treatment of type 2 diabetes. *Diabetes Care,* Vol. 23, No. 8, (August 2000), pp. 1154-1161.

Price, D.L.; Rhett, P.M.; Thorpe, S.R. & Baynes, J.W. (2001). Chelating activity of advanced glycation end-product inhibitors. *J Biol Chem,* Vol. 276, No. 52, (December 2001), pp. 48967-48972

Ravelojaona, V.; Peterszegi, G.; Molinari, J.; Gesztesi, J.L. & Robert, L. (2007). Demonstration of the cytotoxic effect of Advanced Glycation Endproducts (AGE-s). *J Soc Biol,* Vol. 201, No. 2, n.d., pp.185-188

Rolo, A.P. & Palmeira, C.M. (2006). Diabetes and mitochondrial function: Role of hyperglycemia and oxidative stress. *Toxicol Appl Pharm*, Vol. 212, No. 2, (April 2006), pp. 167-178

Sell, D.R.; Lane, M.A.; Johnson, W.A.; Masoro, E.J.; Mock, O.B.; Reiser, K.M.; Fogarty, J.F.; Cutler, R.G.; Ingram, D.K.; Roth, G.S. & Monnier, V.M. (1996). Longevity and the genetic determination of collagen glycoxidation kinetics in mammalian senescence. *Proc Natl Acad Sci USA,* Vol. 93, No. 1, (January 1996), pp. 485-490

Shen, G X. (2010). Oxidative stress and diabetic cardiovascular disorders: roles of mitochondria and NADPH oxidase. *Can J Physiol Pharmacol,*Vol. 88, No. 3, (March, 2010), pp. 241-248

Singh, R.; Barden, A.; Mori, T. & Beilin, L. (2001). Advanced glycation end-products: a review. *Diabetologia,* Vol. 44, No. 2, (February 2001), pp.129-146

Srinivasan, K. & Ramarao, P. (2007). Animal models in type 2 diabetes research: an overview. *Indian J Med Res,* Vol. 125, No. 3, (March 2007), pp.451-472

Szewczyk, A. & Wojtczak, L. (2002). Mitochondria as a pharmacological target. *Pharmacol Rev,* Vol. 54, No. 1, n.d., pp. 101–127

Tessier, F.J. & Birlouez-Aragon, I. (2010). Health effects of dietary Maillard reaction products: the results of ICARE and other studies. *Amino Acids*, (October 2010), DOI: 10.1007/s00726-010—0776-z

Thornalley, P.J. (1998). Cell activation by glycated proteins. AGE receptors, receptor recognition factors and functional classification of AGEs. *Cell Mol Biol, Vol.* 44, No. 7, (November, 1998), pp. 1013-1023

Turner, R.C.; Cull, C.A.; Frighi, V. & Holman. R.R. (1999). Glycemic control with diet, sulfonylurea, metformin, or insulin in patients with type 2 diabetes mellitus: progressive requirement for multiple therapies (UKPDS 49). UK Prospective Diabetes Study (UKPDS) Group. *JAMA,* Vol. 281, No. 21, (June, 1999), pp.2005-2012

Urios, P.; Grigorova-Borsos, A.M.; Peyroux, J. & Sternberg, M. (2007). Inhibition of advanced glycation by flavonoids. A nutritional implication for preventing diabetes complications?. *J Soc Biol,* Vol. 201, No. 2, n.d., pp. 189-198

Vlassara, H. & Palace, M.R. (2002). Diabetes and advanced glycation endproducts. *J Intern Med,* Vol. 251, No. 2 (February 2002), pp. 87-101

Vojtassak, J.; Blasko, M.; Danisovic, L. Sr.; Carsky, J.; Durikova, M.; Repiska, V.; Waczulikova, I. & Bohmer, D. (2008). In vitro evaluation of the cytotoxicity and genotoxicity of resorcylidene aminoguanidine in human diploid cells B-HNF-1. *Folia Biol,* Vol. 54, No. 4, n.d., pp.109-114

Voziyan, P.A.; Khalifah, R.G.; Thibaudeau, C.; Yildiz, A.; Jacob, J.; Serianni, A.S. & Hudson, B.G. (2003). Modification of proteins in vitro by physiological levels of glucose: pyridoxamine inhibits conversion of Amadori intermediate to advanced glycation end-products through binding of redox metal ions. *J Biol Chem,* Vol. 278, No. 47, (November 2003), pp. 46616-46624

Voziyan, P. A. & Hudson, B.G. (2005). Pyridoxamine as a multifunctional pharmaceutical: targeting pathogenic glycation and oxidative damage. *Cell Mol Life Sci,* Vol. 62, No.15, (August 2005), pp.1671-1681

Waczulikova, I.; Sikurova, L.; Bryszewska, M.; Rekawiecka, K.; Carsky, J. & Ulicna, O. (2000). Impaired erythrocyte transmembrane potential in diabetes mellitus and its possible improvement by resorcylidene aminoguanidine. *Bioelectrochemistry.* Vol. 52, No. 2, (December 2000), pp. 251-256

Waldbaum, S. & Patel, M. (2009). Mitochondria, oxidative stress, and temporal lobe epilepsy. *Epilepsy Res*, Vol. 88, No. 1, (January 2010), pp. 23-45

Watala, C.; Dobaczewski, M.; Kazmierczak, P.; Gebicki, J.; Nocun, M.; Zitnanova, I.; Ulicna, O.; Durackova, Z.; Waczulikova, I.; Carsky, J. & Chlopicki, S. (2009). Resorcylidene aminoguanidine induces antithrombotic action that is not dependent on its antiglycation activity. *Vascul Pharmacol,* Vol. 51, No. 4, (October 2009), pp. 275-283

Watkins, P.B. & Whitcomb, R.W. (1998). Hepatic dysfunction associated with troglitazone. *N Engl J Med,* Vol. 338, No. 13, (March 1998), pp. 916-917

Westwood, M.E.; Argirov, O.K.; Abordo, E.A.; & Thornalley, P.J. (1997). Methylglyoxal-modified arginine residues--a signal for receptor-mediated endocytosis and degradation of proteins by monocytic THP-1 cells. *Biochim Biophys Acta* Vol. 1356, No. 1, (March 1997), pp. 84-94, 1997

Wierusz-Wysocka, B.; Wysocki, H.; Byks, H.; Zozulinska, D.; Wykretowicz, A. & Kazmierczak, M. (1995). Metabolic control quality and free radical activity in diabetic patients. *Diabetes Res Clin Pract,* Vol. 27, No. 3, (March, 1995), pp.193-197

Wirth, D.D.; Baertschi, S.W.; Johnson, R.A.; Maple, S.R.; . Miller, M.S.; Hallenbeck, D.K. & Gregg, S.M. (1998). Maillard reaction of lactose and fluoxetine hydrochloride, a secondary amine. *J Pharm Sci,* Vol. 87, No. 1, (January 1998), pp.31-39

Yamamoto, M.; Ozawa, K. & Tobe T. (1981). Roles of high blood glucose concentration during hemorrhagic shock in alloxan diabetic rats. *Circ Shock,* Vol. 8, No. 1, n.d., pp. 49-57

6

Red Palm Oil and Its Antioxidant Potential in Reducing Oxidative Stress in HIV/AIDS and TB Patients

O. O. Oguntibeju, A. J. Esterhuyse and E. J. Truter
Oxidative Stress Research Centre, Department of Biomedical Sciences,
Faculty of Health & Wellness Sciences, Cape Peninsula University of Technology,
South Africa

1. Introduction

1.1 HIV and TB

Scientific evidence has shown that HIV infection is caused by a retrovirus, the Human Immunodeficiency Virus (HIV) which is a ribonucleic acid (RNA) virus so designated because of its genome that encodes an unusual enzyme, reverse transcriptase (RT) that enables the virus to make copies of its own genome as DNA in its host's cells (human T4 helper lymphocytes) (Oguntibeju *et al.*, 2008).

The drastic increase in the number of people infected with HIV is not peculiar to a particular racial group, country or community despite multidimensional efforts which have been made to combat this scourge (Weiss, 1996; Oguntibeju *et al.*, 2007a). It is reported that the virus selectively attacks and depletes T-lymphocyte bearing $CD4^+$ cells (T-helper cells) causing a predisposition to opportunistic infections and malignancies (Weiss, 1996) and ultimately resulting in Acquired Immunodeficiency Syndrome (AIDS).

The cellular receptors to HIV are cells that express the $CD4^+$ T cell receptor ($CD4^+$ T-cells or T4-cells) as well as other white blood cells including monocytes and macrophages. Glial cells in the central nervous system, chromaffin cells in the intestine and Langerhans cells in mucous membranes and skin that express $CD4^+$ T cell receptors can also be infected (Paxon *et al.*, 1996). The possibility that there are other cellular targets apart from $CD4^+$T-cells is proved by the likelihood of neurons that can be infected. This creates the possibility of the presence of co-receptors in addition to $CD4^+$ T cells to mediate fusion between HIV and its target cells (Grossman and Heberman, 1997).

Recognition of the $CD4^+$ T-cells by HIV-1 envelope glycoprotein (gp120) to which the virus binds and enters host cells to initiate rapid replication cycles (Oguntibeju *et al.*, 2007b) depicts significant cytopathic consequences of HIV infection of $CD4^+$ T-cells (Bartlett, 1998) and is an important factor in the initiation of HIV infection. The shed virions which are immunogenic, stimulate B cells to produce humoral antibodies and plasma cells through lymphoid hyperplasia that ultimately results in decreased number of infected cells as the $CD4^+$ T-cells migrate through the germinal cells. The depletion in the number of $CD4^+$ T-cells exceeds the formation of new cells and may maintain this phase for many years resulting in general disorganization of the lymphoid nodes, loss of lymphoid function and integrity.

2. Physiological and biochemical mechanisms of the role of oxidative stress in HIV/AIDS & TB disease complications

After initial infection of the human host, the pace of immunodeficiency development, susceptibility to infection and malignancies become manifest and are generally associated with the rate of CD4$^+$ decline (Enger *et al.*, 1996). The rate of CD4$^+$ decline varies considerably from person to person and is not constant throughout all the stages of HIV infection. Though the virological and immunological process that take place during the period of rapid fall in the number of CD4$^+$ T-cells are poorly understood, Koot *et al.* (1996) reported that acceleration of the decline of CD4$^+$ T-cells heralds the progression of the disease that is associated with the increasing rate of HIV-1 replication *in vivo* and declining cell-mediated immune response. Studies have shown that the host immunological alterations due to HIV infection result in progressive development of opportunistic infections and malignancy and is chiefly mediated /induced by deregulation of a cytokine profile production of ROS which also plays a role in the viral replication. *In vitro* studies have shown activation of viral replication by induction of TNFαβ (Allard *et al.*, 1998).

Das and co-workers (1990) stated that excessive production of reactive oxygen species (ROS) such as superoxide anions, OH$^-$ radicals and H_2O_2, may be related to increased activation of PMN leucocytes during infection. This is influenced by the pro-oxidant effect of TNFα produced by the activated macrophages during the course of HIV infection and secretion of pro-inflammatory cytokines IL-1, IL-6, and IL-8 (Kiedzierska and Crowe, 2001). Gil *et al.* (2003) further established the presence of substantial oxidative stress in HIV infection which they attributed to the role of viral proteins that increases ROS intracellularly, therefore increasing the apoptotic index and depleting the CD4$^+$ T-lymphocyte population. The ROS thus produced can attack the double bonds in polyunsaturated fatty acids, inducing lipid peroxidation which may result in more oxidative cellular damage to the membrane lipids, proteins or DNA. Chronic oxidative stress experienced by patients infected by HIV leads to a condition in which there is increased consumption of antioxidants (such as Vitamin C, and E , selenium, and carotenoids) as well as micronutrients/ trace elements (Januga *et al.*, 2002). Stephen (2006), therefore, concluded that persistent chronic inflammation such as found in HIV infection places a long-term strain on antioxidant defenses, impair immune functions, increases the severity of the disease as well as increases the antioxidant requirement by the infected individual.

Progression of HIV to AIDS in developed countries after initial infection is about 10-12 years for adults in the absence of antiviral therapy. However, some individuals manifest full blown AIDS within 5 years of infection whereas others survive long term (>10 years) asymptomatic HIV-1 infection without a significant decline in CD4$^+$ T-cell count. Such delay in the progression of HIV to AIDS may be attributed to either infection with genetically defective HIV-1 variants or effective host antiviral immune response where the individual has active cytotoxic T-cell responses against HIV-1 infected cells (Haase, 1999).

Ever since Robert Koch made the landmark discovery that tuberculosis is caused by the infectious agent *Mycobacterium tuberculosis* (Koch, 1882), it has remained a major global health threat. Although in developed countries the rates of infection has fallen in the past century, the number is now again increasing which results in over 2000 deaths in developed countries annually due to changes in social structures in cities, the HIV epidemic, and failure to improve treatment programs (Frieden *et al.*, 1995). The increased death rate recorded as a result of poverty, poor living conditions and inadequate medical care in developing/Third World countries is further compounded by the emergence of multi-drug resistance where antibiotics are either of inferior quality, or are not used for a sufficient period of time to control the disease (O'Brien, 2001).

The recent increase in reported pulmonary tuberculosis (PTB) cases globally can be attributed to the increased susceptibility to opportunistic infections in HIV-infected persons. The highest prevalence of cases is reported to be in Asia (China, India, Indonesia, Bangladesh and Pakistan) and Africa with over 90% of global TB infections and deaths annually. TB cases occur predominantly in the economically productive 15-49 year age group (Dye *et al.*, 1999). Like HIV infection, TB also has a long latency period with symptomatic presentation occurring from 3 months to decades after establishment of the infection (Jagirdar and Zagzag, 1996).

TB is caused by an obligate pathogen that does not replicate outside its host environment (Mathema *et al.*, 2006) and is spread by aerosolization of droplets bearing *M. tuberculosis* particles released from the lung or larynx during coughing, sneezing, or talking in poorly ventilated areas. The particles of 1-5 μm in diameter, are inhaled and phagocytosed by resident alveolar macrophages. A vigorous immune response involving cytokines and a large number of chemokines ensues (Roach *et al.*, 2002). Protective immunity is characterized by granuloma formation that consists of primarily activated *M. tuberculosis* infected macrophages and T-cells. Medlar (1955), noted tissue necrosis and cavitations in over 10% of presumed immuno-competent patients and postulated that this was due to non-containment of continual bacterial replication (doubling time of 25-32 hours) that resulted in disease symptoms and its associated pathology. This response presumably initially limits infection to the primary site of invasion (the lung parenchyma and local draining lymph nodes known as the Ghon complex) in the majority of immuno-competent individuals (Bloom and Murray, 1992,).

Increased reactive oxygen species (ROS) has been reported in patients with TB. Excessive endogenously produced ROS in activated phagocytes of TB patients that escape to its surroundings can damage tissue or cellular DNA as well as impair immune function (Madebo *et al.*, 2003). It has been shown that the bactericidal potency of the myeloperoxidase-H_2O_2-halide system of neutrophilic granules demonstrates the bactericidal activities of the phagocytes that invariably produce increased ROS and reactive nitrogen intermediates (RNI) during phagocytic respiratory burst. Lower antioxidant potential as shown by a significant reduction of enzymatic antioxidants (superoxide dismutase, catalase) and non-enzymatic antioxidants (glutathione) as well as high malondialdehyde (MDA) concentrations suggest increases in the generation of ROS due to lipid peroxidation (Reddy *et al.*, 2004).

Di Massio and co-workers (1991), reported significantly reduced vitamin C and α-tocopherol levels in TB patients. These are integral components of antioxidants, which, when present in sufficient quantity, may act synergistically to protect cells from oxidative stress induced damage in TB patients. Several factors such as inadequate nutrients, malnutrition, nutrient malabsorption, low food intake, inadequate nutrient release from the liver, acute infections including other than HIV, may be the cause of low or impaired antioxidant capacity in TB patients (Das *et al.*, 1990).

Presentation of disease is variable as regards the pathology as well as infections in a variety of tissues such as the meninges, lymph nodes, and tissue of the spine, where response to antibiotic medication/treatment to clear the bacilli from tissues, partial reversal of the granulomatous process and subsequent clearance from the sputum may be found in clinical cases (Jargirdar and Zagzag, 1996). The progression and nature of disease may be affected by factors such as conditions that negatively impact on the host immune system (for example, poorly controlled diabetes mellitus, renal failure, chemotherapy, malnutrition or intrinsic host susceptibility (Madebo *et al.*, 2003). Host susceptibility has been known to affect endogenous re-activation and exogenous re-infection by the bacilli.

3. Reactive oxygen species and reactive nitrogen species and their effects on biological macromolecules and organs

Reactive oxygen species / reactive nitrogen species (ROS/ RNS) are constantly being formed in living organisms (Ceconi *et al.*, 2003). In the course of oxygen metabolism, 1- 5% of all inhaled oxygen becomes ROS (Berk, 2007). Endogenously, ROS are produced from various sources such as mitochondria, activated macrophages and leucocytes, oxidase enzyme (NADPH), cyclo-oxygenase and lipoxygenase (Zalba *et al.*, 2006). Reactive oxygen species have oxidation ability and are classified either as free radicals (superoxide anion $O_2^{\cdot-}$, hydroxyl radical $OH^{\bullet}$, nitric oxide NO) or as non-free radicals (hydrogen peroxide H_2O_2, peroxynitrite $ONOO^-$) (Higashi *et al.*, 2006). Previous studies have shown the involvement of ROS in physiological and pathophysiological conditions (Fortuño *et al.*, 2005; Berk, 2007; Heistad *et al.*, 2009). At low concentrations, ROS are involved in normal cell signaling pathways (smooth muscle and endothelial cell growth, apoptosis and survival) and in the remodeling of vessel walls (Fortuño *et al.*, 2005; Heistad *et al.*, 2009). At high concentrations, ROS are identified as harmful compounds and constitute an important risk factor for the development of many diseases such as cardiovascular diseases (Maxwell & Lip, 1997; Heistad *et al.*, 2009).

4. Oxidative stress

Oxidative stress occurs when there is a dysfunction in the overall balance between the production of reactive oxygen and nitrogen species and the antioxidant defense mechanisms (Ceconi *et al.*, 2003; Berk, 2007; Barbosa *et al.*, 2008).

Oxidative stress is believed to play a critical role in the complications and pathophysiology of HIV/AIDS, TB and cardiovascular diseases (Heistad *et al.*, 2009). In the context of oxidative stress in HIV/AIDS and TB, the major vascular ROS is the superoxide anion ($O_2^{\cdot-}$) which is predominantly generated by the NADPH oxidase enzyme (Fortuño *et al.*, 2005). Superoxide is normally dismutased to hydrogen peroxide (H_2O_2) by a family of superoxide dismutase (intracellular Cu/Zn SOD, MnSOD or extracellular Cu/Zn SOD) (Hamilton *et al.*, 2004). Hydrogen peroxide is converted into oxygen and water by catalase enzymes or by glutathione peroxidase (GPx) in the presence of reduced glutathione (Hamilton *et al.*, 2004; Zalba *et al.*, 2006). In the pathophysiological process of oxidative stress, excess superoxide has many effects, superoxide combines with NO to form peroxynitrite. Peroxynitrite is a highly toxic oxidant which causes damage to cells of the vascular wall through oxidation of lipids (lipid peroxidation), proteins (protein nitrosilation) and nucleic acids with superoxide. This causes vascular dysfunction by removing the protective effects of NO (Heistad *et al.*, 2009), initiates the development of vascular inflammatory state (Hamilton *et al.*, 2004), facilitates the oxidation of LDL, causing development of artherosclerotic lesions (Zalba *et al.*, 2006) and triggers apoptotic cell death (Ceconi *et al.*, 2003).

Accumulating evidence has suggested that oxidative stress, mainly through lipid peroxidation, represents an important risk factor in the development of cardiovascular diseases and complications in HIV/AIDS and TB (Waterfall *et al.*, 1997; Ceconi *et al.*, 2003). In fact, lipid peroxidation leads to membrane disruption and release of highly reactive free radicals (such as MDA) which can severely alter the cellular function (Ceconi *et al.*, 2003) (Table 1).

S/N	**Oxidants**	**Reactions**
1.	Production of superoxide	O_2 + electron $\longrightarrow$ O_2
2.	NADPH – oxidase	$2O_2 + NADPH \longrightarrow 2O_2^- + NADP + H^+$
3.	Superoxide dismutase	$O_2^- + O_2^- + 2H^+ \longrightarrow H_2O_2 + O_2$
4.	Calalase	$H_2O_2 \longrightarrow 2H_2O + O_2$
5.	Myeloperoxidase	$H_2O_2 + x^- + H^+ \longrightarrow HOX + H_2O$
6.	Glutathione peroxidase (Se-dependant)	2GSH + R-O-OH $\longrightarrow$ GSSG + H2O + ROH
7.	Fenton reaction	$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} OH + OH^-$
8.	Iron-catalyzed Haber Weiss reaction	$O_2^- + H_2O_2 \longrightarrow O_2 + OH + OH^-$
9.	Glucose-6-phosphate dehydrogenase	G-6-P + NADP $\longrightarrow$ 6-Phosphogluconate + NADPH + H^+
10.	Glutathione reductase	G-S-S-G+NADPH + H^+ $\longrightarrow$ 2GSH + NADH

Source: Murray, 2000

Table 1. Reactions in relation to oxidative stress in blood cells and various tissues.

5. Activation of oxygen

Oxygen is essential for energy metabolism and respiration but is has been implicated in many disease and degenerative conditions (Ceconi *et al.*, 2003). Activation of oxygen may occur by two different mechanisms: absorption of sufficient energy to reverse the spin on one of the unpaired electrons and monovalent reduction. Non-activated oxygen is a bi-radical. It can be activated by either reversing the spin on one of the unpaired electrons to form the singlet state or by reduction. In the monovalent reduction of oxygen, superoxide (O_2^-), hydrogen peroxide (H_2O_2), hydroxyl radical (OH) and finally, water (H_2O) is formed. Superoxide forms the hydroxyl radical (OOH) which is the protonated form of the superoxide anion radical (Gebick and Bielski, 1981; Ceconi *et al.*, 2003).

Numerous enzymes (peroxidases) use hydrogen peroxide as a substrate in oxidation reactions involving the synthesis of complex organic molecules. Haber and Weiss (1994) identified the hydroxyl radical as the oxidizing species in the reaction between H_2O_2 and ferrous salts.

$$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} + OH + OH^-$$

Most oxygen is consumed by the cytochrome oxidase enzyme in the mitochondrial electron transport system. Isolated mitochondria produce H_2O_2 and O^{2-} in the presence of NADH (Loschen *et al.*, 1974). The various Fe-S-proteins and NADH dehydrogenase have also been implicated as possible sites of superoxide and hydrogen peroxide formation (Waterfall *et al.*, 1997). Various oxidative processes including oxidation hydroxylations, dealkylations, deaminations, dehalogenation and desaturation occur in the smooth endoplasmic reticulum. Mixed function oxygenases that contain a heme moiety add an oxygen atom into an organic substrate using NADPH as the electron donor. The generalized reaction catalyzed by cytochrome P_{450} is:

$$RH + NADPH + H^+ + O_2 \longrightarrow ROH + NADP^+ + H_3O$$

Superoxide is produced by microsomal NADPH dependent electron transport involving cytochrome P_{450} (Valko *et al.*, 2007). One possible site at which this may occur is shown in Figure 1:

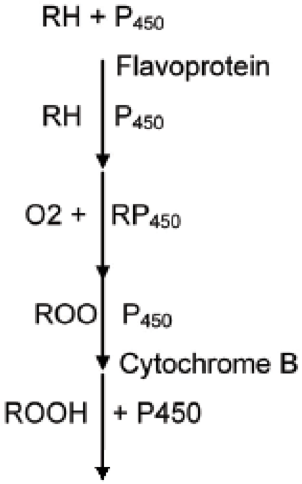


Fig. 1. Schematic presentation of the cytochrome P450 electron transport (Valko *et al.*, 2007).

In the peroxisomes and glyoxysomes, compartmentalized enzymes involved in the B-oxidation of fatty acids and glyoxylic acid cycle includes glycolate oxidase, catalase and various peroxidases. Glycolate oxidase produces H_2O_2 in a two-electron transfer from glycolate to oxygen (Lindqvist *et al.*, 1991). Xanthine oxidase, urate oxidase and NADH oxidase generate superoxide as a consequence of the oxidation of their substrates (Fridovich, 1970). The xanthine oxidase reaction is often used *in vitro* as a source of superoxide producing one mole of superoxide during the conversion of xanthine to uric acid (Fridovich, 1970). A superoxide-generating NADPH oxidase activity has been clearly identified in plasmalemma-enriched fractions (Valko *et al.*, 2007). These flavopoteins may produce superoxide by the redox cycling of certain quinones or nitrogenous compounds and NADPH oxidase reduces Fe^{3+} to Fe^{2+} converting it to a form that can be transported. Dysfunction of NADPH oxidase results in the formation of superoxide (Maxwell & Lip, 1997).

6. Mechanism of action of antioxidants in improving the immune status of AIDS and TB patients

Antioxidants are compounds that dispose, scavenge and suppress the formation of free radicals or oppose their actions. Free radicals, primarily the ROS, superoxide and hydroxyl radicals, which are highly reactive, with an unpaired electron in an atomic or molecular orbit, are generated under physiological conditions during aerobic metabolism (Semba & Tang, 1999; Champe *et al.*, 2005). Because free radicals are potentially toxic, they normally are inactivated or scavenged by antioxidants before they can damage lipids, proteins or nucleic acids.

Superoxide dismutase (SOD), catalase and glutathione peroxidase (GSH-Px) are the primary antioxidant enzymes involved in the direct elimination of ROS whereas glutathione transferase, glucose-6-phosphate dehydrogenase (G6PD) and copper-binding ceruloplasmin are secondary antioxidant enzymes which assist in maintaining a steady concentration of glutathione and NADPH required for optimal functioning of the primary antioxidant enzymes (Kiremidjian-Schumacher *et al.*, 1994; Champe *et al.*, 2005). Antioxidant enzymes require micronutrients such as selenium, iron, copper, zinc and manganese as cofactors for optimal catalytic activity and to act as effective antioxidant defence mechanisms. If homeostasis between the rate of formation of free radicals and the rate of neutralization of free radicals is not maintained, oxidative damage, known as oxidative stress occurs which further damages the already compromised immune system and consequently enhances HIV and TB progression (Cunningham-Rundles, 2001; Cunningham-Rundles *et al.*, 2005) and it has been reported that most of these antioxidants are derived from dietary sources (fruits and vegetables including red palm oil (Maxwell & Lip, 1997; Ebong *et al.*, 1999; Edem, 2002; Van Rooyen *et al.*, 2008).

7. Mechanism of action of red palm oil in improving the immune status of patients with AIDS and TB

Several studies have illustrated that red palm oil (RPO) is a rich cocktail of lipid-soluble antioxidants such as carotenoids (mostly α- and β-carotene, lycopenes), vitamin E (in the form of α- , β-, δ- tocotrienols and tocopherol) and ubiquinone (mostly coenzyme Q_{10}) (Ebong *et al.*, 1999; Edem, 2002; Van Rooyen *et al.*, 2008).

Feeding experiments using various animal models have highlighted that red palm oil is beneficial to health by reducing oxidative stress (Ebong *et al.*, 1999). Many studies have demonstrated the protective effects of red palm oil in a cardiac ischaemia/perfusion model of oxidative stress (Esterhuyse *et al.*, 2005; Bester *et al.*, 2006; Engelbrecht *et al.*, 2006) and modulation of the serum lipid profile in rats.

RPO is widely used as cooking oil in West and Central Africa and plays an essential role in meeting energy and essential fatty acid needs in many regions of the world. It contains many beneficial antioxidants and micronutrient compounds such as tocopherol, tocotrienol, lycopene, squalene, co-enzyme Q_{10}, physterol, glycolipids, phosphatides, calcium, phosphorus, iron, riboflavin, chlorophil, xanthophil, flavonoids, phospholipids, and carotenoid in addition to the equal proportion of saturated and unsaturated fatty acids such as oleic acid, linolenic acid, palmitic acid, linoleic acid, stearic acid and arachidic acid. It is known to be the richest source of carotenoids in terms of provitamin A equivalents i.e α and β carotenes (Sundram *et al.*, 2003) with its wide range of protective properties against disease and aging as well as being modulators for cellular processes / functions where photo-oxidative processes predominate by acting as scavengers of oxygen and peroxyl radicals (Van Rooyen *et al.*, 2008). Sebinova and co-workers (1991) documented the increased protection derived from a combination of tocopherol and tocotrienol and further revealed that tocotrienol offers a more efficient protection than tocopherol as it is preferentially consumed by ROS. It has been shown that fresh RPO has no adverse effect on body weight and morphology of body tissues. It also lowers the level of serum lipids and inhibits tumour growth (Kritchevsky, 2000), enhances intestinal uptake of protein and the metabolism of sulphur-amino acids and promotes reproductive capacity (Ebong *et al.*, 1999). Calcium, phosphorous, iron, riboflavin, chlorophyll, xanthophylls, flavonoids and phospholipids and equal proportion of saturated and unsaturated fatty acids have also been identified as part of its constituents (Sundram *et al.*, 2003).

A number of human feeding studies reported that palm oil diets show a reduction of blood cholesterol values ranging from 7%-38% (Mattson and Grundy, 1985; Bonanome and Grundy, 1988). A comparative study in young Australian adults showed that the total blood cholesterol, triglycerides and HDL-cholesterol levels of those fed on palm oil (palm olein) and olive oil were lower than those fed on the usual Australian diet (Choudhury *et al.*, 1995). A double-blind cross-over study (Sundram, 1997) showed that a palm olein rich oil diet is identical to an oleic-acid rich diet.

A study on fifty-one Pakistani adults showed that those given palm oil rich diets performed better than those on sunflower oil. Palm oil was found to increase HDL-cholesterol and Apo A-1 levels (Farooq *et al*, 1996). A group in Beijing, China compared the effects of palm oil, soybean oil, peanut oil and lard (Zhang *et al*, 1997a; Zhang *et al*, 1997b) and reported that palm oil caused a significant increase in the HDL-cholesterol as well as a significant reduction in LDL-triglycerides.

Sundram and co-workers (1992) performed a dietary intervention study on a free-living Dutch population who normally consumes diets high in fats. Using a double blind cross-over study design consisting of two periods of six weeks of feeding, the normal fat intake of a group of 40 male volunteers was replaced by 70% palm oil. The palm oil diet did not raise serum total cholesterol and LDL-cholesterol, and caused a significant increase in the HDL-cholesterol as well as a significant reduction in LDL-triglycerides.

The effects of palm olein and of canola oil on plasma lipids were examined in double blind experiments in healthy Australian adults (Truswell *et al*, 1992). Palm oil performed better than canola oil in raising the HDL-cholesterol (Truswell *et al*, 1992). Other studies have demonstrated that RPO supplementation has beneficial or neutral effects on serum total cholesterol (Zhang *et al.*, 1997a).

A cross-over feeding study showed that the blood cholesterol, triglycerides, HDL-cholesterol and LDL-cholesterol levels of palm olein and olive oil diets were comparable (Ng *et al.,* 1992). A Malaysian study (Ng *et al.*, 1991) was conducted to compare the effects of a diet containing palm oil (olein), corn oil and coconut oil on serum cholesterol. Coconut oil was found to raise serum total cholesterol by more than 10% whereas both the corn and palm oil diets reduced the total cholesterol; the corn oil diet reduced the total cholesterol by 36% and palm oil by 19%. A similar cholesterol-lowering effect of palm oil was observed in 110 students in a study conducted in Malaysia (Marzuki *et al.*, 1991). The study compared the effect of palm oil with that of soybean oil. Volunteers fed on palm oil (olein) and soybean oil for five weeks, with a six-week wash-out period, showed comparable blood cholesterol levels. However, the blood triglycerides were increased by 28% in those on the soybean oil diet.

Thus, the impact of palm oil on serum lipids is more like that of a mono-unsaturated rather than saturated oil. There appears to be several explanations: (1) Palm oil is made up of 50% unsaturated fats and the saturated fatty acids present are palmitic (90%) and stearic (10%). Stearic acid as well as palmitic acid do not raise blood cholesterol levels in people whose blood cholesterol levels are in the normal range (Hayes, 1993; Hayes *et al*., 1995; Hayes *et al*., 1991, Khosla and Hayes, 1994; Khosla and Hayes, 1992). (2) The vitamin E, particularly the tocotrienol present in palm oil can suppress the synthesis of cholesterol in the liver (Qureshi *et al*., 1991a; Qureshi *et al*, 1991b; Qureshi *et al*, 1980; Mcltosh *et al*, 1991). (3) The position of the saturated and unsaturated fatty acid chains in a triglyceride backbone of the palm oil molecule determines whether the fat will elevate the cholesterol level in the blood (Kritchevsky, 1988; Kritchevsky, 1996). In palm oil, 87% of the unsaturated fatty acid chains are found in position 2 of the carbon atom of the triglyceride backbone molecule (Ong & Goh, 2002). This could explain why palm oil is not cholesterol-elevating. (4) It has an anti-clotting effect and prevents the formation of thrombi in the blood vessels. Hornstra (1988) in the Netherlands first demonstrated that palm oil has an anti-clotting effect, and is as anti-thrombotic as the highly unsaturated sunflower seed oil. A human study (Kooyenga *et al.,* 1997) showed that tocotrienol (from palm oil) supplementation can reduce stenosis of patients with carotid atherosclerosis.

Holub *et al*, (1989) reported that the vitamin E in palm oil inhibits platelets from "sticking" to each other. Other supporting evidence showed that a palm oil diet increases the production of a hormone, prostacyclin that prevents blood-clotting or decreases the formation of a blood-clotting hormone, thromboxane (Sundram *et al*., 1990; Ng *et al*., 1992).

Kurfeld *et al*. (1990) in the United states, using a rabbit model, compared the effects of palm oil with hydrogenated coconut oil, cottonseed oil, hydrogenated cottonseed oil, and an American fat blend containing a mixture of butterfat, tallow, lard, shortening, salad oil, peanut oil and corn oil. At the end of a 14-month feeding period, coconut oil fed rabbits showed the most atherosclerotic lesions, while palm oil-fed rabbits showed less lesions compared to the coconut oil fed rabbits.

More than 70% of the vitamin A intake in Third World countries comes from fruits and vegetables in the form of carotenoids (Van Rooyen *et al.,* 2008). In humans and animals, carotenoids, an important constituent of palm oil, play an important role in protection against photo-oxidative processes by acting as oxygen and peroxyl radical scavengers. Their synergistic action with other antioxidants makes them an even more potent compound.

It has been suggested that different individual compounds exhibiting a variety of anti-oxidant activities may provide additional protection against oxidative stress when ingested simultaneously (Esterbauer *et al*., 1991). A combination of lipophilic anti-oxidants present in red palm oil results in an inhibition of lipid peroxidation which is significantly greater than the sum of the individual effects (Zhang *et al.*, 1995). This suggests that a cocktail of anti-oxidants may have a far more profound anti-oxidative effect due to the synergistic action of the individual compounds (Zhang *et al.*, 1995).

The antioxidant properties of RPO has been attributed to the synergistic actions of carotenoids and vitamin E in the presence of lycopene in a natural food environment and this might provide the ultimate dietary supplement to fight disease associated with oxidative stress (Van Rooyen *et al.*, 2008).

Conclusively, it could be said that oxidative stress plays a role in inflammatory and chronic diseases such as HIV/AIDS and TB and contribute significantly to depletion of immune factors, micronutrients and also promotes the progression of disease. Red palm oil could potentially retard the process due to its unique characteristics and also as it is known to be rich in several important antioxidants.

8. References

Allard JP, Aghdassi E, Chau J, Tam C, Koracs CM, Salit E & Walmsley SL (1998). Effect of vitamin E and C supplementation on oxidative stress and viral load in HIV-infected subjects. AIDS 12:1653-1659.

Barbosa KB, Bressan J, Zulet MA & Martínez-Hernández JA (2008). Influence of dietary intake on plasma biomarkers of oxidative stress in humans. Anales del Sistema Sanitario de Navarra 31(3): 259-80.

Bartlett JD (1998). Natural history and classification of HIV infection Chapter 2 In: Medical Management of Infection 1-15. Bartlett, JG (Ed.). Baltimore: Port City Press.

Berk, BC (2007). Novel approaches to treat oxidative stress and cardiovascular diseases. Trans Am Clin & Climatol Assoc 118: 209-14.

Bester DJ, Van Rooyen J, Du Toit EF & Esterhuyse AJ (2006). Red palm oil protects against the consequences of oxidative stress when supplemented with dislipidaemic diets. Med Tech SA 20: 3-8.

Bloom BR & Murray (1992). Tuberculosis: Commentary on a re-emergent killer. Sci 257: 1055-1064.

Bonanome A & Grundy SM (1988): Effect of dietary stearic acid and plasma cholesterol and lipoprotein levels. New Eng J Med 318:1124-1128.

Ceconi, C., Boraso, A., Cargnoni, A. & Ferrari, R. (2003). Oxidative stress in cardiovascular disease: myth or fact? Arch of Biochem & Biophys 420 (2): 217-21.

Champe PC, Harvey RA & Ferrier DR (2005). Biochemistry. RA Harvey & PC Champe (Eds) 3rd Ed Lippincott Williams & Wilkins, Ambler, PA, USA.

Choudhury N, Tan L & Truswell AS (1995): Comparison of palm olein and olive oil : Effects on plasma lipids and vitamin E in young adults. Am J Clin Nutr 61: 1043-1051

Cunningham-Rundles S (2001). Nutrition and the mucosal immune system. Current Opin Gastroenterol 17: 171-176.

Cunningham-Rundles S, McNeeley DF & Moon A (2005). Mechanisms of nutrient modulation of the immune response. J Allergy Clin Immunol 115 (6):1119-1128; quiz 1129.

Das UN, Podma M, Sogar PS, Ramesh G & Koratkar R (1990). Stimulation of free radical generation in human leucocytes by various agents including NF is a calmodulin – dependent process. Biochem Biophys Res Commun 67: 1030-6.

Di Massio P, Murphy M & Seis H (1991). Antioxidant defense systems: the role of carotenoids, tocopherols and thios. Am J Nutr 53: (suppl) 1945-2005.

Dye CS, Scheele P, Dulin V, Pathanira S and Raviglione MC.(1999). Global burden of tuberculosis, estimated incidence, prevalence and mortality. WHO. Global Surveillance and Monitoring Project, JAMA 282: 677-686.

Ebong PE, Owu DU & Isong EU.(1999). Influence of palm oil (Elaesis guineensis) on health. Plant Foods Hum Nutr 53:209-222.

Edem DO (2002). Palm oil: biochemical, physiological, nutritional, haematological and toxicological aspects of: a review. Plant Foods Hum Nutr 57: 319-41.

Engelbrecht AM, Esterhuyse AJ, du Toit EF. Lochner A & Van Rooyen J (2006). P38-MAPK and PKB/Akt, possible role players in red palm oil-induced protection of the isolated perfused rat heart? J Nutr Biochem 17: 265-271.

Enger CN, Graham J & Peng Y (1996). Survival from early, intermediate and late stages of HIV infection. J Am Med Assoc 275:1329-1334.

Esterbauer H, Schaur JS & Zollner H (1991). Chemistry and biochemistry of 4-hydroxynonenal, malondialdehyde and related aldehydes. Free Rad Biol Med 11:81-128.

Esterhuyse AJ, du Toit EF, Van Rooyen J (2005). Dietary red palm oil supplementation protects against the consequences of global ischemia in the isolated perfused rat heart. Asian Pac J Clin Nutr 14(4*): 340-347 (a).*

Farooq A; Amir ID & Syed AA (1996). Dietary fats and coronary risk factors modification-a human study. Proc PIPOC

Fortuño A, San José G, Moreno MU, Díez J & Zalba G (2005). Oxidative stress and vascular remodeling. Exp Physiol 90 (4): 457-66.

Fridovich I (1970). Quantitative aspects of the production of superoxide anion radical by milk xanthine oxidase. J Biol Chem 245:4053-4057.

Frieden TR, Fujiwara PI, Washko RM & Hamburg MA (1995). Tuberculosis in New York City-turning the tide. N Engl J Med 333:229-233.

Gebick JM & Bielski BHJ (1981). Comparison of the capacities of the dihydroxyl and superoxide radicals to initiate chain oxide of linoleic acid. J Am Chem Soc 103:7020-7022.

Gil L, Martinez G., Gonzalez I, Tarinas A, Alvarez A, Giuliani A, Molina R, Tapanes R, Perez J & Leon OS (2003). Contribution of characterization of oxidative stress in HIV/AIDS patients. Pharmacol Res 47 (3): 217 – 24.

Grossman Z & Heberman RB (1997). T-cell homeostasis in HIV infection is neither failing nor blind: Modified cell count reflects an adaptive response of the host. Nat Med 3: 486-490.

Haase AT (1999). Population biology of HIV-infection: viral load and CD4+T-cell demographics and dynamics in lymphatic tissues. Ann Rev Immunol 17: 625-656.

Haber F & Weiss J (1994). The catalylized decomposition of hydrogen peroxide by iron salts. Proc Royal Soc A 147-332.

Hamilton CA, Miller WH, Al-Benna S, Brosnan MJ, Drummond RD, McBride MW & Dominiczak AF (2004). Strategies to reduce oxidative stress in cardiovascular disease. Clin Sci 106 (3): 219-34.

Hayes KC, Pronczuk A & Khosla P (1995). A rationale for plasma cholesterol modulation by dietary fatty acids: Modelling the human response in animals. J Nutr Biochem 6:188-194.

Hayes KC (1993). Specific dietary fat without reducing saturated fatty acids does not significantly lower plasma cholesterol concentrations in normal males. Am J Clin Nutr 55:675-681.

Hayes, KC, Pronczuk A, Lindsey S & Diersen-Scchade D (1991). Dietary saturated fatty acids (12:0,14:0,16:0) differ in their impact on plasma cholesterol and lipoproteins in human primates. Am J Clin Nutr 53:491-498.

Heistad DD, Wakisaka Y, Miller J, Chu Y & Pena-Silva R (2009). Novel aspects of oxidative stress in cardiovascular diseases. Cir J 73 (2): 201-07.

Higashi Y, Jitsuiki D, Chayama K & Yoshizumi M (2006). Edaravone (3-methyl-1-phenyl-2-pyrazolin-5-one), a novel free radical scavenger, for treatment of cardiovascular diseases. Recent Pat Cardio Drug Discov 1 (1):85-93.

Holub BJ, Silicilia F & Mahadevappa VG (1989). Effect of tocotrienol derivatives on collagen and ADP-induced human platelet aggregation. Presented at PORIM. Int Palm Oil Dev Conf, Sept 5-6, Kuala Lumpur.

Hornstra G (1988). Dietary lipids and cardiovascular disease: Effects of palm oil. Oleagineux 43:75-8.1

Jagirdar J & Zagzag D (1996). Pathology and insights into pathogenesis of TB. pp. 467-482 in Tuberculosis 1st Ed. Little Brown & Company. New York N. Y.

Januga P, Jaruga B, Gackowski D, Olczak A, Halota W, Pawlowska G & Olinski R (2002). Supplementation with antioxidant vitamins prevent oxidative modification of DNA in lymphocytes of HIV infected patients. Free Rad Biol Med 32: 414-420.

Khosla P & Hayes KC (1992). Comparison between effects of dietary saturated (16:0), monounsaturated (18:1) and polyunsaturated (18:2) fatty acids on plasma lipoprotein metabolism in Cebus and Rhesus monkeys fed cholesterol-free diets. Am J Clin Nutr 55:51-62.

Khosla P & Hayes KC (1994). Cholesterolaemic effects of the saturated fatty acids of palm oil. Food Nutr Bull 15:119-125.

Kiedzierska K & Crowe SM (2001). Cytokines and HIV: Interaction and clinical implications. Antivir Chem Chemother 12: 133 – 150.

Kiremidjian-Schumacher L, Roy M, Wishe HI, Cohen MW & Stotzky G (1994). Supplementation with selenium and human immune cell functions: Effect on cytotoxic lymphocytes and natural killer cells. Biol Trace Elem Res. 41 (1-2):115-127.

Koch R (1882). Die aetiology der Tuberculese. Klinische Wochenschr. 19:221-230.

Koot M, Van't B, Wout A & Koot NA (1996). Relation between changes in cellular load evaluation of viral phenotype and the clonal composition of virus type infection. J Infect Dis 173:349-354.

Kooyenga DK, Gerler M, Watkins TR, Gapor A, Diakoumakis E & Bierenbaum ML (1997). Palm oil antioxidant effects in patients with hyperlipidaemia and carotid stenosis-2 year experience. Asia Pac J Clin Nutr.6:1:72-75.

Krichevsky D (1988). Effects of triglyceride structure and lipid metabolism. J Nutr Rev 46:177.

Kritchevsky D (2000). Impact of red palm oil on human nutrition and health. Food & Nutr Bull, 21: 182-88.

Kritchevsky D (1996)). Influence of triglyceride structure on experimental atherosclerosis in rabbits. FASEB J 10:A187.

Kurfeld D, Davidson LM & Lopex-Guisa JM (1990). Palm and other edible oils: atherosclerosis study in rabbits. FASEB 4: 23-26.

Lindquist Y, Branden CL, Mathews FS & Lederer F (1991). Spinach glycolate oxidase and yeast flagrant cytochrome are structurally homologous and evolutionary related enzymes with distinctly different function and flavin mononucleotive binding. J Biol Chem 266:3198-3207.

Loschen G, Azzi S, Richter C & Flohebp L (1974). Superoxide radicals as precursors of mitochondrial hydrogen peroxide. FASEB 42: 68-72.

Madebo T, Lindtjфrn B, Aukrust P & Berge R (2003). Circulating anti-oxidants and lipid peroxidation products in untreated tuberculosis patients in Ethiopia. Am J Clin Nutr 78: 117 – 22.

Marzuki A, Arshad F, Tariq AR & Kamsiah J (1991). Influence of dietary fat on plasma lipid profiles of Malaysian adolescents. Am J Clin Nutr 53:1010S-1014S.

Mathema B, Kurepina NE, Bifani PJ & Kreiswirth BN (2006). Molecular epidemiology of tuberculosis: Current insight. Clin Microbiol Rev 19: 658-685.

Mattson FH & Grundy SM (1985). Comparison of effects of dietary saturated, monounsaturated and polyunsaturated fatty acids on plasma lipids and lipoproteins in man. J Lipid Res 26:194-202.

Maxwell SR & Lip GY (1997). Free radicals and antioxidants in cardiovascular disease. Brit J Clin Pharmacol 44 (4): 307-17.

Mclntosch GH, Whyte J. McAthur R & Nestle PJ (1991). Barley and wheat foods; Influence on plasma cholesterol concentrations in hypercholesterolaemic men. Am J Clin Nutr 1205-1209.

Medlar EM (1955) Necropsy studies of human pulmonary TB. Ann Rev Tuberc 71 (2): 29 – 55.

Murray RK (2000). Muscle and the cytoskeleton. In Harper's Biochemistry (Murray, R. K., Granner, D. K.,Mayes, P. A., and Rodwell, V. W. Eds., 25th ed.), pp. 715-736. New York, McGraw-Hill Press.

Ng TKW, Hayes KC, de Witt GF, Jegathesan M, Satgunasingham N, Ong, ASH & Tan DTS (1992). Palmitic and oleic acids exert similar effects on serum lipid profile in the normo-cholesterolaemic humans. J Am Coll Nutr 11:383-390.

Ng TKW, Hassan K, Lim JB, Lye, MS & Ishak R (1991). Non-hypercholesterolaemic effects of a palm oil diet in Malasysian volunteers. Am J Clin Nutr 53:1015S-1020S.

O'Brien RJ (2001). Tuberculosis: scientific blueprint for TB drug development.Global Alliance for TB. Drug Dev, New York, N.Y.

Oguntibeju OO, van den Heever WMJ & Van Schalkwyk FE (2008). Potential effects of nutrient supplement on the anthropometric profiles of HIV-positive patients: complementary medicine could have a role in the management of HIV/AIDS. Afr J Biomed Res 11: 13-22.

Oguntibeju OO, van der Heever WMJ & Van Schalkwyk FE (2007a). Interplay between socio-demographic variables, nutritional and immune status of HIV-positive/AIDS patients. Pak J Biol Sci 10 (20): 3592 – 98.

Oguntibeju OO, van der Heever WMJ & Van Schalkwyk FE (2007b). A review of the epidemiology, biology and pathogenesis of HIV. J Biol Sci 7(8):1296-1304.

Ong ASG & Goh SH (2002). A healthful and cost-effective dietary component. United Nat Food & Nutr Bull 23 (1): 11-19.

Paxon WA, Martin SR & Tse D (1996). Relative resistance to HIV-1 infection of CD4 lymphocytes from persons who remain uninfected despite multiple sexual exposures. Nat Med 2: 412-417.

Qureshi A, Qureshi N & Hasler-Rapaczl (1991a). Dietary tocotrienol reduces concentrations of plasma cholesterol, apoprotein B, thromboxane B2 and platelet factor 4 in pigs with inherited hyper-lipidemias. Am J Clin Nutr 53:1042S-1046S.

Qureshi A, Burger WC, Prentice N, Bird HR & Sunde ML (1980). Regulation of Lipid metabolism in chicken liver by dietary cereals. J Nutr 10:388-393.

Qureshi A, Qureshi N & Wright JJK (1991b): Lowering of serum cholesterol in hypercholesterolemic humans by tocotrienols (palmvitee). Am J Clin Nutr 53:1021S-1026S.

Reddy YN, Murthy SV, Krishna DR & Prabhakar S (2004). Role of radicals and antioxidants in tuberculosis patients. Ind J Tuberc 151: 213-18.

Roach DR., Bean AG, Demangel C, Frances MP, Briscoe H & Brilton WJ (2002). TNF regulates chemokine induction essential for cell recruitment, granuloma formation and clearance of mycobacterium infection. J Immunol 168: 4620 – 27.

Sebinova E, Kagan V, Han D & Parker L (1991). Free radical recycling and inter-membrance mobility in the antioxidant properties of alpha– tocopherol and beta– tocotrienol. Free Rad Biol Med 10: 263 – 275.

Semba RD & Tang L (1999). Micronutrients and the pathogenesis of human immunodeficiency virus infection. Br J Nutr 81: 181-189.

Stephen MS (2006). The pathogenesis of HIV infection: the stupid may not be so dumb afterall. Retrov 3: 60.

Sundram K, Ismail A, Hayes KC, Jeyamalar R & Pathmanathan R (1997). Trans-celeidial fatty acids adversely affect the lipoprotein profile relative to specific saturated fatty acids in humans. J Nutr 127:3:5143-5203.

Sundram K, Sambanthamurthi R, Tan YA (2003). Palm Fruit Chemistry. Asia Pac J Clin Nutr 12 (3):` 355 – 362.

Sundram K, Hornstra G & Houwelingen AC (1992). Replacement of dietry fat with palm oil: Effect on human serum lipid, lipropotein and apoprotein. Brit J Nutr 68:677-692.

Sundram K, Khor HT & Ong ASH (1990). Effect of dietary palm oil and its fractions on rat plasma LDL and high density lipoprotein. Lipids 25(4):187-193.

Truswell AS, Choudhury N & Roberts DCK (1992). Double-blind comparison of plasma lipids in healthy subjects eating potato crisps fried in palm olein or canola oil. Nutr Res 12: S34-S52.

Valko M, Leibfritz D, Moncol J, Cronin MT, Mazur M. & Telser J (2007). Free radicals and antioxidants in normal physiological functions and human disease. Int J Biochem Cell Biol 39 (1): 44-84.

Van Rooyen J, Esterhuyse AJ, Engelbrecht A & du Toit EF (2008). Health benefits of natural carotenoid rich oil: a proposed mechanism of protection against ischaemia reperfusion injury. Asia Pac J Clin Nutr 17: 316 – 19.

Waterfall AH, Singh G, Fry JR & Marsden CA (1997). The measurement of lipid peroxidation *in vivo*. Brain Res Protocols 2: 217–22.

Weiss R (1996). HIV receptors and the pathogenesis of AIDS. Sci., 272: 1885-1886.

Zalba G, Fortuño A & Díez J (2006). Oxidative stress and atherosclerosis in early chronic kidney disease. Nephrol Dialy Transplant 21 (10): 2686-90.

Zhang J, Wang CR, Dai J, Chen XS & Ge KY (1997a): Palm oil diet may benefit mildly hyper-cholesterolaemic Chinese adults. Asia Pac J Clin Nutr 6:1:22-25.

Zhang J, Wang P, Wang CR Chen XS & Ge KY (1997b): Non-hypercholesterolaemic effects of a palm oil diet in Chinese adults. J Nutr 127(3): 5095-5135.

Zhang J, Wang P, Wang P, Wang CR, Chen XS & Ge KY (1995). The effects of palm oil on serum lipids and thrombosis in Chinese adults. Paper presented at the 7th Asian Congr Nutr, /Beijing, China Oct. 7-11.

7

Medical Plant and Human Health

Ahmed Morsy Ahmed
Faculty of Agriculture,
Ain shams University,
Egypt

1. Introduction

1.1 Primary healthcare and traditional medicine

Primary Health Care (PHC) is the key to the development of a national health policy based on practical, scientifically sound and socially acceptable methods and technology made universally acceptable to individuals and families in the community and through their full participation and at a cost that the community and the country can afford, in order to maintain, at every stage of their development, in the spirit of self-reliance and self determination. It is the first level of contact for the individual, family and the community within the national health care system, bringing health care as close as possible to where people live and work and thus constitutes the first element of a continuing health care process (WHO, 1978a). A health system, based on primary health care was adopted as the means of achieving the goal of health for all. Most developing countries of the world, for which the scheme was designed, have failed to seriously implement it up till this moment.

Traditional Medicine is defined by the World Health Organization (WHO, 1978a) as the sum total of knowledge or practices whether explicable or inexplicable, used in diagnosing, preventing or eliminating a physical, mental or social disease which may rely exclusively on past experience or observations handed down from generation to generation, verbally or in writing. It also comprises therapeutic practices that have been in existence often for hundreds of years before the development of modern scientific medicine and are still in use today without any documented evidence of adverse effects. The explicable form of Traditional Medicine can be described as the simplified, scientific and the direct application of plant, animal or mineral materials for healing purposes and which can be investigated, rationalized and explained scientifically.

1.2 Herbal medicine

The use of *Salix alba*, the willow plant (containing the salicylates) for fever and pains which led to the discovery of aspirin, would belong to this form of Traditional Medicine. Herbal medicines, which squarely belong to this form, are regarded by the World Health Organization, as finished and labeled medicinal products that contain, as active ingredients, aerial or underground parts of identified and proven plant materials, or combination thereof, whether in crude form or as plant preparations. They also include plant juices, gums, fatty oils, essential oils etc (WHO, 1978a). There are several other official modern drugs today, which were originally developed like aspirin through traditional medicine e.g.

morphine, digoxin, quinine, ergometrine, reserpine, atropine, etc and all of which are currently being used by orthodox medicine in modern hospitals all over the world.

The inexplicable form of Traditional Medicine on the other hand, is the spiritual, supernatural, magical, occultic, mystical, or metaphysical form that cannot be easily investigated, rationalized or explained scientifically e.g. the use of incantations for healing purposes or oracular consultation in diagnosis and treatment of diseases. The explanation is beyond the ordinary scientific human intelligence or intellectual comprehension. Plants are reputed in the indigenous systems of medicine for the treatment of various diseases (Arise *et al.*, 2009). Phyto-chemicals isolated from plant sources are used for the prevention and treatment of several medical problems including diabetes mellitus (Waltner- Law *et al.*, 2002). There are more than 800 plant species showing a hypoglycemic activity. The World Health Organization (1980) has also recommended the evaluation of the effectiveness of plants in conditions where safe modern drugs are lacked.

2. Major groups of antimicrobial compounds from plants

Plants have an almost limitless ability to synthesize aromatic substances, most of which are phenols or their oxygen-substituted derivatives (Geissman 1963). Most are secondary metaboltes, of which at least 12,000 have been isolated, a number estimated to be less than 10% of the total (Schultes 178). In many cases, these substances serve as plant defense mechanisms against predation by microorganisms, insects, and herbivores. Some, such as terpenoids, give plants their odors; others (quinones and tannins) are responsible for plant pigment. Many compounds are responsible for plant flavor (e.g., the terpenoid capsaicin from chili peppers), and some of the same herbs and spices used by humans to season food yield useful medicinal compounds.

Simple phenols and phenolic acids, The mechanisms thought to be responsible for phenolic toxicity to microorganisms include enzyme inhibition by the oxidized compounds, possibly through reaction with sulfhydryl groups or through more nonspecific interactions with the proteins (Mason and Wasserman 1987).Phenolic compounds possessing a C_3 side chain at a lower level of oxidation and containing no oxygen are classified as essential oils and often cited as antimicrobial as well. Eugenol is a well-characterized representative found in clove oil. Eugenol is considered bacteriostatic against both fungi (Duke 1985) and bacteria (Thomson 1978). Quinones are aromatic rings with two ketone substitutions. They are ubiquitous in nature and are characteristically highly reactive. Probable targets in the microbial cell are surface-exposed adhesins, cell wall polypeptides, and membrane-bound enzymes. Quinones may also render substrates unavailable to the microorganism. As with all plant-derived antimicrobials, the possible toxic effects of quinones must be thoroughly examined.Kazmi et al.

Flavones, flavonoids, and flavonols activity is probably due to their ability to complex with extracellular and soluble proteins and to complex with bacterial cell walls, as described above for quinones. More lipophilic flavonoids may also disrupt microbial membranes (Tsuchiya 1996). Tannins are found in almost every plant part: bark, wood, leaves, fruits, and roots (Scalbert 1991). This group of compounds has received a great deal of attention in recent years, since it was suggested that the consumption of tannin-containing beverages, especially green teas and red wines, can cure or prevent a variety of ills (Serafini et al 1994) their mode of antimicrobial action, may be related to their ability to inactivate microbial adhesins, enzymes, cell envelope transport proteins, etc. They also complex with polysaccharide (Ya et al 1988). The antimicrobial significance of this particular activity has not been explored. There is also evidence for direct inactivation of microorganisms: low tannin concentrations modify the morphology of germ tubes of *Crinipellis perniciosa* (Brownlee et al 1990). Tannins in plants inhibit insect growth (Schultz 1988) and disrupt digestive events in ruminal animals (Butler 1988).

Terpenoids and Essential Oils are secondary metabolites that are highly enriched in compounds based on an isoprene structure. They are called terpenes, their general chemical structure is $C_{10}H_{16}$, and they occur as diterpenes, triterpenes, and tetraterpenes (C_{20}, C_{30}, and C_{40}), as well as hemiterpenes (C_5) and sesquiterpenes (C_{15}). When the compounds contain additional elements, usually oxygen, they are termed terpenoids. The mechanism of action of terpenes is not fully understood but is speculated to involve membrane disruption by the lipophilic compounds. Accordingly, Mendoza et al. (Mendoza et al 1997) found that increasing the hydrophilicity of kaurene diterpenoids by addition of a methyl group drastically reduced their antimicrobial activity.

Alkaloids are Heterocyclic nitrogen compounds. Alkaloids have been found to have microbiocidal effects (including against *Giardia* and *Entamoeba* species Ghosha et al 1996, the major antidiarrheal effect is probably due to their effects on transit time in the small intestine. Berberine is an important representative of the alkaloid group. It is potentially effective against trypanosomes (Freiburghaus et al 1996) and plasmodia Omulokoli et al 1997). The mechanism of action of highly aromatic planar quaternary alkaloids such as berberine and harmane (Hopp 1976) is attributed to their ability to intercalate with DNA (Phillipson and O'Neill 1987). The mechanism of lectins and polypeptides may be back to the formation of ion channels in the microbial membrane (Terras et al. 1993) or competitive inhibition of adhesion of microbial proteins to host polysaccharide receptors (Sharon and Ofek 1986). Recent interest has been focused mostly on studying anti-HIV peptides and lectins, but the inhibition of bacteria and fungi by these macromolecules, such as that from the herbaceous *Amaranthus*, has long been known (De Bolle et al 1996).

2.1 Ethnobotanical study

Out of around 1076 species recorded so far from Semifinal Biosphere Reserve, more than 200 species are attributed with medicinal uses (Rout, 2004). This system of using herbs and different biological active ingredients in treating various diseases had become a part of their culture till recent years. Entrance of market economy gave rise to exploitation of natural resources and thereby depleting our resources base. The most affected part in this process was medicinal plants, which is most sensitive and delicate in the environment of forest. These medicinal plants gain further importance in the region where modern medical health facilities are either not available or not easily accessible. Although our ancient sages through hit and trial method developed herbal medicines, the reported uses of plant species do not certify efficacy.

The present preliminary report on the uses of some plant species need to pharmacologically screened, chemically analyzed and tested for bioactive activities. Pharmacological screening of plant extracts provides insight to both their therapeutic and toxic properties and helps in eliminating the medicinal plants or practices that may be harmful. The plant parts used for medical preparation were bark, flowers, rhizomes, roots, leaves, seeds, gum and whole plants. In some cases the whole plant including roots was utilized. The use of plant derived products containing high of dietary fiber and complex polysaccharide for the management of diabetes have been proposed (Jenkins et al., 1976). Natural products especially of plant origin have been found to be potential sources of novel molecules for the treatment of diabetes (Farnsworth, 1994; Marles and Farnsworth, 1995).

Considering the rate at which the vegetation is getting depleted in this part of the world, therefore it is needed to document the precious knowledge of these plants and to search for more plants with antidiabetic potential. The search for anti-diabetic agents has been focused on plants because of their ready availability, effectiveness, affordability and probably due to low side effects (Marles and Farnsworth, 1995). Ethnobotanical study has been the method often used to search for locally important plant species with low side effects especially for the discovery of crude drugs (Farnsworth, 1994). The present study therefore is a documentation of plants and plant parts used for the management of diabetes mellitus by traditional healers of the area (Table 1).

Diagnosis methods	**Respondent (%)**
Signs and symptoms	
Loss of body weight	100
Body weakness	100
Excessive urination	100
Presence of sugar in urine	100
Excessive thirsty	53
Duration of treatment	
Short duration	40
Long duration	60
Efficacy of plant treatment on patients	
Disappearance of sugar in urine 100	100
Reduction in body weakness 100	100
Normal body weight 100	100
Reduction in frequency of urination	100
Traditional healers claim of no adverse effect after treatment	
Yes	100
No	0
Traditional healers claim of total cure after treatment	
Yes	80
No	20

Table 1. Diagnosis methods of diabetes mellitus by the herbalists using herbal preparation

2.2 Diabetes mellitus

2.2.1 What is Diabetes?

The word "diabetes" (a Greek word that means "to pass through") was first used by Aretaeous of Capadocia in the 2nd century AD to describe a condition that is characterized by excess of sugar in blood and urine, hunger and thirst (MacFrlance et al.,1997) and the adjective "mellitus" (a latin-greek word that means "honey") was introduced by the English physician John Rollo so as to distinguish the conditions from other polyuric diseases, in which glycosuria does not occur (Rollo 1797). People suffering from diabetes are not able to produce or properly use insulin in the body and therefore chronic hyperglycemia occurs. In addition, the diabetic individual is prone to late onset complications (Fujisawa et al., 2004), such as retinopathy, neuropathy and vascular diseases, that are largely responsible for the morbidity and mortality observed in diabetic patients.

Diabetes mellitus is a chronic metabolic disorder characterized by widespread complications. It is the world's largest endocrine disease associated with increased morbidity and mortality rates (Ghosh and Surawanshi, 2001). The chronic hyperglycemia of diabetes is associated with long-term damage, dysfunction and failure of various organs (Lyra *et al.*, 2006). Liver involvement is one of the leading causes of death in diabetes mellitus. The mortality rate from the hepatic affection is greater than that from the cardiovascular complications. The spectrum of liver implication in diabetes ranges from non-alcoholic fatty liver disease to cirrhosis and eventually hepato-cellular carcinoma (Keith *et al.*, 2004). Liver, an insulin-dependent organ, plays a pivotal role in glucose and lipid homeostasis. It participates in the uptake, oxidation and metabolic conversion of free fatty acids and in the synthesis of cholesterol, phospholipids and triglycerides.

Several mechanisms are implicated in the pathogenesis of the functional and morphological alterations of the liver of diabetic patients (Brixova, 1981; Moller, 2001). There are two main types of diabetes, namely type I and type II (World Health Organization. Definition, Diagnosis and Classification of Diabetes mellitus and its omplications. Part 1: Diagnosis and Classification of Diabetes Mellitus (Department of Non communicable Disease Surveillance, Geneva, 1999)). Type I Diabetes, that is called insulin-dependent diabetes mellitus (IDDM) or juvenile onset diabetes develops when the body's immune system destroys pancreatic β-ells, the only cells in the body that produce the hormone insulin that regulates blood glucose. This type of diabetes usually strikes children and adults and the need for insulin administration is determinant for survival. Type I diabetes accounts for 5% to 10% of all diagnosed cases of diabetes and the risk factors may be autoimmune, genetic, or environmental.

On the other hand, type II diabetes, also called non–insulin-dependent diabetes mellitus (NIDDM) or adult-onset diabetes, accounts for about 90% to 95% of all diagnosed cases of diabetes. It usually begins as insulin resistance, a disorder in which the cells do not use insulin properly and as the need for insulin rises; the pancreas gradually loses its ability to produce it. This type of diabetes is associated with older age, obesity, family history of diabetes, history of gestational diabetes, impaired glucose metabolism, physical inactivity, and race/ethnicity. It must be noted thought that in the last decay type II diabetes in children and adolescents is being diagnosed more frequently (Fagot-Gampagna & Narayan 2001). In the case of the IDDM, insulin is of crucial importance for the survival of the patients. On the other hand, in the case of NIDDM the treatment includes medicines, diets and physical training.

2.2.2 Ant diabetic medicines: Important medicinal plants used in diabetes treatment; important active

2.2.2.1 Compounds involved in diabetes mellitus treatment

Many kinds of antidiabetic medicines have been developed for the patients and most of them are chemical or biochemical agents aiming at controlling or/and lowering blood glucose to a normal level. Despite the impressive advances in health sciences and medical care, there are many patients who are using alternative therapies alone or complementary to the prescribed medication. Traditional plant remedies or herbal formulations exist from ancient times and are still widely used, despite all the controversy concerning their efficacy and safety (Huxtable 1990; Fugh-Berman 2000), to treat hypoglycemic and hyperglycemic conditions all over the world.

It must be noted that many ethno-botanical surveys on medicinal plants used by the local population have been performed in different parts of the world and there is a considerable number of plants described as antidiabetic. In addition a variety of compounds have been isolated (alkaloids, glycosides, terpenes, flavonoids, etc) but further studies need to be done so as these 'leads' to develop into clinically useful medicines. To date, met form in (a biguanide) is the only drug approved for treatment of type II diabetes mellitus. It is a derivative of an active natural product, galegine, isolated from the plant Galega officinalis L. (Witters 2001). The followinf table illustrate some of these plant involved in diabetes mellitus treatment

Scientific name	Family	Active constituent	Part Used	Folk Medical Uses
Abelmoschus moschatus	**Malvaceae**	myricetin	aerial part	hypoglycemic action decrease the plasma glucose concentrations myricetin has an ability to enhance glucose utilization to lower plasma glucose in diabetic rats with deficient insulin levels.
Azadirachta indica	**Meliaceae**		dried powder of root and leaves	caused significant lowering of blood sugar and reduction in serum lipids
Cornus macrophylla	**Cornaceae**			prevent and treat diabetic complications
Achyranthes aspera L.	**Amaranthaceae**		powdered whole plant	hypoglycemic effect
Achyrocline satureioides	**Asteraceae**	achyrofuran	powdered whole plant	lowered blood glucose levels
Acosmium panamense	**Fabaceae**		water and butanolic extracts	lowered the plasma glucose levels
Aegle marmelose	**Rutaceae**		leaf extract	similar hypoglycemic activity to that of insulin treatment Treatment with the leaf extract showed improved functional state of pancreatic β cells. The results indicate the potential hypoglycemic effect of the leaf extract, possibly involved in processes for the regeneration of damaged pancreas.

Table 2. Important medicinal Plants used in Diabetes treatment.

Scientific name	Family	Active constituent	Part Used	Folk Medical Uses
Allium sativum L	**(Liliaceae**	allicine		decreased the concentration of serum lipids, blood glucose and activities of serum enzymes like alkaline phosphatase, acid phosphatase and lactate dehydrogenase and liver G6Pase
Allium cepa L.	**(Liliaceae**	S-methyl cysteine sulfoxide		lowered the levels of malondialdehyde, hydroperoxide and conjugated dienes in tissues exhibiting antioxidant effect
		(SMCS)		on lipid peroxidation in experimental diabetes The hypoglycemic and hypolipidaemic actions of Allium cepa were associated with antioxidant activity, since onion decreased SOD activities while no increased lipid hydroperoxide and lipoperoxide concentrations were observed in diabetic rats treated with Allium cepa
Aloe vera	**(Liliaceae**		leaf pulp and gel extracts	contain a hypoglycemic agent which lowers the blood glucose levels showed hypoglycemic activity on type I and II diabetic maintained the glucose homeostasis by controlling the carbohydrate metabolizing enzymes (Rajasekaran et al., 2004).
Aloe barbadensis	**Liliaceae**		leaves	the bitter principle may be mediated through stimulating synthesis and/or release of insulin from the β cells of Langerhans
Andrographis paniculata	**Acanthaceae**	andrographolide	crude ethanolic extract	antidiabetic effect may be attributed, at least in part, to increased glucose metabolism andrographolide can increase the glucose utilization to lower plasma glucose in diabetic rats lacking insulin.

Table 2. Important medicinal Plants used in Diabetes treatment. (continuation)

Scientific name	Family	Active constituent	Part Used	Folk Medical Uses
Angylocalyx pynaertii	**Leguminosae**	sugar-mimic alkaloids	pod extract	specific inhibitors of alpha-L fucosidase with no significant inhibitory activity towards other glycosidases
Areca catechu		arecolin		Arecoline have hypoglycemic activity in animal model of diabetes upon s.c. administration
Averrhoa bilimbi L.	**Oxalidaceae**		ethanolic extract	ethanolic extract has hypoglycemic, hypotriglyceridemic, anti-lipid peroxidative and anti-atherogenic properties in STZ-diabetic rats.
Bauhinia forficata L.	**Leguminosae**	kaempfer itr in	n-butanol extract	blood glucose-lowering effect in normal and diabetic rats
Beta vulgaris L. var. cicla	**Chenopodiaceae**	Betavulgaroside I, II, III and IV	roots	hypoglycemic effects that was demonstrated by a per os glucose tolerance test in rats after their per os administration
Bidens pilosa	**Asteraceae**	3-beta-D-glucopyranosyloxy-1-hydroxy-6(E)-tetradecene-8,10,12-triyne	Butanol extract of whole plant	inhibited the differentiation of naive helper T cells (Th0) into Th1 cells but enhanced their transition into type II helper T (Th2) cells using an in vitro T cell differentiation assay
Bryonia alba L.	**Curcubiaceae**	derivatives of trihydroxyoctadecadienoic acid	roots	Restore the disordered lipid metabolism of alloxan-diabetic rats
Bumelia sartorum	**Sapotaceae**	bassic acid	ethanol extract of root bark	altered glucose tolerance in alloxan-induced diabetic rats, enhanced glucose uptake in skeletal muscle and significantly inhibited glycogenolysis in the liver.
Caesalpinia ferrea Mart	**Leguminosae**	ellagic acid and 2′-(2,3,6-trihydroxy-4-carboxyphenyl) ellagic aci		might contribute to the relief of the long-term diabetic complications
Camellia sinensis	**Theaceae**	epigallocatechin-3-O-gallate		may be beneficial in the prevention of diabetes mellitus
Cassia tora	**Fabaceae**	a-tocopherol, ascorbic acid and maltodextrin	soluble fiber extracted	improve serum lipid levels in type II diabetic subjects without serious adverse effects

Table 2. Important medicinal Plants used in Diabetes treatment.(continuation)

Scientific name	Family	Active constituent	Part Used	Folk Medical Uses
Cecropia obtusifolia Bertol	**Cecropiaceae**	isoorientin and 3-caffeoylquinic acid	water and butanolic extracts prepared from leaves	beneficial effects on carbohydrate and lipid metabolisms when it was administered as an adjunct on patients with type II diabetes with poor response to conventional medical treatment.
Cleome droserifolia (Capparidaceae)				possessed a postprandial hypoglycemic effect but also suppressed the hepatic glucose release in the fasting state in a comparable way to this of insulin. The plant also possessed hypo-cholesterolemic effect, most pronounced on the LDL cholesterol.
Cnidium officinale Makino	**Apiaceae**			inhibited the high glucoseinduced proliferation of GMCs partially through TGF-beta1 production
Coccinia indica	**Cucurbitaceae**		ethanolic extract of Coccinia indica leaves	normalized blood glucose and caused marked improvement of altered carbohydrate metabolizing enzymes during diabetes.
Commelina communis	**Commeliaceae**	pyrrolidine alkaloid, 2,5-dihydroxymethyl-3,4-dihydroxypyrrolidine and four piperidine alkaloids, 1-deoxymannojirimycin, 1-deoxynojirimycin, alpha-homonojirimycin and 7-O-beta-D-glucopyranosyl alpha-homonojirimycin	methanolic extract	inhibitoryactivity against alpha-glucosidase

Table 2. Important medicinal Plants used in Diabetes treatment. (continuation)

Scientific name	Family	Active constituent	Part Used	Folk Medical Uses
Croton cajucara	**Euphorbiaceae**	trans-hydrocortin		The mentioned compound also effectively lowered the blood sugar levels in glucose fed normal rats.
Cryptolepis sanguinolenta	**Asclepiadaceae**	cryptolepine		antihyperglycemic effect of cryptolepine leaded to a significant decline in plasma insulin concentration, associated with evidence of an enhancement in insulin mediated glucose disposal. Finally, cryptolepine increased glucose uptake by 3T3-L1 cells.
Dioscorea dumetorum	**Dioscoreaceae**	dioscoretine	methanol extrac	normal and alloxan diabetic rabbits produced significant hypoglycemic effects and the hypoglycemic effects were compared to those of to butamide (Iwu et al., 1998)
Galega officinalis L.	**Leguminosae**	Metformin		Treat symptoms now ascribed to type II diabetes
Gentiana olivieri Griseb	**Gentianaceae**	isoorientin		Isoorientin exhibited significant hypoglycemic and antihyperlipidemic effects
Gymnema sylvestre	**Asclepiadaceae**	gymnemic acid I	Water soluble extracts the leaves	Gymnemic acids II and III showed potent inhibitory activity on glucose uptake.
Hintonia latiflora	**Rubiaceae**	coutaraegenin		coutareagenin, one of the active substances contained in the Hintonia latiflora bark, produces a reduction of the diabetic elevated blood sugar levels
Hydnocarpus wightiana	**Arcariaceae**	Hydnocarpin, Luteolin, And isohydnocarpin	acetone extract of the seed hulls	the presence of antioxidant molecules along with their enzyme inhibitory activities in the acetone extract of Hydnocarpus wightiana seed hulls may be responsible for the antidiabetic Properties

Table 2. Important medicinal Plants used in Diabetes treatment. (continuation)

Scientific name	Family	Active constituent	Part Used	Folk Medical Uses
Lagerstroemia speciosa L.	**Lythraceae**	Lagerstroemin and flosin B	hot-water extract from banaba leaves	effects on controlling of the level of plasma glucose in type II diabetes mellitus.
Larrea tridentata	**Zygophyllaceae**	masopropol		lower plasma glucose concentration in two mice models of type II diabetes after per os administration (Luo et al., 1998)
Myrcia multiflora	**Myrtaceae**	myrciacitrin I	ethyl acetate-soluble portion,	show inhibitory activities on ALR2 and alpha-glucosidase as well as on the increase of serum glucose level in sucrose-loaded rats and in alloxan-induced diabetic mice (Yoshikawa et al., 1998).
Paeonia lactiflora Pall.	**Panunculaceae**	paeoniflorin		the mentioned glucoside reduced the elevation of blood sugar in glucose challenged rats. Increase of glucose utilization by paeoniflorin can thus be considered (Hsu et al., 1997).
Pandanus odorus Ridl.	**Pantadaceae**	4-hydroxybenzoic acid		4-hydroxybenzoic acid caused a decrease in plasma glucose levels dose-dependently. The compound did not affect serum insulin level and liver glycogen content in the diabetic model, but increased glucose consumption in normal and diabetic rat diaphragms.
Phyllanthus sellowianus	**Euphorbiaceae**	isoqu ercitrin and rutin		hypoglycemic and diuretic agent reduction in blood glucose levels
Pueraria thunbergiana	**Leguminosae**	kaikasaponin III		kaikasaponin III (as a mixture of isomers), may exhibit its hypoglycemic and hypolipidemic effects by up-regulating or down-regulating antioxidant mechanisms via the changes in Phase I and II enzyme activities .
Stevia rebaudiana Bertoni	**Compositae**	stevioside		Stevioside was able to regulate blood glucose levels by enhancing not only insulin secretion, but also insulin utilization in insulin-deficient rats; the latter was due to decreased PEPCK gene expression in rat liver by stevioside's action of slowing down Gluconeogenesis (Chen et al., 2005).

Table 2. Important medicinal Plants used in Diabetes treatment. (continuation)

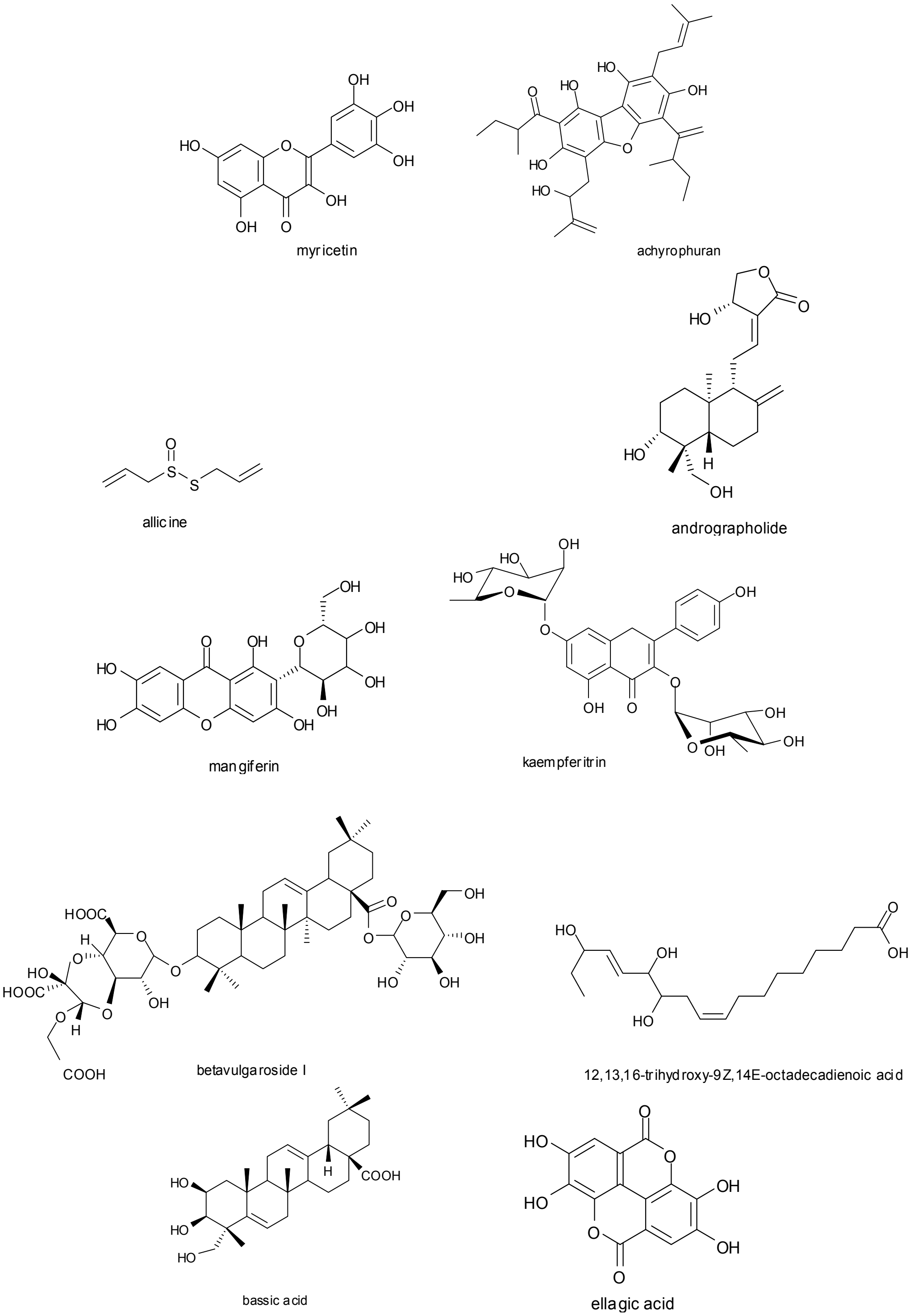
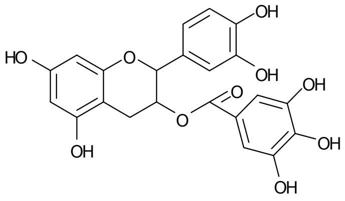
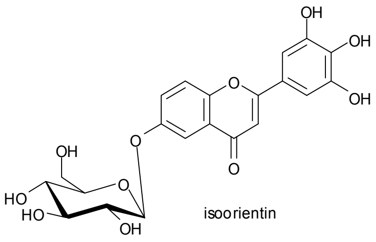

myricetin
achyrophuran
allicine
andrographolide
mangiferin
kaempferitrin
betavulgaroside I
12,13,16-trihydroxy-9Z,14E-octadecadienoic acid
bassic acid
ellagic acid

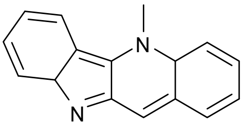
Fig. 1. Some important active compound involved in diabetes mellitus treatment.



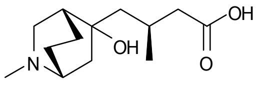
epigallocatechin-3-O-gallate



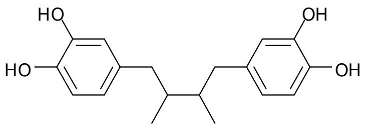
isoorientin



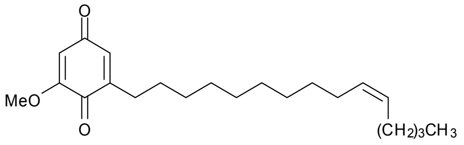
cryptolepine



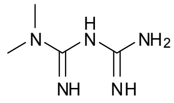
dioscoretine



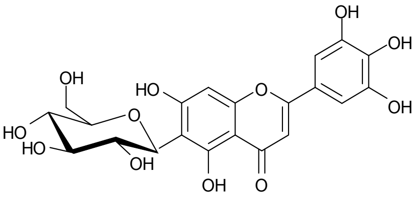
masopropol



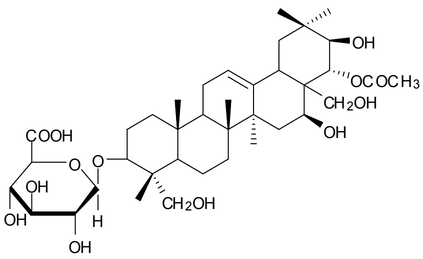
maesanin



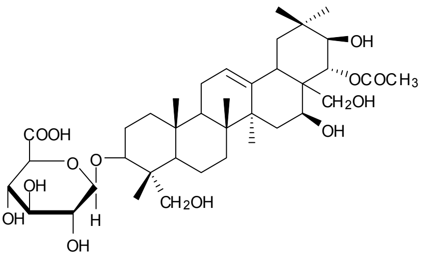
Metformin



isoorientin

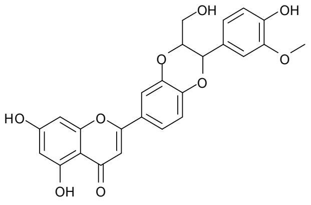


gymnemic acid I

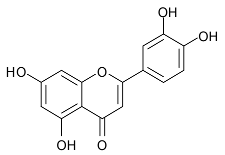


gymnemic acid II

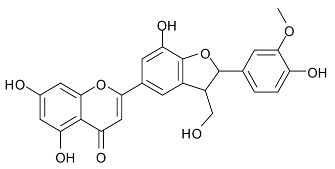
Fig. 1. Some important active compound involved in diabetes mellitus treatment. (continuation)



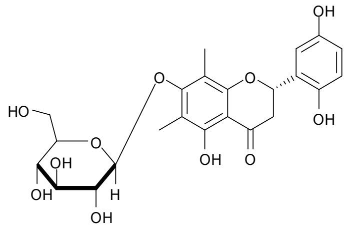
hydnocarpin



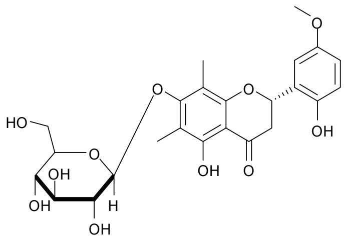
luteolin



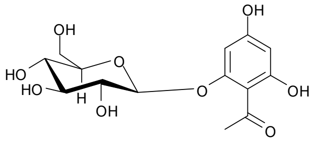
isohydnocarpin



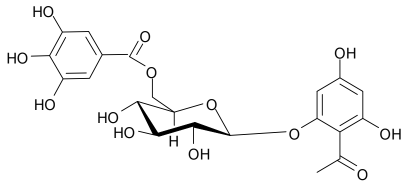
myrciacitrin I



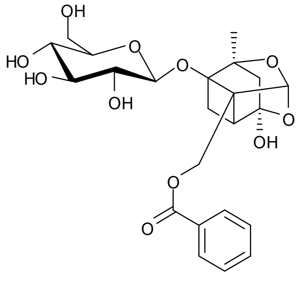
myrciacitrin II



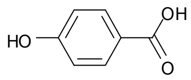
myrciaphenon A



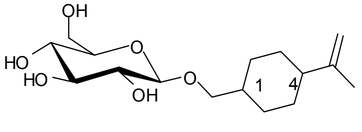
myrciaphenon B



paeoniflorin



4-hydroxybenzoic acid



perilloside C:1,4-trans
perilloside D:1,4-cis

Fig. 1. Some important active compound involved in diabetes mellitus treatment. (continuation)

3. Some biomedical application for some plant used in treating hypoglycemic and hyperglycemic conditions

A- Amelioration of hyperglycemia in STZ-diabetic rats.

Several approaches were conducted in order to investigate the hypoglycemic and hypolipidemic activities of plant extract on experimental animals (STZ-diabetic rats). The derived data showed that some extracts are potent in the amelioration of hyperglycemia in STZ-diabetic rats and is a potential source for the isolation of new per os active agent(s) for anti-diabetic therapy like what observed in Averrhoa bilimbi ethanolic extract which has hypoglycemic, hypotriglyceridemic, anti-lipid peroxidative and anti-atherogenic properties in STZ-diabetic rats (Purshparaj et al., 2001). Another way was by the inhibition of myeloperoxidase activity and decreased lipid peroxidation, induced by ascorbyl radical either in microsomes or in asolectin and phosphatidylcholine liposomes as what observed in Bauhinia forficate due to the kaempferitrin effect .that showed high reactivity with the free radical 1,1-diphenyl-2-picryl hydrazyl (DPPH),

In the plant *Beta vulgaris L. var. cicla (Chenopodiaceae*) the anti-diabetic effect in experimental animals may be to the reduction in blood glucose levels by the regeneration of the β cells (Bolkent et al., 2000). Betavulgarosides I, II, III, IV and oleanolic acid oligoglycosides produced hypoglycemic effects that was demonstrated by a per os glucose tolerance test in rats after their per os administration (Yoshikawa et al., 1996). Or through stimulating the synthesis/release of insulin from the β cells of Langerhans like the effect of *Biophytum sensitivum* and/or also mediated through enhance secretion of insulin from the β cells of Langerhans or through an Extra pancreatic mechanism (*Catharanthus roseus L. (Apocynaceae*)). Or by increasing the number of islet β cells and that of islet α cells was decreased in STZ-diabetic rats like what happen in Dendrobium candidum (Orchidaceae), Its mechanism of action probably involves the stimulation of the secretion of insulin from β cells and the inhibition of the secretion of glucagons from α cells and it can probably decrease the decomposition of liver glycogen and increase the synthesis of liver glycogen.

The Administration of *Coccinia indica (Cucurbitaceae*) leaf extract, an indigenous plant used in Ayurvedic medicine in India, to normal and STZ-diabetic animals exhibited significant hypoglycemic and antihyperglycemic effect and reversed the associated with diabetes biochemical alterations (Venkateswaran & Pari 2002). The results indicated that the per os administration of Coccinia indica leaf extract to diabetic animals normalized blood glucose and caused marked improvement of altered carbohydrate metabolizing enzymes during diabetes. The antioxidant effects of an ethanolic extract of Coccinia indica leaves was studied in STZ-diabetic rats (Venkateswaran & Pari 2003). Per os administration of Coccinia indica leaf extract resulted in a significant reduction in thiobarbituric acid reactive substances and hydroperoxides and a significant increase in reduced GSH, SOD, CAT, glutathione peroxidase and glutathione-S-transferase in liver and kidney of STZ-diabetic rats, which clearly showed the extract's antioxidant property. The ethnopharmacological use of *Gongronema latifolium* in ameliorating the oxidative stress found in diabetics and indicating promise of possible use in lessening morbidity in affected individuals.

The obtained result by Kim et al 2006 suggest that the administration of *Chrysanthemum coronarium* and *Morus alba,* have a hypoglycemic effect in diabetic rats and their effect was equivalent to that of glibenclamide. The administration of *C.unshiu* shows more antihyperlipidemic effect than antidiabetic effect. The effect of aqueous extracts from four medicinal plants on the blood glucose levels of experimental animals was determined at various time intervals for 9 h after oral administration at 100 mg dose kg-1 b.wt. (Fig. 2). There was a significant elevation in the blood glucose level by 3.3-5 times during experimental time period in alloxan-induced diabetic rats, when compared to normal rats. The administration of *C. coronarium* extract caused the blood glucose levels of diabetic rats to 83.4, 67.6, 75.1, 81.1 and 74.3% at the time interval of 1, 3, 5, 7 and 9 h, respectively ($p<0.05$). Maximum reduction of 32.4% was observed 3 h after treatment. The administration of *M. alba* extract produced the most significant reduction ($p<0.05$) among four medicinal plants in the blood glucose levels of 34, 41, 33 and 35% at 3, 5, 7 and 9 h respectively.

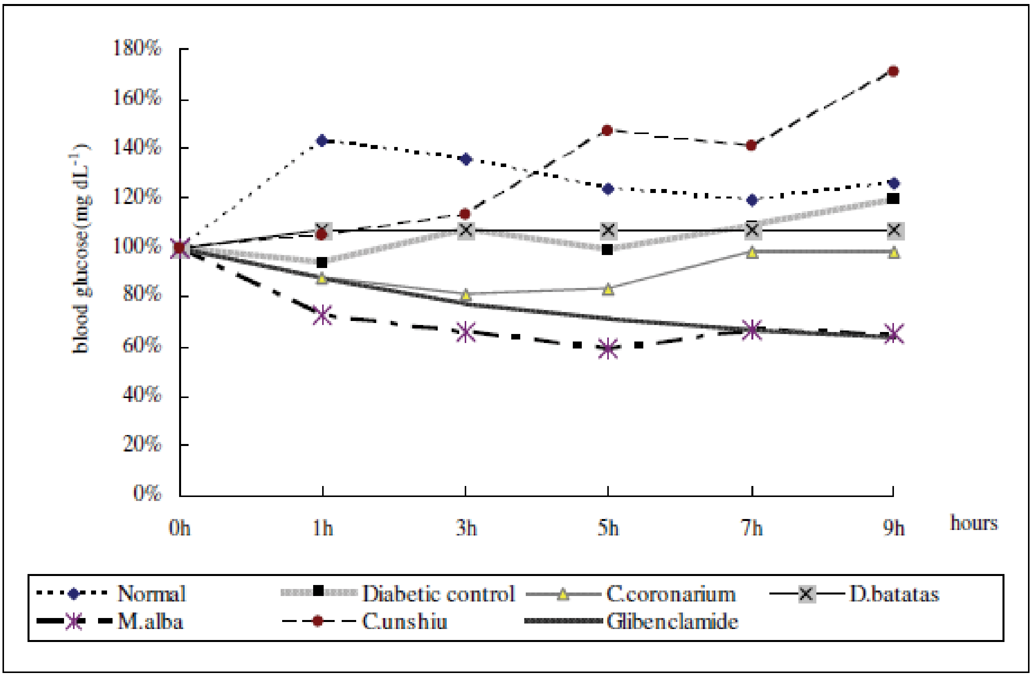


Fig. 2. Percentage of effect of experimental plants on plasma glucose levels compared with 0 h data in alloxaninduced diabetic rats (%). Values are mean percent of blood glucose concentration (n= 6). According to the study obtained by Kim et al., 2006.

B- Hypoglycemic, hypolipidemic hypotriglyceridemic, anti-lipid peroxidative and anti-atherogenic properties of these agents in different medicinal plant:

Insulin resistance (hyperinsulinemia) is now recognized as a major contributor to the development of glucose intolerance, dyslipidemia and hypertension in type II diabetic patients. Aqueous extracts of Pterocarpus marsupium bark (Fbaceae) and Trigonella foenum-graecum. (Leguminosae) seeds have been shown to exert hypoglycemic/antihyperglycemic effect in experimental as well as in clinical settings (Grover et al., 2005). Results of this study, in addition to previous clinical benefits of

Pterocarpus marsupium seen in type II diabetic subjects, are suggestive of usefulness of its bark in insulin resistance, the associated disorder of type II diabetes. Though several antidiabetic principles (epicatechin, pterosupin, marsupin and pterostilbene) have been identified.

The administration of trihydroxyoctadecadienoic acid and its derivatives obtained from the roots of the native Armenian plant *Bryonia alba (Curcubiaceae*) was found to restore the disordered lipid metabolism of alloxan-diabetic rats (Karageuzyan et al., 1998). These derivatives of trihydroxyoctadecadienoic acid can correct major metabolic abnormalities typical of severe diabetes mellitus they can influence the profile of the formation of stable prostaglandins by actions such as downstream of prostaglandin endoperoxides. In *Camellia sinensis (Theaceae*) species, the blood glucose lowering activity of Camellia sinensis was studied by many workers. It has recently been reported that the major green tea polyphenolic constituent, epigallocatechin 3- gallate, mimics the cellular effects of insulin including the reductive effect on the gene expression of rate-limiting gluconeogenic enzymes in a cell culture system (Koyama et al., 2004). Per os administration of green tea that contains this polyphenolic constituent caused a reduction in the levels of mRNAs for gluconeogenic enzymes, PEPCK and G6Pase in the mouse liver. Epigallocatechin 3-gallate alone was also found to down-regulate the gene expression of these enzymes in vitro.

Bumelia sartorum (Sapotaceae), the hypoglycemic effect may be similar to chlorpropamide and possibly due to an enhanced secretion of insulin from the islets of Langerhans or an increased utilization of glucose by peripheral tissues. Besides hypoglycemic activity, the ethanol extract also elicited significant anti-inflammatory activity, but did not show any significant effects on blood pressure, respiration or on the various isolated tissue preparations studied. Bassic acid, an unsaturated triterpene acid isolated from an ethanol extract of Bumelia sartorum root bark, elicited significant hypoglycemic activity in alloxan-diabetic rats and altered the pattern of glucose tolerance in these animals when administered per os (Naik et al., 1991). In addition, bassic acid treatment increased significantly the glucose uptake process and glycogen synthesis in isolated rat diaphragm. Bassic acid treatment increased plasma insulin levels significantly in alloxandiabetic rats. It was therefore suggested that the hypoglycemia activity of bassic acid may be mediated through enhanced secretion of insulin from the pancreatic β cells.

Compelling evidence of *Bidens pilosa (Asteraceae)* suggests that infiltrating CD4 type I helper T (Th1) cells in the pancreatic islets play a pivotal role in the progression of diabetes in non-obese diabetic mice. Treatment with a butanol fraction of Bidens blood glucose and insulin in non-obese diabetic mice in a dose-dependent manner and elevated the serum IgE levels regulated by Th2 cytokines (Chang et al., 2004). Moreover, the butanol fraction inhibited the differentiation of naive helper T cells (Th0) into Th1 cells but enhanced their transition into type II helper T (Th2) cells using an in vitro T cell differentiation assay. The butanol fraction of *Biderns pilosa* and its polyacetylenes can prevent diabetes possibly via suppressing the differentiation of Th0 cells into Th1 cells and promoting that of Th0 cells into Th2 cells.

The results of a reported study of *Gongronema latifolium (Asclepiadaceae)* suggest that the extracts from leaves could exert their antidiabetic activities through their antioxidant properties (Ugochukwu & Bababy 2002). The aqueous and ethanolic extracts were tested in order to evaluate their effect on renal oxidative stress and lipid peroxidation in non-diabetic and STZ-diabetic rats after per os administration (Ugochukwu & Makini 2003). The ethanolic extract appeared to be more effective in reducing oxidative stress, lipid peroxidation, and increasing the GSH/GSSG ratio, The antihyperglycemic effects of aqueous and ethanolic extracts from Gongronema latifolium leaves was also investigated on glucose and glycogen metabolism in liver of non-diabetic and STZ-diabetic rats (Ugochukwu & Bababy 2003). The data showed that the ethanolic extract from the plant's leaves had antihyperglycemic potency, which was suggested to be mediated through the activation of HK, PFK, G6PD and inhibition of GK in the liver.

According to the study occurred by Mohammad Khalil, et al. (2010), it is found found that, the treatment of the diabetic rats with *Citrullus colocynthis* pulp extract, in the livers showed, more or less, an improvement in the histological architecture with persistence of the cytoplasmic vacuoles in some hepatocytes that could be attributed to the residual adverse effect of the diabetic affliction. But the noticed apparent general improvement signifies that *Citrullus colocynthis* could possess cyto-protective abilities on the hepatocytes. The present findings are supported by those announced by Aburjai *et al*. (2007) who confirmed the anti-diabetic properties of *Citrullus colocynthis* extract. Also, Sebbagh *et al*. (2009) stated that this plant could improve the streptozotosin-induced diabetes in rats. Bujanda *et al*. (2008) owed the similar steatotic inhibitory effect of resveratol to inhibition of the tumor necrosis factor alpha production, lipid per-oxidation and oxidative stress.

The damaged sinusoids, the aemorrhage and the inflammatory cell infiltration encountered in the liver of the diabetic animals, might be due to the hyperglycemic state. Seifalian *et al*. (1999), in an analogous study in rabbits, stressed on that the sinusoidal affection is correlated with the severity of fat accumulation in the parenchymal cells. According to the findings of Khan and Chakarabarti (2003); Hayden *et al*. (2005) and Ban and Twigg (2008), hyperglycemia is the main offending factor in the onset of the micro-vascular diabetic complications. Fortunately, following *Citrullus colocynthis* intake, in their study, the damaged sinusoids, the haemorrhage and the inflammatory cell infiltration subsided indicating a beneficial role of such a remedy. These findings are supported by those claimed by Alarcon-Aguilar *et al*. (2000). Despite the obvious antidiabetic effect of *Citrullus colocynthis* pulp extract, its use in normal rats, in the current work, caused hepatocytic poration and few hapatocytes had condensed or fragmented nuclei indicating minimal hepatotoxicity.

This herbal therapy of *Gymnema sylvestre (Asclepiadaceae*) appears to bring about blood glucose homeostasis through increased serum insulin levels provided by repair/regeneration of the endocrine pancreas. Also, the effectiveness of the water extract from the leaves of *Gymnema sylvestre*, in controlling hyperglycemia was investigated in type II diabetic patients as conventional per os anti-hyperglycemic agents (Baskaran et al., 1990). The obtained data suggested that the β cells may be regenerated / repaired in type II diabetic patients on the extract's supplementation. This was supported by the raised insulin levels in the serum of patients after the supplementation. Furthermore, the antihyperglycemic action of a crude saponin fraction and five triterpene glycosides (gymnemic acids I-IV and gymnemasaponin V) derived from the methanol extract of leaves of *Gymnema sylvestre* was investigated in STZ-diabetic mice (Sugihara et al., 2000).

The results indicated that insulin-releasing action of gymnemic acid IV, administered i.p., may contribute to the antihyperglycemic effect by the leaves of Gymnema sylvestre.

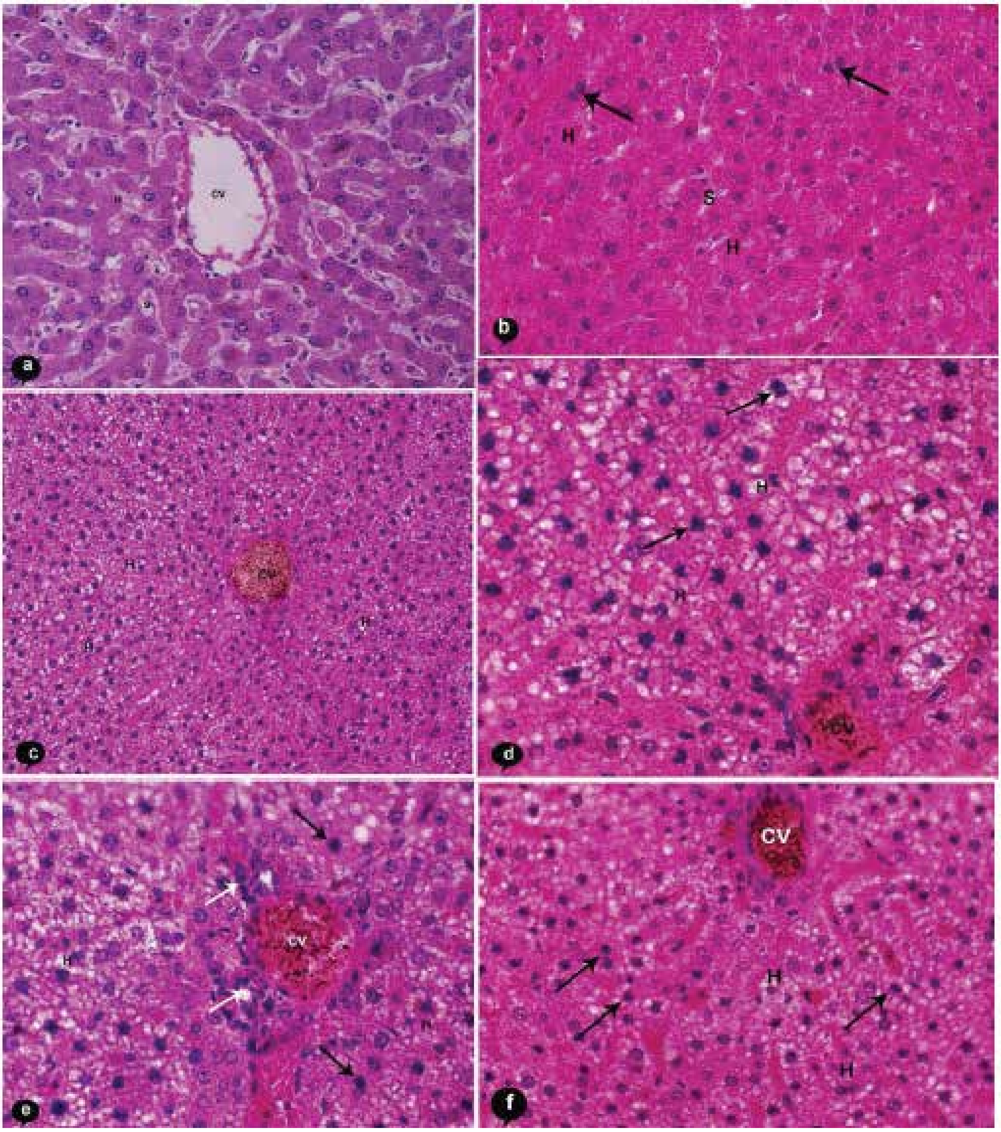


Fig. 3. Photomicrographs of the livers of control, *Citrullus colocynthis*- treated, diabetic and *Citrullus colocynthis*treated diabetic rats, H&E stain; (a) Liver of control rats, ×40; (b) Liver of *Citrullus colocynthis* treated rats, showing few hapatocytes had condensed or fragmented nuclei ×40; (c) Liver of diabetic rats, showing disorganized hepatic cords, reduced sinusoids and many hepatocytes having cytoplasmic vacuolization, ×20; (d) Higher magnification of (c), showing most hepatocytes with cytoplasmic vacuolar degeneration and pyknotic nuclei (arrows), ×40; (e) Liver diabetic rats, showing a central inflammatory cell infiltration (white arrows) and hepatocytes with cytoplasmic vacuolar degeneration and pyknotic nuclei (black arrows), ×40; (f) Liver of diabetic rats treated with *Citrullus colocynthis*, showing recovered hepatocytes with less cytoplasmic vacuolization compared with diabetic animals. Few hepatocytes with cytoplasmic vacuolization and pyknotic nuclei are still seen (arrows), ×40. According to the study occurred by Mohammad Khalil, et al. (2010).

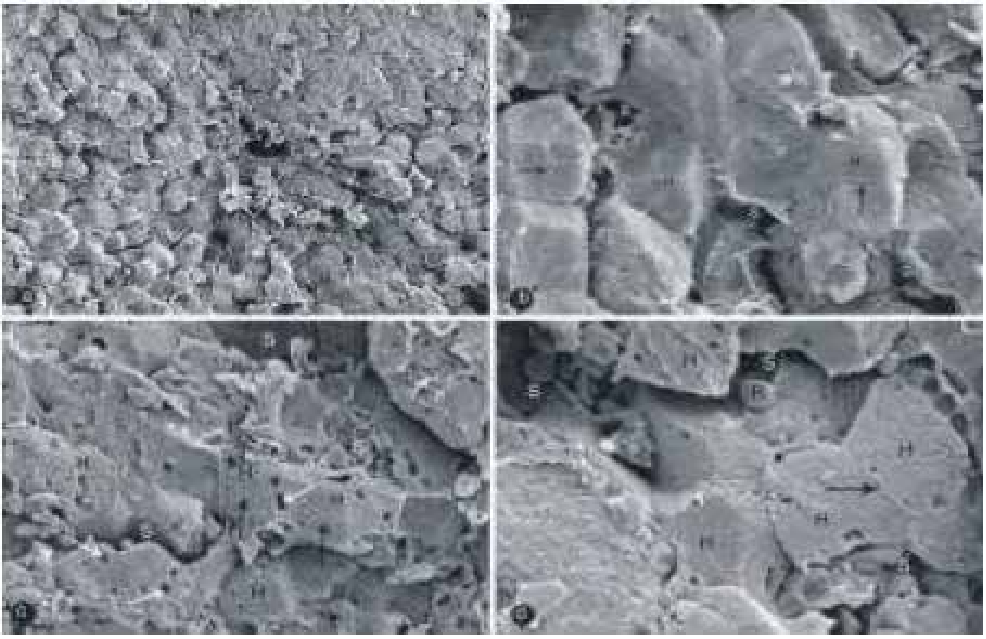


Fig. 4. Scanning electron micrographs of livers of control and *Citrullus colocynthis*-treated rats. (a) SEM of the liverof control rats showing normal hepatic cord around a Central Vein (CV), ×450; (b) Larger magnification of (a) showing normal Sinusoids (S) and normal Hepatocytes (H) with normal intercellular boundaries (arrows), ×1500 (c); SEM of the liver of *Citrullus colocynthis*-treated rats, showing rows of Hepatocytes (H) containing few pores (white arrows). Note the blood Sinusoids (S) and the intercellular boundaries (black arrows), ×1100; (d) Higher magnification of (c), showing blood Sinusoids (S), Hepatocytes (H), intercellular spaces (black arrows) and erythrocytes (R), ×1500. According to the study occurred by Mohammad Khalil, et al. (2010).

Gymnemic acid IV may be an anti-obese and antihyperglycemic prodrug. The inhibitory activity of each triterpene glycoside on the glucose uptake in rat small intestine fragments was examined, in order to determine its impact on the increase of serum glucose level in glucose-loaded rats (Yoshikawa et al., 1997). It was found that Gymnemic acids II and III showed potent inhibitory activity on glucose uptake. Gymnemoside b and gymnemic acids III, V, VII were found to exhibit a little inhibitory activity against glucose uptake, and the principal constituents, gymnemic acid I and gymnemasaponin V, lacked this activity. It is noteworthy that, although Gymnema saponin constituents such as Gymnemic acids II and III show no effect on serum glucose levels in oral-loaded rats, they exhibit potent inhibitory activity on the glucose uptake and further studies need to be contacted.

The hypoglycemic activity of a decoction from *Juniperus communis* (juniper berries) both in normal glycemic and in STZ-diabetic rats was studied (Sanchez et al., 1994). Juniper decoction decreased glycemic levels in normal glycemic rats through an increase of peripheral glucose utilization or a potentiation of glucose induced insulin secretion.

The per os administration of the decoction to STZ diabetic rats resulted in a significant reduction both in blood glucose levels and in the mortality index, as well as the prevention

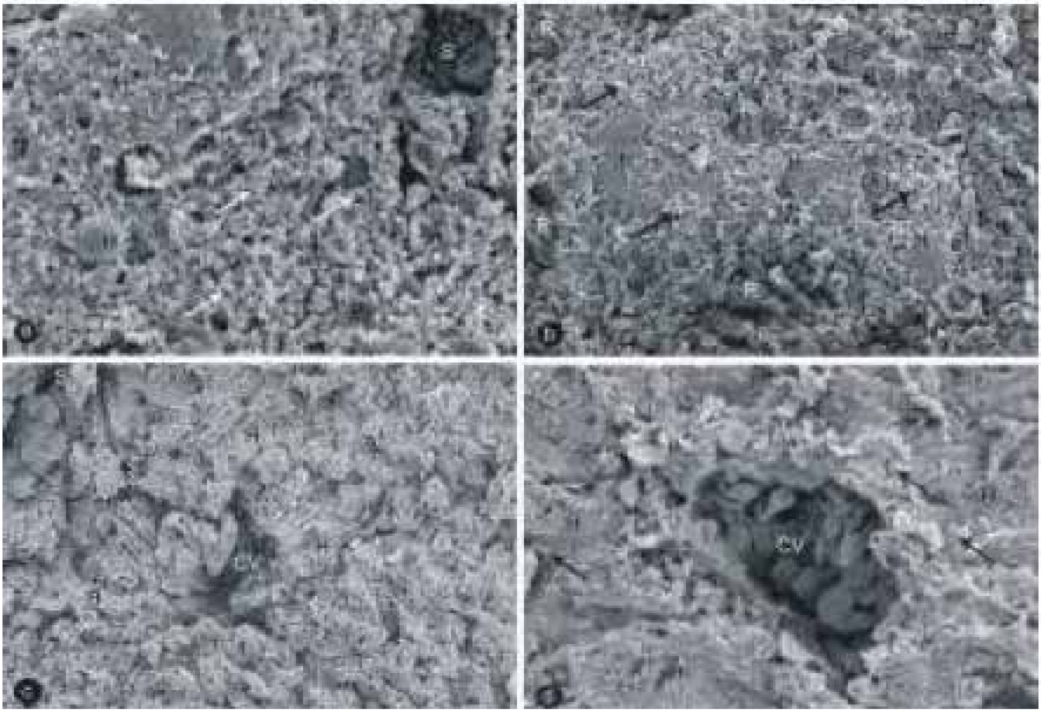


Fig. 5. Scanning electron micrographs of livers of diabetic and *Citrullus colocynthis*-treated diabetic rats. (a) SEM of the liver of diabetic rats showing degenerating Hepatocytes (H) having a plenty of lipid droplets (white arrows). Damaged blood Sinusoids (S) are also seen as well as some erythrocytes (R) are spread inside hepatocytes, ×450; (b) SEM of the liver of diabetic rats showing hemorrhage of erythrocytes (R) between Hepatocytes (H). Lipid degeneration (black arrows) of hepatocytes is also seen, ×450; (c) SEM of the liver of *Citrullus colocynthis*-treated diabetic rats, showing intact Hepatocytes (H) and blood Sinusoids (S). Only fewhepatocytes around the central vein have some vacuoles, ×450; (d) SEM of the liver of *Citrullus colocynthis* treated diabetic rats (300 mg) with higher magnification, showing intact Hepatocytes (H) and a Central Vein (CV) containing a lot of erythrocytes. A few hepatocytes around the central vein have a few lipid droplets (arrows), ×1100. According to the study occurred by Mohammad Khalil, et al. (2010).

of the loss of body weight. This effect seemed to be mediated by the peripheral action of juniper while in Olea europaea (Oleaceae) plant species, the hypoglycemic activity may result from two mechanisms: (a) potentiation of glucose induced insulin release, and (b) increased peripheral uptake of glucose.

The hypoglycemic effect of the rhizomes of *Polygala senega* (Polygalaceae) was proposed that after i.p. administration was without altering the insulin levels and with the need of the presence of insulin in order to act. In addition, one of the active components of this hypoglycemic effect was identified as a triterpenoid glycoside, senegin-II (as a mixture of isomers). *Solanum lycocarpum (Solanaceae*) has been widely employed for diabetes management, obesity and to decrease cholesterol levels. Some of the polysaccharides slowed gastric emptying and act on the endocrinous system affecting the liberation of gastrointestinal hormones, lowering blood glucose levels. The hypocholesterolemic activity could be due to the increased fecal bile acid excretion as well as to the action of the short-chain fatty acids, coming from fermentation, on the synthesis of deltaaminolevulinate and by the increase of the cholesterol 7-alpha-hydroxylase and 3-hydroxy-3-methylglutaryl CoA reductase synthesis (DalAgnol & Lino von Poser 2000).

4. Conclusions and future directions

Scientists from divergent fields are investigating plants anew with an eye to their antimicrobial usefulness. A sense of urgency accompanies the search as the pace of species extinction continues. Laboratories of the world have found literally thousands of phytochemicals which have inhibitory effects on all types of microorganisms in vitro. More of these compounds should be subjected to animal and human studies to determine their effectiveness in whole-organism systems, including in particular toxicity studies as well as an examination of their effects on beneficial normal microbiota. It would be advantageous to standardize methods of extraction and in vitro testing so that the search could be more systematic and interpretation of results would be facilitated. Also, alternative mechanisms of infection prevention and treatment should be included in initial activity screenings. Disruption of adhesion is one example of an anti-infection activity not commonly screened for currently. Attention to these issues could usher in a badly needed new era of chemotherapeutic treatment of infection by using plant-derived principles.

5. References

Aburjai, T., M. Hudaib, R. Tayyem, M. Yousef and M. Qishawi, 2007. Ethnopharmacological survey of medicinal herbs in Jordan, the Ajloun heights egion. J. Ethnopharmacol., 110: 294-304. PMID: 7097250

Alarcon-Aguilar, F.J., M. Jimenez-Estrad, R. Reyes- Chilpa and R. Roman-Ramos, 2000. Hypoglycemic effect of extracts and fractions from Psacalium decompositum in healthy and alloxan-diabetic mice. J. Ethnopharmacol., 72:

Arise, R.O., S.O. Malomo, J.O. Adebayo and A. Igunnu, 2009. Effects of aqueous extract of eucalyptus globules on lipid peroxidation and selected enzymes of rat liver. J. Med. Plants Res., 3: 77-81

Brixova, E., 1981. Experimental and clinical liver steatosis. Folia Fac. Med. Univ. Comenian Bratisl., 19: 9-90.

Ban, C.R. and S.M. Twigg, 2008. Fibrosis in diabetes complications: Pathogenic mechanisms and circulating and urinary markers. Vasc. Health Risk Manage., 4: 575-596. PMID: 18827908

Bolkent S, Yanardag R, Tabakoglu-Oguz A, Ozsoy-Sacan A. "Effects of chard (Beta bulgaris L. var. cicla) extract on pancreatic B cells in diabetic rats: a morphological and biochemical study". *J. Ethnopharmacol.* 2000; 73: 251-259.

Baskaran K, Kizar Ahamath B, Radha Shanmugasundaram K, ShanmugasundaramER. "Antidiabetic effects of a leaf extract from Gymnema sylvestre in non-insulin-dependent diabetes mellitus patients". *J. Ethnopharmacol.* 1990; 30(3): 295-300.

Bujanda, L., E. Hijono, M. Larzabal, M. Beraza and P. Aldazabal *et al.*, 2008. Resveratrol inhibits non-alcoholic fatty liver disease in rats. BMC Gastroenterol., 8: 40-40. DOI: 10.1186/1471-230X-8-40

Chang SL, Chang CL, Chiang YM, Hsieh RH, Tzeng CR, Wu TK, Sytwu HK, Shyur LF, Yang WC. "Polyacetylenic compounds and butanol fraction from Bidens pilosa can modulate the differentiation of helper T cells and prevent autoimmune diabetes in non-obese diabetic mice". *Planta Med.* 2004; 70(11): 1045-1051.

Chen TH, Chen SC, Chan P, Chu YL, Yang HY, Cheng JT. " Mechanism of the hypoglycemic effect of stevioside, a glycoside of Stevia rebaudiana". *Planta Med.* 2005; 71(2): 108-113. Dall'Agnol R, Lino von Poser G. "The use of complex polysaccharides in the management of metabolic diseases: the case of Solanum lycocarpum fruits". *J. Ethnopharmacol.* 2000; 71(1-2): 337-341.

Fagot-Campagna A. & Narayan K. "Type 2 diabetes in children". *Br. Med. J.* 2001: 322: 377-387.

Fugh-Beerman A. "Herb-drug interactions". *Lancet* 2000; 355: 134-138.

Fujisawa T, Ikegami H, Kawaguchi Y, Nojima K, Kawabata Y, Ono M, Nishino M, Noso S. Taniguchi H, Horiki M, Itoi-Babaya M, Babaya N, Inoue K, Ogihata T. "Common genetic basis between type I and type II diabetes mellitus indicated by interview-based assessment of family history ". *Diabetes Res. Clin. Pract.* 2004; 66S:S91-S95

Ghosh, S. and S.A. Surawanshi, 2001. Effect of *Vinca rosea* extracts in treatment of alloxan diabetes in male albino rats. Indian J. Exp. Biol., 39: 748759.PMID:12018575

Grover JK, Vats V, Yadav SS. "Pterocarpus marsupium extract (Vijayasar) prevented the alteration in metabolic patterns induced in the normal rat by feeding an adequate diet containing fructose as sole carbohydrate". *Diabetes Obes. Metab.* 2005; 7(4): 414-417.

Huxtable RJ. "The harmfull potential of herbal and other plant products". Drug Safety 1990; 5(Suppl. 1): 126-136.

Hsu FL, Lai CW, Cheng JT. "Antihyperglycemic effects of paeoniflorin and 8-debenzoylpaeoniflorin, glycosides from the root of Paeonia lactiflora". *Planta Med.* 1997; 63(4): 323-325.

Hayden, M.R., J.R. Sowers and S.C. Tyagi, 2005. The central role of vascular extracellular matrix and basement membrane remodeling in metabolic syndrome in type and 2 diabetes: The matrix preloaded. Cardiovasc. Diabetol., 4: 9-9. DOI: 10.1186/1475-2840-4-9

Khan, Z.A. and S. Chakarabarti, 2003. Endothelins in chronic diabetic complications. Can. J. Physiol. Pharmol., 81: 622-634. DOI: 10.1139/Y03-053

Keith, K.G., V. Fonseca, M.H. Tan and A. Dalpiaz, 2004. Narrative review: Hepatobiliary disease in type 2 diabetes mellitus. Ann. Intern. Med., 141: 946-956. PMID: 15611492

Koyama Y, Abe K, Sano Y, Ishizaki Y, Njelekela M, Shoji Y, Hara Y, Isemura M. "Effects of green tea on gene expression of hepatic gluconeogenic enzymes in vivo". *Planta Med.* 2004; 70(11): 1100-1102.

Karageuzyan KG, Vartanyan GS, Agadjanov MI, Panossian AG, Hoult JR. "Restoration of the disordered glucose-fatty acid cycle in alloxan-diabetic rats by trihydroxyoctadecadienoic acids from Bryonia alba, a native Armenian medicinal plant". *Planta Med.* 1998; 64(5): 417-422.

Ji Su Kim, Jung Bong Ju , Chang Won Choi and Sei Chang Kim (2006): Hypoglycemic and Antihyperlipidemic Effect of Four Korean Medicinal Plants in Alloxan Induced Diabetic Rats American Journal of Biochemistry and Biotechnology 2 (4): 154-160, 2006 ISSN 1553-3468

Luo J, Chuang T, Cheung J, Quan J, Tsai J, Sullivan C, Hector RF, Reed MJ, Meszaros K, King SR, Carlson TJ, Reaven GM. "Masoprocol (nordihydroguaiaretic acid): a new antihyperglycemic agent isolated from the creosote bush (Larrea tridentata)". *Eur. J. Pharmacol.* 1998; 346(1): 77-79.

Luo J, Fort DM, Carlson TJ, Noamesi BK, nii-Amon-Kotei D, King SR, Tsai J, Quan J, Hobensack C, Lapresca P, Waldeck N, Mendez CD, Jolad SD, Bierer DE, Reaven GM. " Cryptolepis sanguinolenta: an ethnobotanical approach to drug discovery and the isolation of a potentially useful new antihyperglycemic agent". *Diabet. Med.* 1998; 15(5): 367-374.

Lyra, R., M. Oliveira, D. Lins and N. Cavalcanti, 2006. Prevention of type 2 diabetes mellitus. Arq. Bras. Endocrinol. Metabo,. 50: 239-249.

MacFarlance IA, Bliss M, Jackson JG, Williams G. "The history of diabetes mellitus. In textbook of diabetes". (eds Pichup, j. Williams G.) Blackwell, London 1997, 2nd edn, vol. I, pp 1-21.

Moller, D.E., 2001. New drug targets for type 2 diabetes and the metabolic syndrome. Nature, 414: 821-827. PMID: 11742415

Mohammad Khalil, Gamal Mohamed, Mohammad Dallak, Fahaid Al-Hashem, Hussein Sakr, Refaat A. Eid, Mohamed A. Adly, Mahmoud Al-Khateeb, Saleh Banihani, Zuhair Hassan and Nabil Bashir (2010):The Effect of *Citrullus colocynthis* Pulp Extract on the Liver of Diabetic Rats a Light and Scanning Electron Microscopic Study. American Journal of Biochemistry and Biotechnology 6 (3): 155-163, 2010 ISSN 1553-3468

Naik SR, Barbosa Filho JM, Dhuley JN, Deshmukh V. "Probable mechanism of hypoglycemic activity of bassic acid, a natural product isolated from Bumelia sartorum". *J. Ethnopharmacol.* 1991; 33(1-2): 37-44.

Pushparaj PN, Tan BK, Tan CH. "The mechanism of hypoglycemic action of the semi-purified fractions of Averrhoa bilimbi in streptozotocin-diabetic rats". *Life Sci.* 2001; 70(5): 535-547.

Rajasekaran S, Sivagnanam K, Ravi K, Subramanian S. "Hypoglycemic Effect of Aloe vera Gel on Streptozotocin-Induced Diabetes in Experimental Rats". *J. Med. Food.* 2004; 7(1): 61-66.

Rollo J. "An account of two cases of the diabetes mellitus, with remarks as they arose during the progress of the cure". C. Dilly, London, 1797.

Sanchez de Medina F, Gamez MJ, Jimenez I, Jimenez J, Osuna JI, Zarzuelo A."Hypoglycemic activity of juniper berries". Planta Med. 1994; 60(3): 197-200.

Sebbagh, N., C. Cruciani-Guglielmacci, F. Cruciani, Guglielmacci, M.F. Berthault and C. Rouch *et al.*, 2009. Comparative effects of *Citrullus colocynthis*, sunflower and olive oil-enriched diet on streptozotocin-induced diabetes in rats. Diabetes Metab., 35: 178-184. PMID: 19264524

Seifalian, A.M., C. Piasecki, A. Agawel and B.R. Davidson, 1999. The effect of graded steatosis on flow in the hepatic on flow in the hepatic parenchymal microcirculation. Transplantation, 68: 780-784. PMID: 10515377

Sugihara Y, Nojima H, Matsuda H, Murakami T, Yoshikawa M, Kimura I."Antihyperglycemic effects of gymnemic acid IV, a compound derived fromGymnema sylvestre leaves in streptozotocin-diabetic mice". *J. Asian Nat. Prod. Res.* 2000; 2(4): 321-327.

Ugochukwu NH, Babady NE. "Antihyperglycemic effect of aqueous and ethanolic extracts of Gongronema latifolium leaves on glucose and glycogen metabolism in livers of normal and STZ-diabetic rats". *Life Sci.* 2003; 73(15): 1925-1938.

Ugochukwu NH, Babady NE. "Antioxidant effects of Gongronema latifolium in hepatocytes of rat models of non-insulin dependent diabetes mellitus". *Fitoterapia* 2002; 73: 612-618.

Ugochukwu NH, Cobourne MK. "Modification of renal oxidative stress and lipid peroxidation in STZ-diabetic rats treated with extracts from Gongronema latifolium leaves". *Clinica Chimica Acta* 2003; 336: 73-81

Venkateswaran S, Pari L. "Effect of Coccinia indica leaves on antioxidant status in STZ-diabetic rats". *J. Ethnopharmacol.* 2003; 84: 163-168.

Venkateswaran S, Pari L. "Effect of Coccinia indica on Blood Glucose, Insulin and Key Hepatic Enzymes in Experimental Diabetes". *Int. J. Pharmacogn.* 2002; 40(3): 165-170.

Waltner-Law, M.E., X.L. Wang, B.K. Law, R.K. Hall and M. Nawano *et al.*, 2002. Epigallocatechin gallate, a constituent of green tea, represses hepatic glucose production. J. Biol. Chem., 277: 34933-34940. PMID: 12118006

WHO (1978a). Alma Ata Declaration. Primary Health Care. Health for all series No.1

World Health Organization, 1980. WHO expert committee on diabetes mellitus.

Witters L. "The blooming of the French lilac". *J. Clin. Invest.* 2001; 108: 1105-1107.

Yoshikawa M, Murakami T, Kadoya M,. "Medicinal foodstuffs. III. Sugar beet. (1): Hypoglycemic oleanolic acid oligoglycosides,betavulgarosides I, II, III, and IV, from the root of Beta vulgaris L. (Chenopodiaceae)". *Chem. Pharm. Bull. (Tokyo)* 1996; 44: 1212-1217.

Yoshikawa M, Murakami T, Kadoya M, Li Y, Murakami N, Yamahara J, Matsuda H. "Medicinal foodstuffs. IX. The inhibitors of glucose absorption from the leaves of Gymnema sylvestre R. BR. (Asclepiadaceae): structures of gymnemosides a and b". *Chem. Pharm. Bull. (Tokyo).* 1997; 45(10): 1671-1676.

Yoshikawa M, Shimada H, Nishida N, Li Y, Toguchida I, Yamahara J, Matsuda H. "Antidiabetic principles of natural medicines. ALR2 and alpha-glucosidase inhibitors from Brazilian natural medicine, the leaves of Myrcia multiflora DC. (Myrtaceae): structures of myrciacitrins I and II and myrciaphenones A and B". *Chem. Pharm. Bull. (Tokyo).* 1998; 46(1): 113-119.

8

In Vitro Leukocyte Adhesion in Endothelial Tissue Culture Models Under Flow

Scott Cooper, Melissa Dick, Alexander Emmott,
Paul Jonak, Léonie Rouleau and Richard L. Leask
*Department of Chemical Engineering, McGill University,
Canada*

1. Introduction

Atherosclerosis, an inflammatory disease which causes thickening and stiffening of arteries, is a major cause of death in the United States (Lloyd-Jones et al., 2010). These deaths occur because of vessel occlusion created by atherosclerotic plaques and thrombus shedding, leading to heart attack, ischemia, or stroke. Atherosclerosis is expected to be the leading cause of death worldwide within 10 years (Lloyd-Jones et al., 2010).

Inflammation plays a significant role in the initiation and progression of atherosclerosis. The cells that line the arteries, endothelial cells (ECs), mediate the inflammatory process. The forces created by blood flow affect the inflammatory response of ECs and the interaction with blood components. This chapter summarizes the background in our recent studies on the response of ECs to blood forces and the interaction of inflammatory cells.

2. Background

2.1 Pathogenesis of atherosclerosis

Vascular anatomy

Arteries have three tissue layers: the intima, media, and adventitia. The intima is lined with a monolayer of ECs in direct contact with blood. The ECs act as a protective membrane, allowing diffusion from the blood stream into the artery. ECs are capable of expressing specific genes in response to physical stresses which cause the vessel to remodel leading to the development of atherosclerosis. In addition, the intima can contain other cells (smooth muscle cells, fibroblasts and inflammatory cells), an extracellular matrix (ECM), and is only a few cell layers thick in healthy tissue. The internal elastic membrane, consisting of a layer of elastic connective tissues, separates the intima and media. The media layer is mainly comprised of smooth muscle cells (SMCs) and ECM. Although the media is involved in atherosclerosis, remodelling is less evident. The adventitia is relatively unaffected by atherosclerosis. It is separated from the media by the external elastic membrane comprised mainly of collagen, providing structural support yet allowing for artery expansion when required (Waller et al., 1992).

Over time, an atherosclerotic plaque grows by the accumulation of lipids, inflammatory cells, vascular cells and matrix material in the intima. It often produces a fibrous cap, over a necrotic lipid core, which can weather and rupture over time. The artery is able to compensate for some intimal thickening by expanding outwards, instead of allowing for the plaque to impede blood flow. Eventually the vessel can no longer expand outwards, and negative remodelling can occur. Blood flow is therefore disturbed through the formation of a stenosis (Shah, 2006). This may lead to ischemia and angina pectoris (Libby, 2002).

Atherosclerosis can occur in any size of artery. However, clinical manifestations frequently occur in medium and large arteries when the EC layer is breached by erosion or disruption of the fibrous cap. Disruption may occur from a thinning of the fibrous cap as there is increased lipid accumulation, inflammatory cell recruitment and matrix metalloproteinase (MMP) expression, as well as the expression of cytokines inhibiting collagen synthesis. When the plaque is opened up to the blood stream, platelets cause blood coagulation and thrombus formation. There are two possible outcomes after thrombus formation. First, the thrombus may be broken down and reabsorbed. A second outcome is that the thrombus is disrupted by the blood flow and detached from the site of injury. This embolism may then travel through the vasculature to small arteries, where it causes ischemia and potential heart attack or stroke (Libby, 2002).

The role of inflammation

Over the past two decades, it has been recognized that inflammation plays a critical role in the development and progression of atherosclerosis (Libby, 2002). Indeed, the localization of plaques to regions of disturbed blood flow (curvature, bifurcations, and branches) has been linked to an inflammatory response of ECs due to hemodynamic forces (Libby, 2002; Shah, 2006). In these areas, ECs become inflamed, causing an influx of leukocytes (Shah, 2006). It has been found that nuclear factor κB (NF-κB), a transcription factor responsible for expressing genes involved in the inflammatory cascade, is activated at sites of disturbed blood flow (Van der Heiden et al., 2010).

Additionally, monocytes, part of the family of leukocytes, are attracted into the intima through the EC layer due to the existence of a chemical gradient. During inflammation, a chemokine called monocyte chemoattractant protein-1 (MCP-1) is expressed within the intima layer (Libby, 2002). MCP-1 is expressed constitutively, by both the EC layer, and the SMCs within the intima (Schwartz et al., 1991). The receptor for MCP-1 on the monocyte (the CCR2 receptor) is attracted to the MCP-1 within the intima, and monocytes migrate into the intima through diapedesis (Libby, 2002).

Also flowing in the blood stream are low-density lipoproteins (LDL), including cholesterol. LDLs are brought across the EC layer and into the intima. Reactive oxygen species within the intima, including $^{\cdot}OH$ and $^{\cdot}O_2$, oxidize the LDLs, turning them into oxidized low-density lipoproteins (Ox-LDL). These Ox-LDL molecules are also responsible for stimulating ECs and SMCs to secrete additional MCP-1 (Schwartz et al., 1991).

Once inside the intima, monocytes begin to express characteristics of macrophages, activated by the presence of macrophage colony stimulating factor (M-CSF) (Libby, 2002). M-CSF also activates Ox-LDL receptors on the macrophages, turning them into scavengers for Ox-LDL (Libby et al., 2002). Macrophages begin to take up Ox-LDL, filling themselves with lipids and transforming into foam cells (Ross, 1993; Shah, 2006). The accumulation of macrophage foam cells, as well as collagen, elastin, and proteoglycans, within the intimal layer is known as the fatty streak, and is an early indication of a complex atherosclerotic lesion (Ross, 1993; Schwartz et al., 1991). It has been shown that the progression of a foam cell to a more advanced lesion may be halted or reversed, possibly through a decrease in blood LDL levels (Schwartz et al., 1991).

In advanced lesions, macrophages are unable to take up any additional Ox-LDL, and these lipid molecules begin to accumulate within the intima instead (Schwartz et al., 1991). Ox-LDLs are toxic, and begin to injure and kill ECs, SMCs, and macrophages. When the macrophage is injured, its lipid contents are released into the intimal layer, forming a lipid core, also comprised of enzymes, cytokines, and growth factors that have accumulated within the intima (Schwartz et al., 1991; Shah, 2006).

In order to protect the body from the growing necrotic lipid core within the intima, a fibrous plaque is formed over the lesion. This prevents direct contact between the accumulation of cells within the intima and blood flow. The fibrous plaque is composed primarily of collagen, elastin, and proteoglycans that were found in the original fatty streak (Ross, 1999). T lymphocytes and macrophages release MMPs, which break down the extracellular matrix within the intima, allowing for these components to be used within the fibrous plaque (Libby, 2002).

2.2 Mechanisms in leukocyte-endothelium adhesion

2.2.1 Intercellular adhesion at the arterial surface

The initiation of atherosclerosis is hypothesized to start with endothelial injury, which triggers inflammatory pathways integral to the progression of the disease (Ross et al., 1977). Leukocytes, including neutrophils and monocytes, preferentially adhere to sites of inflammation. The events that take place during leukocyte recruitment are shown in the representative drawing, Figure 1.

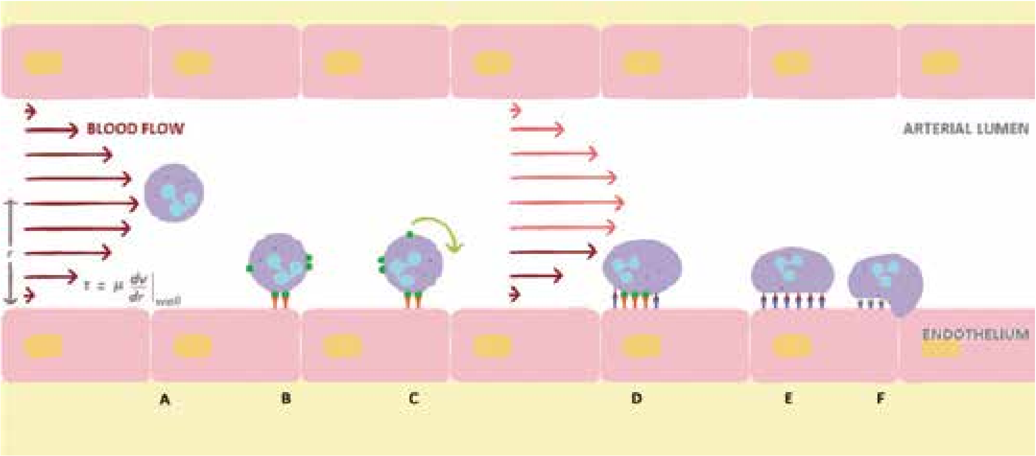


Fig. 1. Leukocyte recruitment to a site of endothelial injury. When injured, the endothelium expresses an increase in cell adhesion molecules (CAMs). (A) Leukocytes circulate within the blood stream. (B) Ligands on the leukocytes attach to selectins on the endothelium, effectively tethering the leukocyte. (C) The leukocyte begins to roll across the endothelium, reducing its velocity by forming and breaking selectin-ligand bonds. (D) Leukocytes begin to make bonds between CAMs and integrins. This firmly attaches the leukocyte to the endothelium. (E) Leukocytes migrate towards a cellular junction through CAM-integrin interactions. (F) CAM-CAM interactions allow the leukocyte to migrate through the endothelium by diapedesis.

Leukocyte tethering and rolling

Leukocytes circulating within the bloodstream must make their way to the site of injury, located on the endothelium. Although leukocytes flow in close contact with the endothelial layer, they do not stick until the inflamed endothelium starts to express adhesive molecules (Kelly et al., 2007). Tethering and rolling of leukocytes along the endothelial wall is due to a class of cell adhesion molecules (CAMs) known as selectins. There are three selectins involved: E-selectin, P-selectin, and L-selectin. Both E- and P-selectins are expressed on the endothelium; E-selectin is synthesized and expressed after endothelial stimulation, whereas P-selectin is expressed constitutively and stored, then quickly released upon stimulation (Kelly et al., 2007). L-selectin differs in that it is not expressed on the endothelium, but instead is constitutively expressed on the leukocyte surface. Both E- and P-selectin recognize carbohydrate ligands on the surface of leukocytes, while L-selectin recognizes a series of ligands expressed on the endothelium. When the selectins come in contact with their ligands they will bind, tethering the leukocyte to the endothelium (Miyasaka et al., 1997). Tethering facilitates leukocyte rolling along the EC surface. The velocity of the travelling leukocyte will be reduced as more selectin-ligand bonds form, allowing leukocyte adhesion (Kelly et al., 2007; Kubes & Kerfoot, 2001).

Leukocyte adhesion

Firm adhesion begins when integrins, another class of CAMs, are activated on the surface of leukocytes by chemoattractant cytokines, termed chemokines. Chemokines are secreted by circulating leukocytes and ECs (Kelly et al., 2007). One class of integrin responsible for neutrophil adhesion is the β_2-integrin. When activated by cytokines, certain β_2-integrins bind to intracellular adhesion molecules 1 and 2 (ICAM-1 and ICAM-2, respectively) on the EC surface, effectively adhering the leukocyte to the EC. Also activated by cytokines, $\alpha_4\beta_1$-integrin will bind with vascular cell adhesion molecule 1 (VCAM-1) (Kelly et al., 2007; Rao et al., 2007). The bound integrin-CAM complex results in firm adhesion of the leukocyte to the endothelial surface (Miyasaka et al., 1997). Although under transcriptional regulation, both adhesion molecules ICAM-1 and VCAM-1 have been shown to be upregulated at sites in the vasculature prone to developing atherosclerotic lesions (Iiyama et al., 1999).

Leukocyte migration

Once the leukocytes have firmly adhered to the endothelium, they may migrate from the lumen of the blood vessel into the subintimal space. Initially, leukocytes must make their way to the closest EC junction through a process termed locomotion. The movement of the leukocyte is made possible through interactions with leukocyte integrins and both ICAM-1 and -2 located on the endothelial surface (Schenkel et al., 2004). At the junction, the leukocytes will encounter another cell adhesion molecule, called platelet endothelial CAM 1 (PECAM-1). PECAM-1 is expressed both on the leukocyte and the endothelium. An interaction between the complementary PECAM-1 molecules allows leukocytes to migrate through the gap junction by diapedesis (Rao et al., 2007; Schenkel et al., 2004), a process also called transmigration.

2.2.2 Idealized arterial hemodynamics

Hemodynamics, the mechanics of blood flow, influence many of the physiological processes of the vascular system (Glagov et al., 1988). From an engineering perspective, blood flow through medium and small arteries (such as the right and left coronary arteries) is often simplified by assuming steady laminar flow in a straight, rigid vessel (Ku, 1997; Nichols & O'Rourke, 1990). Additionally, blood is assumed to be a Newtonian fluid to simplify the flow dynamics to Hagen-Poiseuille flow (Nichols & O'Rourke, 1990). Such assumptions allow us to reduce the governing equations describing pressure-driven flow into a one-dimensional velocity profile in the form of Hagen-Poiseuille flow (Nichols & O'Rourke, 1990). For a cylindrical vessel model of arterial perfusion, the velocity profile as a function of a radial dimension is described by:

$$v(r) = 2\left(\frac{Q}{\pi R^2}\right)\left[1 - \left(\frac{r}{R}\right)^2\right] \tag{1}$$

where Q is the volumetric flow rate of blood and R is the hydraulic radius of the vessel.

Wall shear stress (WSS) is a tangential force per unit area of a fluid-wall interface that results from flow parallel to the vessel wall. For fluids with constant dynamic viscosity μ (Newtonian fluid), the WSS is the product of the viscosity and the shear rate γ, evaluated at the vessel wall:

$$\tau_w = \mu\gamma|_{r=R} \tag{2}$$

The wall shear rate of a fluid is the velocity gradient evaluated at the fluid-wall interface. For Hagen-Poiseuille flow through a cylindrical vessel (Equation 1), the wall shear rate is expressed as:

$$\gamma|_{r=R} = \left.\frac{dv(r)}{dr}\right|_{r=R} = \frac{4Q}{\pi R^3} \tag{3}$$

Combining Equations (2) and (3) produces an expression for WSS that is dependent on both the vessel geometry and the volumetric flow rate:

$$\tau_w = \mu\frac{4Q}{\pi R^3} \tag{4}$$

This equation is accepted as a reasonable model of the average WSS for arteries that are absent from serious geometric disturbances (Ku, 1997). Arterial WSS values range from 5 to 70 dyne/cm^2 with average WSS values of approximately 15 dyne/cm^2 being observed in coronary arteries (Glagov et al., 1988; Malek et al., 1999). Moderate levels of steady, laminar shear stress (> 10-15 dyne/cm^2) are believed to induce an atheroprotective EC phenotype while low shear stresses (< 4 dyne/cm^2) are believed to induce an atheroprone EC phenotype (Malek et al., 1999). An atheroprone EC phenotype describes one which facilitates the disease pathway marked by an increase in adhesion molecules and a decrease in vasodilators (as described in Sections 2.1 & 2.2) (Libby, 2002; Malek et al., 1999).

The fluid flow regime is determined by the dimensionless Reynolds number (Re), which represents the ratio of inertial to viscous forces. For flow in a cylindrical channel, the Reynolds Number is described by:

$$Re = \frac{\rho DU}{\mu} \tag{5}$$

where ρ is the fluid density, D is the hydraulic diameter, and U is the average fluid velocity. A Reynolds Number below 2300 indicates laminar flow that will behave predictably while a value above this threshold suggests the presence of flow disturbances. The average arterial conditions are within a laminar flow regime (Nichols & O'Rourke, 1990) and vary depending on the artery and metabolic demand (Myers et al., 2001; Nichols & O'Rourke, 1990).

Localized hemodynamics of leukocyte adhesion

The progression of leukocyte adhesion is strongly influenced by local hemodynamic forces. As the cell is passing along the wall, the torque imparted on the cell by the blood stream causes the cell to spin. As a result, the state of loose attachment with selectin-ligand bonds constantly forming and breaking has become known as cell rolling. The blood stream imposes not only torque but also shear stress on the slow moving cell. In turn, the membrane of the cell will try to distribute this stress by elongating in the direction of flow, allowing for increased binding with the vessel wall. Firrell and Lipowsky found that leukocytes rolling along rat arteriolar walls would elongate by around 140%, allowing their contact area to jump from approximately 14 μm^2 to 50 μm^2 (Firrell & Lipowsky, 1989).

Modelling of bond forces and leukocyte attachment is well documented in the literature (Cozens-Roberts et al., 1990; Evans et al., 2004; Lawrence et al., 1997; Tees & Goetz, 2003). For successful adhesion, a fine balance between the adhesive force and the hemodynamic force must be met. This adhesive force is dependent on several factors, including: the receptor density of both cells, the rate of reaction with respect to both bond formation and dissociation and the strength of the bonds and their response to strain. For instance, the bonds between E-, P-, and L-selectins and their respective ligands behave as catch-slip bonds (Lawrence et al., 1997; Marshall et al., 2003; Sarangapani et al., 2004). Though receptor-ligand bonds spontaneously dissociate (Tees & Goetz, 2003), slip bond behaviour describes an increasing probability of dissociation with increasing tensile force. Catch-slip bond behaviour, on the other hand, strengthens with increasing tensile force until some optimal force has been met. Furthermore, the rate at which this force is increased, known as the force gradient, also affects the strength of the selectin-ligand bonds (Evans et al., 2004). This is of particular interest to the study of stenotic arteries as plaque formation leads to distinct regions of varying force gradients.

2.3 Leukocyte adhesion in Parallel-Plate Flow Chambers (PPFCs)

The flow between two parallel plates has often been used to investigate the effects of blood flow on ECs and their interactions with blood components. Traditionally, parallel-plate flow chambers (PPFCs) have been used to provide an environment suited for tissue and suspension culture experiments under laminar flow. In classical PPFCs, fluid is driven through a channel formed by two narrowly separated plates in parallel. ECs are cultured on the bottom surface of the upper plate (often, a glass coverslip) while suspension cultures of leukocytes (neutrophils or monocytes) are prepared in the perfusion medium and their movement visualized within the chamber (Lawrence et al., 1987). This allows for regional or complete surface quantification of cells that are either adherent or, if observed in real-time, leukocytes that are undergoing rolling adhesion. A schematic of a PPFC is presented in Figure 2.

In a well defined PPFC, the WSS can be accurately characterised for steady, laminar flow of a Newtonian fluid as a function of a constant measurable volumetric flow rate Q:

$$\tau_w = \frac{6\mu Q}{wh^2} \tag{6}$$

where w is the width of the plate perpendicular to flow and h is the height of the interstitial gap between the two plates. For a constant volumetric flow rate, the velocity profile is parabolic and the WSS is uniform across the upper and lower plates save for the boundaries defined by the gasket (where the flow field approaches zero) and in the region of developing flow at the inlet of the chamber (Bacabac et al., 2005; Lawrence et al., 1987). In practice, PPFCs are designed with a large *w/h* ratio allowing most of the flow field to be homogenous over the surface of the cells (Bacabac et al., 2005).

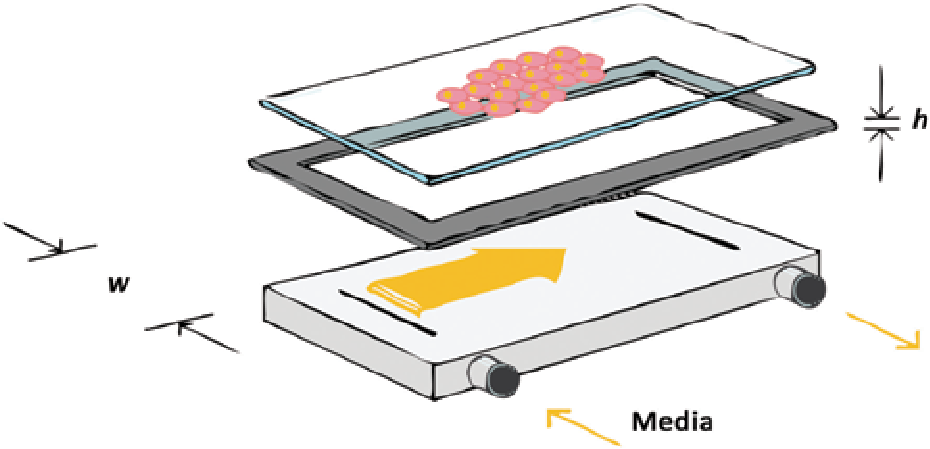


Fig. 2. Schematic of a Parallel-Plate Flow Chamber.

Studies of leukocyte adhesion using PPFCs have become the benchmark for revealing the role of shear in the adhesion pathway. Early results revealed a discreet shear dependence on both non-specific (Forrester & Lackie, 1984) and adhesion molecule-mediated adhesion (Alon et al., 1995; Finger et al., 1996; Lawrence et al., 1997). Further studies highlighted the role of endothelial dysfunction and inflammation as a precursor to adhesion when conditioned with flow (Alcaide et al., 2009; Sheikh et al., 2003; Sheikh et al., 2005). These findings complement the paradigm of leukocyte adhesion at sites of vascular inflammation; however, they do not address focal adhesion in non-uniform shear fields. This was considered by performing flow experiments using a step disturbance across the plate of the flow chamber (Burns & DePaola, 2005; Chen et al., 2006). The flow fields created by the step introduce regions of flow reversal and spatial WSS gradients to represent physiological hemodynamics (Burns & DePaola, 2005; Chen et al., 2006). Despite disturbed flow, leukocyte adhesion is increased in areas of high WSS gradients, with the highest incidence in re-attachment zones (Chen et al., 2006).

3. Asymmetric stenosis tissue culture model

3.1 Experimental methods

3.1.1 Asymmetric stenosis model design

Parallel-plate flow chambers are not ubiquitous when characterizing the role of WSS in endothelial dysfunction and leukocyte adhesion. Cone-plate viscometers (Shankaran & Neelamegham, 2001), animal models (Walpola et al., 1993, 1995) and three-dimensional (3D) tissue culture models (Hinds et al., 2001) have an increasing presence in the field.

A 3D model of an idealized coronary artery with an eccentric stenosis has been developed by our research group to reveal the effect of spatial WSS gradients on both endothelial inflammation and leukocyte adhesion (Rouleau et al., 2010a, 2010b). The eccentric stenosis geometry with a 50% occlusion (i.e. 50% area reduction, orthogonal to flow) has been chosen to represent a clinically relevant atherosclerotic lesion (Brunette et al., 2008; Wexler et al., 1996). The model measures 10 cm in length with an internal diameter of 3.175 mm, Figure 3. Sylgard™ 184 silicone elastomer is cast and cured in PVC moulds to create semi-compliant structures that maintain their geometric integrity through the stages of sterilization and cell culture preparation.

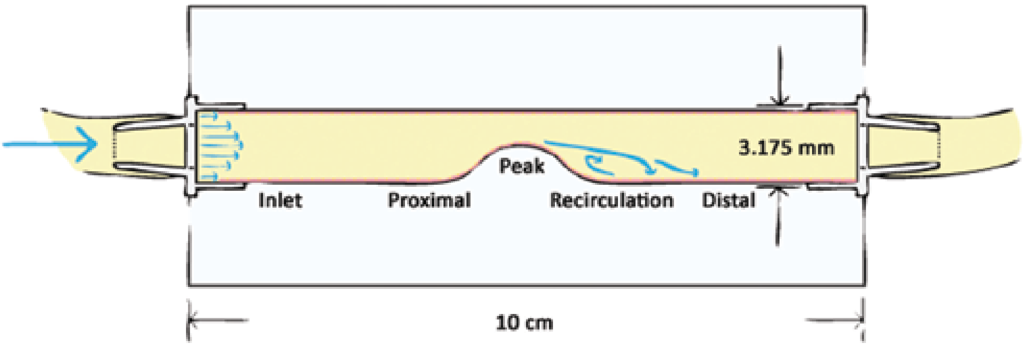


Fig. 3. Three-dimensional asymmetric stenosis model schematic with regional classifications defined using computational fluid dynamics and photochromic molecular flow visualization (Section 3.1.3).

3.1.2 Perfusion design

The perfusion flow loop consists of a media reservoir with tubing, flow dampeners, and an 8-roller peristaltic pump head with a programmable drive to produce steady or pulsatile laminar flow at the entrance of the models, Figure 4. ECs are cultured in the internal lumen of the model until they form a continuous, confluent monolayer. Inlet WSS values of 4.5, 9 and 18 dynes/cm² were chosen to represent moderate physiological shear in coronary arteries whereas inlet WSS values of 1.25 and 6.25 dynes/cm² were chosen to represent moderate to high shear for *in vitro* neutrophil adhesion. The WSS field was experimentally and numerically determined (Rouleau et al., 2010b).

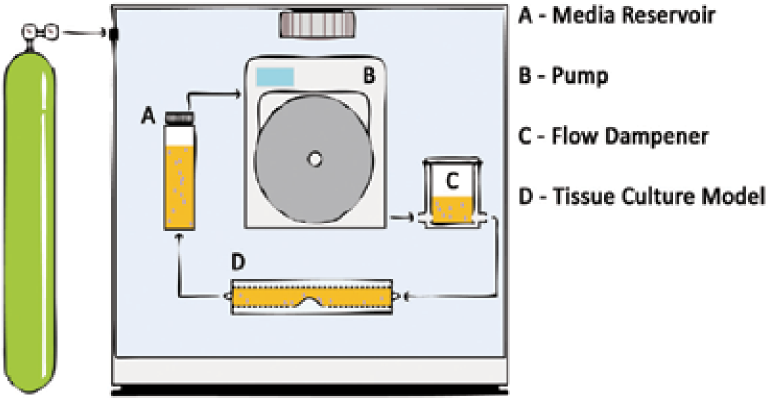


Fig. 4. Schematic of a steady-flow perfusion experiment with an asymmetric stenosis model.

3.1.3 Computational fluid dynamics and photochromic molecular flow visualization

Computational fluid dynamics (CFD) is a theoretical branch of research that relies on the power of modern computers to estimate fluid behaviour. The popularity of CFD lies in its ability to simulate physical experiments, thereby providing direction for further work or allowing for quick testing of key variables. With respect to cardiovascular flow studies, CFD is a numerical solution to a continuum of the Navier-Stokes (NS) equation. Although CFD is a powerful tool, it is only as accurate as the input data (e.g. geometry and mechanical properties). Defining the geometry and mechanics of healthy and diseased tissue is a constant endeavour in biomedical engineering (Choudhury et al., 2009; Tremblay et al., 2010). If the proper information is available, then CFD simulations are feasible, however experimental validation is still crucial. CFD has been performed for our stenosis model, yielding flow profiles at 6.25 dynes/cm^2 and 1.25 dynes/cm^2, respectively (Figures 5 & 6).

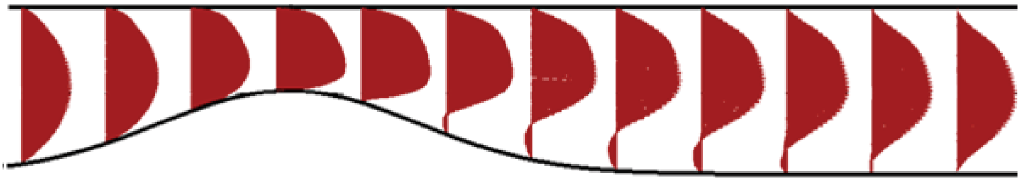


Fig. 5. Velocity profile, with an appreciable recirculation zone, in the asymmetric tissue culture model at 6.25 dynes/cm^2.

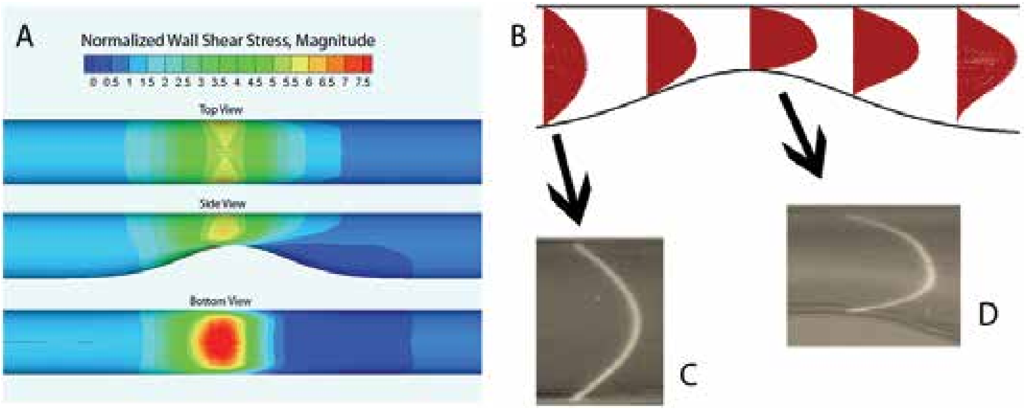


Fig. 6. Flow analysis of the asymmetric tissue culture model. (A) CFD normalized WSS contour plot. (B) Velocity profile at 1.25 dynes/cm^2. (C,D) PMFV velocity profiles at the inlet and peak, respectively.

We have used the photochromic molecular flow visualization (PMFV) technique to validate our CFD flow in the stenosis model (Ethier C.R. et al., 2000; Mahinpey et al., 2004). Using the flow profile at a given position, we are able to estimate the shear stress acting on the wall of a channel. Photochromic species reversibly change conformation when excited by a light source, such as a laser, resulting in an observable colour change. A PMFV setup includes a solution of photochromic dye, a laser, and a high-resolution camera. In practice, the photochromic dye solution is pumped through a micro-channel, the laser is triggered and the resulting pulse passes through the solution orthogonal to flow, activating any dye it contacts (Couch et al., 1996; Park et al., 1999). At this moment, a narrow column of visible dye will appear within the solution. The solution is in motion, however, so a fraction of a second later, the excited dye will have displaced with flow. This displacement is recorded by a camera, providing a snapshot of the flow profile and subsequently, the shear stress on the opposing walls. Photochromic visualization results are presented in Figure 6 (c) & (d) to validate our CFD simulation.

3.1.4 Cellular analysis

EC morphology is often a good predictor of EC phenotype. Healthy ECs elongate in the direction of flow, whereas dysfunctional ECs may become randomly oriented and cobblestone in appearance (Dartsch & Betz, 1989). The shape index (SI) is a metric of EC morphology defined as (Nerem et al., 1981):

$$SI = \frac{4\pi \cdot Area}{Perimeter^2} \tag{7}$$

Generally, elongated cells have a lower SI than rounded cells. In concert with the SI, the angle of orientation evaluates the proportion of EC elongation relative to the direction of flow. ECs in regions of observed elongation will be narrowly distributed near 0° (i.e. in the axis of flow) while regions that appear cobblestoned will have a much wider distribution.

Protein and mRNA regulation of inflammatory markers and transcription factors defines the endothelial phenotype and relates to leukocyte adhesion and atherogenesis. Gene regulation is quantified for large cultures using Q-PCR, while the resultant protein expression is observed using Western Blotting. Regional inflammation around the stenosis can be observed using immunostaining and confocal microscopy for adhesion molecules and the translocation of inflammatory transcription factors to EC nuclei. Our analysis includes, but is not limited to, specific inflammatory markers and adhesion molecules, including: ICAM-1, VCAM-1, E-selectin and NF-κB.

3.2 Results and discussion

3.2.1 Endothelial cell morphology

The stenosis model was first used to investigate the morphological effects of shear gradients caused by the stenosis. A morphological response is one of the last measurable changes which occurs in the cascade of events following introduction of flow. As ECs experience a steady WSS, they tend to become elongated and aligned in the direction of flow, representing a healthy endothelium. When exposed to low shear magnitude or WSS spatial gradients, ECs tend to take on a more cobblestone and random morphology which is indicative of an unhealthy endothelium (Helmlinger et al., 1991; Levesque et al., 1986; Levesque & Nerem, 1985; Nerem et al., 1981; Nerem, 1993).

Perfusion experiments to evaluate EC response were run at wall shear stress values of 4.5, 9 and 18 dyne/cm^2 which corresponded to Reynold's numbers of 50, 100 and 200, respectively. It was found that at all times and inlet WSS values the shape indices in the inlet and outlet of the stenosis model were statistically similar to that of a straight model. These results demonstrate that these values can be a good reference point for morphological changes in the regions surrounding the stenosis. Furthermore, the longer the perfusion time, the more elongation was observed in the direction of flow.

Effect of wall shear stress gradients on endothelial cell morphology

The stenosis model allows the observation of the morphological response of the ECs to WSS gradients. Figure 7 shows the WSS patterns within the model as a function of position. A positive shear gradient is found in the proximal region of the stenosis, reaching a peak WSS just upstream of the apex. A negative WSS gradient is observed in the recirculation zone of the stenosis, however, downstream of the flow reattachment point, laminar Hagen-Poiseuille flow resumes.

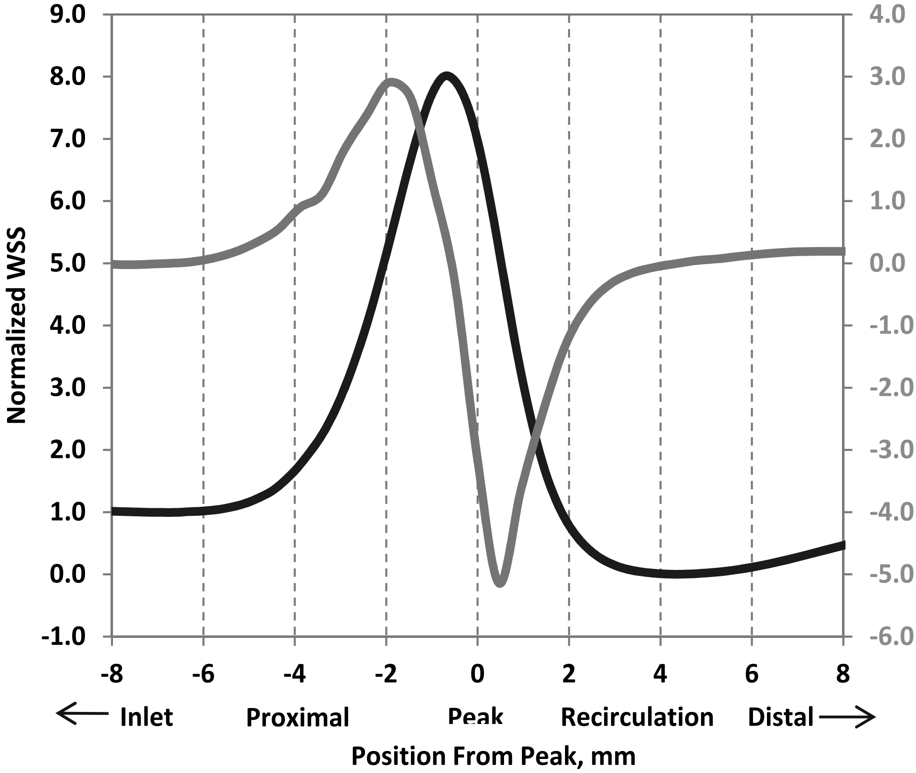


Fig. 7. CFD simulation of the WSS and WSS gradient Profile along the bottom central axis of the stenosis model (– Normalized WSS Magnitude; – Normalized WSS gradient).

It was found that after sufficient time had passed, morphological trends formed throughout the model. The inlet and outlet of the model showed similar shape indices as the straight tube controls, making them acceptable internal controls. These uniform internal controls can be compared to the other regions of the models. The deceleration, or recirculation zone (depending on the flow patterns, governed by the Re), showed the highest shape index, perhaps indicative of the most inflammatory response. The acceleration, or proximal zone, showed a slightly elevated shape index compared to the control regions, though this value was still statistically lower than that found in the deceleration zone. It was expected that the increased shear in this zone would result in more elongated ECs, however the results suggest that WSS gradients can have a more drastic effect on endothelium health than WSS magnitude alone.

It can be concluded that the deceleration zone could potentially present an inflamed endothelium and therefore one would predict to see an increase in regional expression of proteins linked to inflammation in that region. In turn, this should also lead to the largest amount of neutrophil adhesion in the recirculation zone.

3.2.2 Regional inflammation and adhesion molecule expression

WSS magnitude and duration was investigated to determine the effect these factors have on inflammatory response. Straight, cylindrical models were perfused under 4.5, 9 and 18 dyne/cm^2 inlet WSS conditions for various time periods (up to 24 hours). VCAM-1 and ICAM-1 mRNA expression decreased with increasing WSS magnitude and time, indicative of an atheroprotective phenotype. An increase in WSS magnitude resulted in a decrease in E-selectin mRNA expression; however, E-selectin expression increased from 0-12 hours, and sharply fell by 24 hours.

Inflammation was considered in the stenosis model by observing regional endothelial CAM expression using immunostaining and confocal microscopy. Perfusions with neutrophils were run at 1.25 dyne/cm^2, revealing an increase in CAM expression at the peak of the stenosis, Figure 8 (Rouleau et al., 2010a). Furthermore, perfusions were run at a higher inlet WSS of 6.25 dyne/cm^2, with no noticeable difference in regional CAM expression. These WSS values are consistent with the conditions used for neutrophil adhesion experiments.

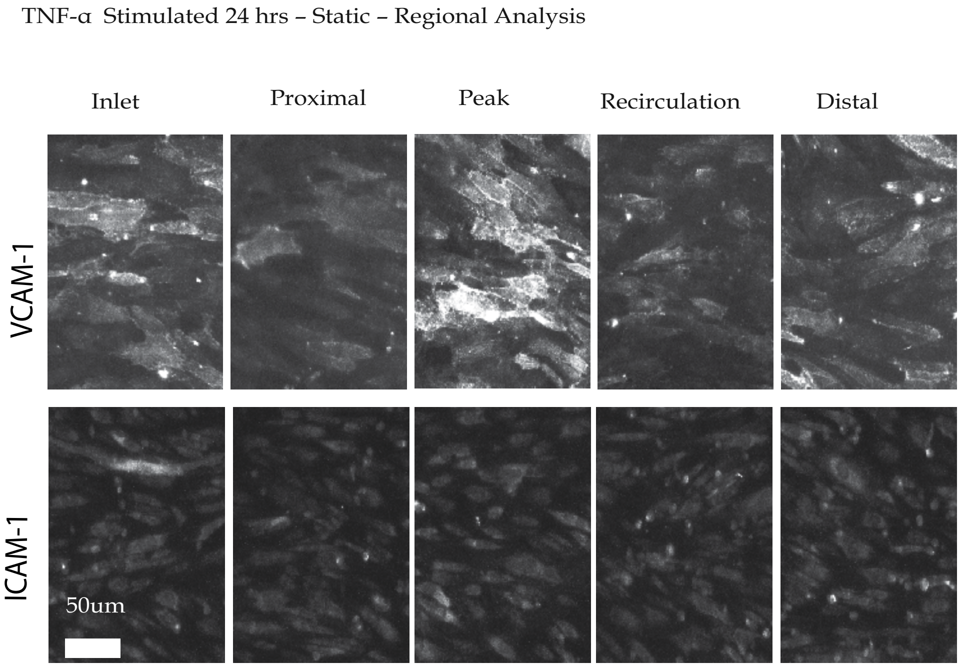


Fig. 8. Regional ICAM and VCAM expression. Copyright Springer, Annals of Biomedical Engineering, 38, 2010, pp. 2797, Neutrophil Adhesion on Endothelial Cells in a Novel Asymmetric Stenosis Model: Effect of Wall Shear Stress Gradients. Rouleau, L.; Copland, I; Tardif, J-C.; Mongrain, R. & Leask, R., Figure 6 with kind permission from Springer Science+Business Media B.V.

Using inlet WSS values of 4.5, 9 and 18 dyne/cm^2, more physiologically relevant hemodynamics were present. During these perfusions, ICAM-1, VCAM-1 and E-selectin levels were quantified and NF-κB translocation was observed to provide a robust picture of the regional inflammation around the stenosis. Similar to the experiments at 1.25 and 6.25 dyne/cm^2, there was an upregulation of ICAM-1 and VCAM-1 at the stenosis peak and in the proximal and recirculation zones. Similarly, E-selectin and NF-κB were also upregulated in these areas. It can then be concluded that there is a higher likelihood of increased neutrophil adhesion around the stenosis of the model.

3.2.3 Regional neutrophil adhesion

In perfusion experiments, it was shown that both WSS magnitude and perfusion time significantly affected the adhesion of a leukocyte cell line (NB4 cells). The trends showed that flow conditioned cells resulted in reduced adhesion of the NB4 cells for both the TNF-α stimulated and non-stimulated ECs. The two WSS magnitudes investigated, 1.25 dynes/cm^2 and 6.25 dynes/cm^2, resulted in a 3 fold and 15 fold decrease in adhesion, when compared to cells that were kept static prior to the adhesion experiments, respectively. Furthermore, experiments were run under even higher WSS conditions (12.5 dynes/cm^2) and it was found that very few cells were able to adhere to the ECs. This data demonstrates the influence of hemodynamic and attractive (ligand-receptor) forces acting on the neutrophils. The higher hemodynamic forces push the neutrophils off of the binding sites, overcoming the attractive forces which form during the adhesion of the neutrophils. For both the low (1.25 dynes/cm^2) and high (6.25 dynes/cm^2) shear stress conditions it was found that an increase in perfusion time from 1 to 6 hours resulted in an increase in adhesion, with a more noticeable increase occurring in non-stimulated ECs, potentially showing that at the shorter time point in TNF-α stimulated cells, a maximum *in vitro* adhesion is reached.

Regional neutrophil adhesion

The stenosis model presents a unique 3D environment which allowed for the investigation of the spatial differences in adhesion on and around a stenosis. Videos of the adhesion assays showed that there was a region of flow recirculation downstream of the stenosis. Immediately downstream of the separation point there was a distinct line of NB4 cell adhesion. It is postulated that this focal neutrophil adhesion was facilitated by low WSS and minimal fluid momentum caused by backflow in the recirculation zone.

For the rest of the analysis, the average adhesion was evaluated for each region of the model, Figure 9. TNF-α stimulation increased adhesion in all regions save for the stenosis peak. It was found that both the WSS magnitude and perfusion duration affected the incidence of adhesion. For example, it was found that at low inlet WSS (1.25 dyne/cm^2), ECs in the recirculation and distal regions showed a significant increase in adhesion from 1 to 6 hours.

It was found that in general, the recirculation zone tended to have the highest cell adhesion. It is hypothesized that the recirculation of NB4 cells results in a higher concentration of cells flowing along the endothelium (Rouleau et al., 2010a). Furthermore, the leukocytes have a lower momentum in the recirculation zone due to the decreased shear. This would allow for an increase in adhesion in this region. The endothelium in this location is also exposed to reduced WSS, leading to an increased inflammatory response.

The proximal and distal regions have lower incidence of adhesion than the recirculation region but more than at the stenosis peak. Interestingly, there seemed to be greater adhesion in the proximal region than the inlet of the model which may be due to a positive wall shear gradient. As the fluid reaches the stenosis, the projected surface area decreases resulting in an increase in WSS. Hinds et al. found a similar result in their studies using monocytes (Hinds et al., 2001). Comparing these two results, it can be seen that there is increased adhesion of leukocytes to ECs in the presence of complex wall shear stress gradients.

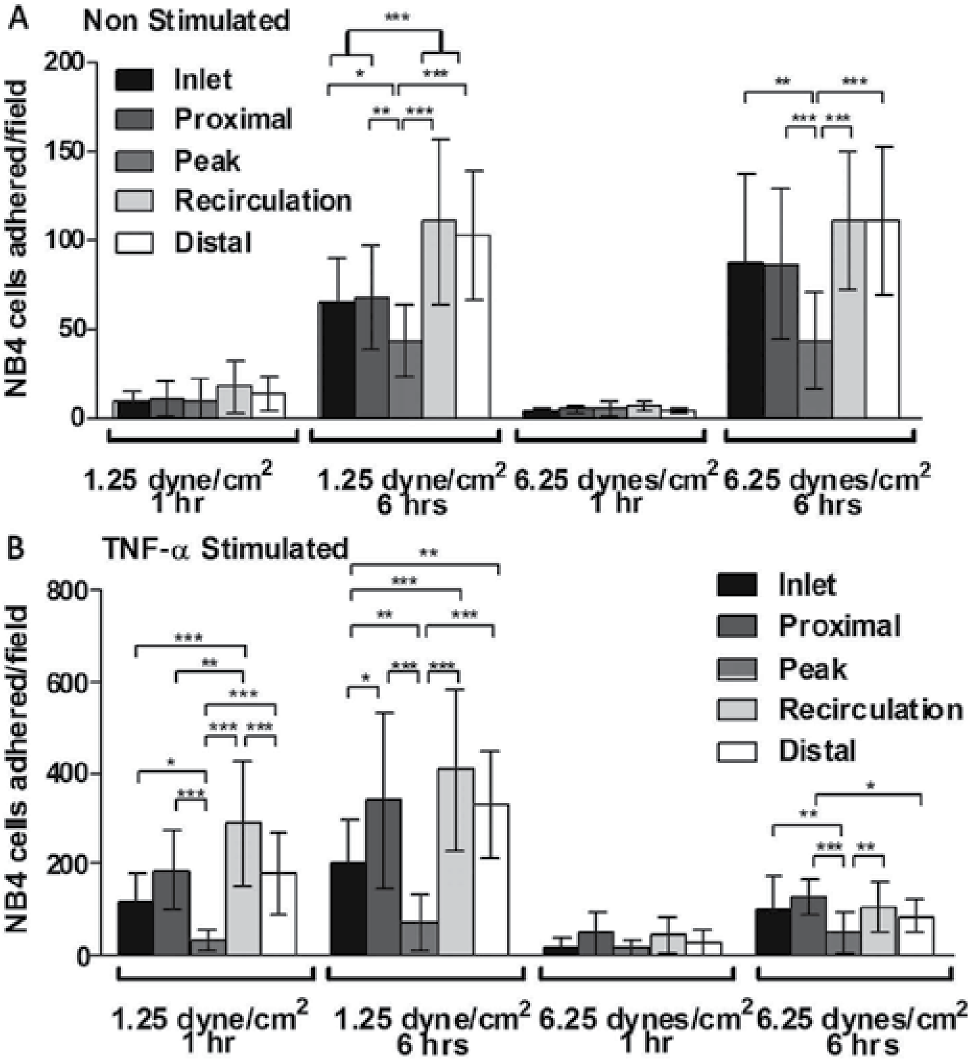


Fig. 9. Regional neutrophil (NB4) adhesion in the asymmetric tissue culture model. Copyright Springer, Annals of Biomedical Engineering, 38, 2010, pp. 2798, Neutrophil Adhesion on Endothelial Cells in a Novel Asymmetric Stenosis Model: Effect of Wall Shear Stress Gradients. Rouleau, L.; Copland, I; Tardif, J-C.; Mongrain, R. & Leask, R., Figure 7 (c) & (d) with kind permission from Springer Science+Business Media B.V.

In all instances, the stenosis peak had relatively low adhesion, which is a result of the high shear forces. By the peak of the stenosis, WSS values were appreciably larger than those found in the inlet (Figure 7). These hemodynamic forces exceed the adhesive force needed for the neutrophils to adhere. Extending static adhesion past 1 hour resulted in little or no additional adhesion. This was likely due to a lack of adhesion sites for the NB4s to bind to. Ultimately, under any conditions, there will be a point where the endothelium becomes saturated with bound NB4s and simply cannot facilitate further adhesion, *in vitro*. Although this leaves a certain limitation on the results of the aforementioned experiments, it also demonstrates the high levels of focal adhesion which can occur as a result of endothelial inflammation.

4. Conclusions

Atherosclerosis is an inflammatory disease. *In vitro* studies of the interaction of inflammatory cells with the endothelium have advanced our understanding of the role of inflammation in atherosclerosis development and progression. Our novel three dimensional dynamic cell culture model of a coronary stenosis has shown the importance of spatial gradients in wall shear stress in EC response and leukocyte attachment. Leukocyte attachment is increased in the proximal and distal regions of the stenosis. The increased attachment occurs in regions where the ECs have an inflamed phenotype. The results suggest that the hemodynamics created by the stenosis geometry create an inflammatory response of the endothelial cells that promotes leukocyte attachment. These results help to explain disease stability in established coronary stenoses.

5. References

Lloyd-Jones, D.; Adams, R.; Brown, T.; Carnethon, M.; Dai, S.; De, S.; Ferguson, T.; Ford, E.; Furie, K.; Gillespie, C.; Go, A.; Greenlund, K.; Haase, N.; Hailpern, S.; Ho, P.; Howard, V.; Kissela, B.; Kittner, S.; Lackland, D.; Lisabeth, L.; Marelli, A.; McDermott, M.; Meigs, J.; Mozaffarian, D.; Mussolino, M.; Nichol, G.; Roger, V.; Rosamond, W.; Sacco, R.; Sorlie, P.; Roger, V.; Thom, T.; Wasserthiel-Smoller, S.; Wong, N. & Wylie-Rosett, J. (2010). Heart disease and stroke statistics--2010 update: a report from the American Heart Association. *Circulation*, 121, 7, (February 2010), pp. (e46-e215)

Waller, B.; Orr, C.; Slack, J.; Pinkerton, C.; Van, T. & Peters, T. (1992). Anatomy, histology, and pathology of coronary arteries: a review relevant to new interventional and imaging techniques--Part I. *Clin Cardiol*, 15, 6, (June 1992), pp. (451-457)

Shah, P. (2006). Pathogenesis of Atherosclerosis, In: *Essential Cardiology*, C. Rosendorff, pp. (409-418), Humana Press, Retrieved from <http://dx.doi.org/10.1007/978-1-59259-918-9_22>

Libby, P. (2002). Inflammation in atherosclerosis. *Nature*, 420, 6917, (December 2002), pp. (868-874)

Van der Heiden, K.; Cuhlmann, S.; Luong, l.; Zakkar, M. & Evans, P. (2010). Role of nuclear factor kappaB in cardiovascular health and disease. *Clin Sci (Lond)*, 118, 10, (May 2010), pp. (593-605)

Schwartz, C.; Valente, A.; Sprague, E.; Kelley, J. & Nerem, R. (1991). The pathogenesis of atherosclerosis: an overview. *Clin Cardiol*, 14, 2 Suppl 1, (February 1991), pp. (I1-16)

Libby, P.; Ridker, P. & Maseri, A. (2002). Inflammation and Atherosclerosis. *Circulation*, 105, 9, (March 2002), pp. (1135-1143)

Ross, R. (1993). The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature*, 362, 6423, (April 1993), pp. (801-809)

Ross, R. (1999). Atherosclerosis--an inflammatory disease. *N Engl J Med*, 340, 2, (January 1999), pp. (115-126)

Ross, R.; Glomset, J. & Harker, L. (1977). Response to injury and atherogenesis. *Am J Pathol*, 86, 3, (March 1977), pp. (675-684)

Kelly, M.; Hwang, J. & Kubes, P. (2007). Modulating leukocyte recruitment in inflammation. *J Allergy Clin Immunol*, 120, 1, (2007), pp. (3-10)

Miyasaka, M.; Kawashima, H.; Korenaga, R. & Ando, J. (1997). Involvement of selectins in atherogenesis: a primary or secondary event? *Ann N Y Acad Sci*, 811, (April 1997), pp. (25-34)

Kubes, P. & Kerfoot, S. (2001). Leukocyte recruitment in the microcirculation: the rolling paradigm revisited. *News Physiol Sci*, 16, (April 2001), pp. (76-80)

Rao, R.; Yang, L.; Garcia-Cardena, G. & Luscinskas, F. (2007). Endothelial-dependent mechanisms of leukocyte recruitment to the vascular wall. *Circ Res*, 101, 3, (August 2007), pp. (234-247)

Iiyama, K.; Hajra, L.; Iiyama, M.; Li, H.; DiChiara, M.; Medoff, B. & Cybulsky, M. (1999). Patterns of vascular cell adhesion molecule-1 and intercellular adhesion molecule-1 expression in rabbit and mouse atherosclerotic lesions and at sites predisposed to lesion formation. *Circ Res*, 85, 2, (July 1999), pp. (199-207)

Schenkel, A.; Mamdouh, Z. & Muller, W. (2004). Locomotion of monocytes on endothelium is a critical step during extravasation. *Nat Immunol*, 5, 4, (April 2004), pp. (393-400)

Glagov, S.; Zarins, C.; Giddens, D. & Ku, D. (1988). Hemodynamics and atherosclerosis. Insights and perspectives gained from studies of human arteries. *Arch Pathol Lab Med*, 112, 10, (October 1988), pp. (1018-1031)

Nichols, W. & O'Rourke, M. (1990). McDonald's Blood Flow in Arteries; theoretical, experimental and clinical principles. Third Edition, (1990), pp. (1-456), Lea & Febiger, Philadelphia

Ku D. (1997). Blood flow in arteries. *Annual Review of Fluid Mechanics*, 29, 1, (1997), pp. (399-434)

Malek, A. M.; Alper, S. L. & Izumo, S. (1999). Hemodynamic shear stress and its role in atherosclerosis. *JAMA*, 282, 21, (December 1999), pp. (2035-2042)

Myers, J.; Moore, J.; Ojha, M.; Johnston, K. & Ethier, C. (2001). Factors influencing blood flow patterns in the human right coronary artery. *Ann Biomed Eng*, 29, 2, (February 2001), pp. (109-120)

Firrell, J. & Lipowsky, H. (1989). Leukocyte margination and deformation in mesenteric venules of rat. *Am J Physiol*, 256, 6 Pt 2, (June 1989), pp. (H1667-H1674)

Lawrence, M.; Kansas, G.; Kunkel, E. & Ley, K. (1997). Threshold levels of fluid shear promote leukocyte adhesion through selectins (CD62L,P,E). *J Cell Biol*, 136, 3, (February 1997), pp. (717-727)

Evans, E.; Leung, A.; Heinrich, V. & Zhu, C. (2004). Mechanical switching and coupling between two dissociation pathways in a P-selectin adhesion bond. *Proc Natl Acad Sci U S A*, 101, 31, (August 2004), pp. (11281-11286)

Cozens-Roberts, C.; Quinn, J. & Lauffenberger, D. (1990). Receptor-mediated adhesion phenomena. Model studies with the Radical-Flow Detachment Assay. *Biophys J*, 58, 1, (July 1990), pp. (107-125)

Tees, D. & Goetz, D. (2003). Leukocyte adhesion: an exquisite balance of hydrodynamic and molecular forces. *News Physiol Sci*, 18, (October 2003), pp. (186-190)

Marshall, B.; Long, M.; Piper, J.; Yago, T.; McEver, R. & Zhu, C. (2003). Direct observation of catch bonds involving cell-adhesion molecules. *Nature*, 423, 6936, (May 2003), pp. (190-193)

Sarangapani, K.; Yago, T.; Klopocki, A.; Lawrence, M.; Fieger, C.; Rosen, S.; McEver, R. & Zhu, C. (2004). Low force decelerates L-selectin dissociation from P-selectin glycoprotein ligand-1 and endoglycan. *Journal of Biological Chemistry*, 279, 3, (January 2004), pp. (2291-2298)

Lawrence, M.; McIntire, L. & Eskin, S. (1987). Effect of flow on polymorphonuclear leukocyte/endothelial cell adhesion. *Blood*, 70, 5, (November 1987), pp. (1284-1290)

Bacabac, R.; Smit, T.; Cowin, S.; Van Loon, J.; Nieuwstadt, F.; Heethaar, R. & Klein-Nulend, J. (2005). Dynamic shear stress in parallel-plate flow chambers. *J Biomech*, 38, 1, (January 2005), pp. (159-167)

Forrester, J. & Lackie, J. (1984). Adhesion of neutrophil leucocytes under conditions of flow. *J Cell Sci*, 70, (August 1984), pp. (93-110)

Finger, E.; Puri, K.; Alon, R.; Lawrence, M.; von Andrian, U. & Springer, T. (1996). Adhesion through L-selectin requires a threshold hydrodynamic shear. *Nature*, 379, 6562, (January 1996), pp. (266-269)

Alon, R.; Hammer, D. & Springer, T. (1995). Lifetime of the P-selectin-carbohydrate bond and its response to tensile force in hydrodynamic flow. *Nature*, 374, 6522, (April 1995), pp. (539-542)

Sheikh, S.; Rahman, M.; Gale, Z.; Luu, N.; Stone, P.; Matharu, N.; Rainger, G. & Nash, G. (2005). Differing mechanisms of leukocyte recruitment and sensitivity to conditioning by shear stress for endothelial cells treated with tumour necrosis factor-alpha or interleukin-1beta. *Br J Pharmacol*, 145, 8, (August 2005), pp. (1052-1061)

Sheikh, S.; Rainger, G.; Gale, Z.; Rahman, M. & Nash, G. (2003). Exposure to fluid shear stress modulates the ability of endothelial cells to recruit neutrophils in response to tumor necrosis factor-alpha: a basis for local variations in vascular sensitivity to inflammation. *Blood*, 102, 8, (October 2003), pp. (2828-2834)

Alcaide, P.; Auerbach, S. & Luscinskas, F. (2009). Neutrophil Recruitment under Shear Flow: It's All about Endothelial Cell Rings and Gaps. *Microcirculation*, 16, 1, (2009), pp. (43-57)

Chen, C. N.; Chang, S.; Lee, P.; Chang, K.; Chen, L.; Usami, S.; Chien, S. & Chiu, J. (2006). Neutrophils, lymphocytes, and monocytes exhibit diverse behaviors in transendothelial and subendothelial migrations under coculture with smooth muscle cells in disturbed flow. *Blood*, 107, 5, (March 2006), pp. (1933-1942)

Burns, M. & DePaola, N. (2005). Flow-conditioned HUVECs support clustered leukocyte adhesion by coexpressing ICAM-1 and E-selectin. *AJP - Heart and Circulatory Physiology*, 288, 1, (January 2005), pp. (H194-H204)

Shankaran, H. & Neelamegham, S. (2001). Effect of secondary flow on biological experiments in the cone-plate viscometer: methods for estimating collision frequency, wall shear stress and inter-particle interactions in non-linear flow. *Biorheology*, 38, 4, (2001), pp. (275-304)

Walpola, P.; Gotlieb, A.; Cybulsky, M. & Langille, B. (1995). Expression of ICAM-1 and VCAM-1 and monocyte adherence in arteries exposed to altered shear stress. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 15, 1, (January 1995), pp. (2-10)

Walpola, P.; Gotlieb, A. & Langille, B. (1993). Monocyte adhesion and changes in endothelial cell number, morphology, and F-actin distribution elicited by low shear stress in vivo. *Am J Pathol*, 142, 5, (May 1993), pp. (1392-1400)

Hinds, M.; Park, Y.; Jones, S.; Giddens, D. & Alevriadou, B. (2001). Local hemodynamics affect monocytic cell adhesion to a three-dimensional flow model coated with E-selectin. *J Biomech*, 34, 1, (January 2001), pp. (95-103)

Rouleau, L.; Copland, I.; Tardif, J.; Mongrain, R. & Leask, R. (2010). Neutrophil adhesion on endothelial cells in a novel asymmetric stenosis model: effect of wall shear stress gradients. *Ann Biomed Eng*, 38, 9, (September 2010), pp. (2791-2804)

Rouleau, L.; Farcas, M.; Tardif, J.; Mongrain, R. & Leask, R. (2010). Endothelial cell morphologic response to asymmetric stenosis hemodynamics: effects of spatial wall shear stress gradients. *J Biomech Eng*, 132, 8, (August 2010), pp. (081013-

Brunette, J.; Mongrain, R.; Laurier, J.; Galaz, R. & Tardif, J. (2008). 3D flow study in a mildly stenotic coronary artery phantom using a whole volume PIV method. *Medical Engineering & Physics*, 30, 9, (November 2008), pp. (1193-1200)

Wexler, L.; Brundage, B.; Crouse, J.; Detrano, R.; Fuster, V.; Maddahi, J.; Rumberger, J.; Stanford, W.; White, R. & Taubert, K. (1996). Coronary artery calcification: pathophysiology, epidemiology, imaging methods, and clinical implications. A statement for health professionals from the American Heart Association. Writing Group. *Circulation*, 94, 5, (September 1996), pp. (1175-1192)

Rouleau, L.; Rossi, J. & Leask, R. (2010). Concentration and time effects of dextran exposure on endothelial cell viability, attachment, and inflammatory marker expression in vitro. *Ann Biomed Eng*, 38, 4, (April 2010), pp. (1451-1462)

Choudhury, N.; Bouchot, O.; Rouleau, L.; Tremblay, D.; Cartier, R.; Butany, J.; Mongrain, R. & Leask, R. (2009). Local mechanical and structural properties of healthy and diseased human ascending aorta tissue. *Cardiovasc Pathol*, 18, 2, (March 2009), pp. (83-91)

Tremblay, D.; Cartier, R.; Mongrain, R. & Leask, R. (2010). Regional dependency of the vascular smooth muscle cell contribution to the mechanical properties of the pig ascending aortic tissue. *J Biomech*, 43, 12, (August 2010), pp. (2448-2451)

Ethier C.; Prakash, S.; Steinman, D.; Leask, R.; Couch, G. & Ojha, M. (2000). Steady Flow Separation Patterns in a 45 Degree Junction. *J Fluid Mechanics*, 411, (2000), pp. (1-38)

Mahinpey, N.; Leask, R.; Ojha, M.; Johnston, K. & Trass, O. (2004). Experimental study on local mass transfer in a simplified bifurcation model: potential role in atherosclerosis. *Ann Biomed Eng*, 32, 11, (November 2004), pp. (1504-1518)

Couch G.; Johnston, K. & Ojha, M. (1996). Full-field flow visualization and velocity measurement with a photochromic grid method. *Meas Sci Technol*, 7, (1996), pp. (1238-1246)

Park.H.; Moore J.; Trass O. & Ojha, M. (1999). Laser photochromic velocimetry estimation of the vorticity and pressure field-two dimenstional flow in a curved vessel. *Experiments in Fluids*, 26, (1999), pp. (55-62)

Dartsch, P. & Betz, E. (1989). Response of cultured endothelial cells to mechanical stimulation. *Basic Res Cardiol*, 84, 3, (May 1989), pp. (268-281)

Nerem, R.; Levesque, M. & Cornhill, J. (1981). Vascular endothelial morphology as an indicator of the pattern of blood flow. *J Biomech Eng*, 103, 3, (August 1981), pp. (172-176)

Levesque, M. & Nerem, R. (1985). The elongation and orientation of cultured endothelial cells in response to shear stress. *J Biomech Eng*, 107, 4, (November 1985), pp. (341-347)

Levesque, M.; Liepsch, D.; Moravec, S. & Nerem, R. (1986). Correlation of endothelial cell shape and wall shear stress in a stenosed dog aorta. *Arteriosclerosis*, 6, 2, (March 1986), pp. (220-229)

Helmlinger, G.; Geiger, R.; Schreck, S. & Nerem, R. (1991). Effects of pulsatile flow on cultured vascular endothelial cell morphology. *J Biomech Eng*, 113, 2, (May 1991), pp. (123-131)

Nerem, R. (1993). Hemodynamics and the vascular endothelium. *J Biomech Eng*, 115, 4B, (November 1993), pp. (510-514)

9

Pain in Osteoarthritis: Emerging Techniques and Technologies for Its Treatment

Kingsley Enohumah
The Rotunda Hospital, Dublin
Ireland

1. Introduction

Osteoarthritis (OA) is described as a condition characterised by use-related joint pain experienced on most days in any given month, for which no other cause is apparent. OA is the commonest disease affecting synovial joints and affects more than 40% of the population over 65 years. It affects primary the knee joints however hip, ankle, shoulder and small joints of hand and feet may be involved.

Previously, OA was considered a wear and tear, degenerative disease that must be accepted as an inevitable consequence of trauma and ageing. With advances in research and understanding of the mechanism of OA progression it is now known as a disease of the synovial joint affecting subchondral bone, synovium, meniscus, ligaments and supporting structures around the joints, including the cartilage.

The pathological changes seen in OA are characterised by focal areas of loss of articular cartilage within the synovial joints, associated with hypertrophy of the bone (osteophytes and subchondral sclerosis) and thickening of the capsule. OA is a chronic, degenerative disease associated with joint pain and loss of function. The primary problem in OA is the damage to the articular cartilage, which triggers a series of other events that culminate in pain and loss/limitation of function in the affected joint.

Undoubtedly, pain, which is the most prominent and disabling presentation of OA, is an increasingly important public health problem especially within an increasing aging population.

2. Epidermiology

OA occurs worldwide with higher prevalence in developed societies. Its twice as common in women as in men with a significant familial tendency.

The prevalence of OA increases with age in a progressive manner with 80% radiolographic changes in people by the age of 65 years. However only about 25-30% are symptomatic. Primary OA is uncommon before the age of 50 years.

World Health Organisation reports that knee OA is ranked fourth most important global cause of disability in women and the eighth most important in men. Annual arthroplasty rate in over the age of 65 in Europeans vary from country to country but are of the order of

0.5–0.7 per 1000. The annual costs attributable to knee OA are immense. There is therefore a burden on health from both morbidity and cost.
OA is a complex disorder with multiple risk factors.

2.1 Risk factors for OA

1. Age > 50 years
2. Crystals in joint fluid or cartilage
3. High bone mineral density
4. History of immobilisation
5. Injury to the joint
6. Joint hypermobility or instability
7. Obesity (weight-bearing joints)
8. Peripheral neuropathy
9. Prolonged occupational or sports stress

3 Anatomy of a joint

A brief review of the basic anatomy of a typical synovial joint is presented here to help understand the mechanisms involved in OA-induced damages of the involved joint that culminate in pain and other symptoms of OA.

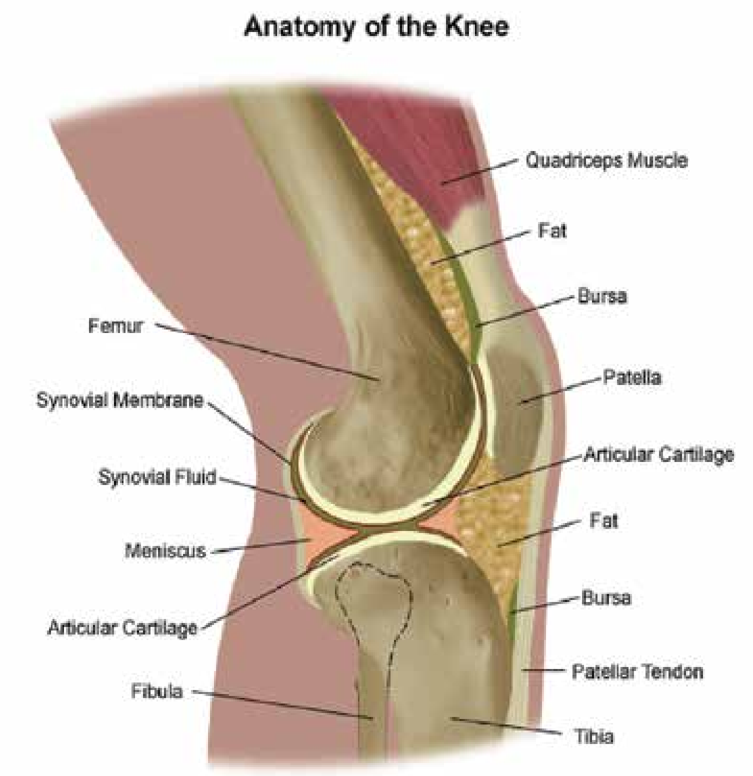

Anatomy of the Knee
Quadriceps Muscle
Fat
Bursa
Femur
Patella
Synovial Membrane
Articular Cartilage
Synovial Fluid
Fat
Meniscus
Bursa
Articular Cartilage
Patellar Tendon
Fibula
Tibia

Fig. 1. A typical synovial joint

A joint is where two bones meet. Articular cartilage covers the bone ends which are lubricated by synovial fluid. Seventy to eighty per cent of the cartilage is made up of water and a type II collagen with proteoglycans and glycosaminoglycans produced by chrondrocytes. The collagen fibres in the cartilage offer tensile strength to the cartilage because of its architectural makeup. The cartilage, however, contains no intrinsic blood vessels. It receives its nutrition from the synovial fluid. The synovial fluid, which is secreted by the synovial membrane lining the inner surface of the joint, facilitates not only movement but also provides nutrients, phagocytosis and other immunologic functions within the joint. The integrity of a joint is therefore dependent upon its architecture, the cartilage, bone and the supporting structures enclosing the joint. OA in simple terms is a result of alterations in the aforementioned architectural structures within the joint with resultant pain, loss of function and instability in the involved joint. Figure 1 shows the diagram of a typical synovial joint.

4. Sources of nociception in a joint

Pain is defined as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage”. Pain, as generally acknowledged, is mainly a signal that the body has been injured.

The term “nociception” was coined by the Nobel Laureate Sherrington to designate a physiological sensory phenomenon. “Nociception” is derived from “nocere”, the Latin word for “to hurt”. Nociceptors are peripheral sensory organs that are activated when nociceptive stimuli cause tissue damage. These nociceptors are unspecialised, naked nerve endings found close to small blood vessels and mast cells. The functional nociceptive unit is therefore made up of the structural triad of capillary, nociceptor and mast cell. This is the unit that is sensitive to tissue damage. There are also a rich supply of myelinated and unmyelinated fibres innervating the joint capsule, ligaments subchondral bone, periosteum and menisci.

In the anatomy of the joint described above, the cartilage does not contain blood vessels but derives its nutrients from the synovium. The subchondral bone, periosteum, synovium, ligaments, and the joint capsule contain nerve endings that could be the source of nociceptive stimuli in OA. Irritation of the periostal as a result of remodelling, denuded bone, compression of soft tissue by osteophytes, microfractures of the subchondral bone, effusion and spasm of surrounding muscles has been shown to contribute to the pain that may be felt by patients with OA. So in effect the bone in the periosteum and bone marrow is richly innervated with nociceptive fibres and represents a potential source of nociceptive pain in patients with OA.

5. Pathology and pathogenesis

OA is a heterogenous spectrum of clinical condition affecting mostly joints. No one mechanism explains the various processes seen in the joint of OA. Factors including inflammation, genetic, injury or trauma and joint mechanics have all been implicated in the pathophysiology of OA. Each joint response is a balance of the anabolic and catabolic factors acting in combination with both the extrinsic and intrisic factors.

Summary of the mechanisms suggested for the pathogenesis of OA are:

- *Matrix loss:* Metalloproteinases (MMPs) such as stromelysin and collagenase which are secreted by the chondrocytes catalyses the degradation of both collagen and proteoglycans resulting in matrix loss

- *Role of inflammatory mediators:* Mediators such as TNF-α and IL-1 stimulate MMPs secretion and this inhibit collagen production.
- *Tissue inhibitors of MMPs:* Tissue inhibitors of MMPs regulate the MMPs. Therefore any disturbance of this regulatory mechanism may lead to increased cartilage degradation and may contribute to the development of OA.
- *Growth factors deficiency:* Growth factors such as insulin-like growth factor and transforming growth factor enhance collagen synthesis and so when these factors are deficient matrix repair is impaired.
- Genetic susceptibility

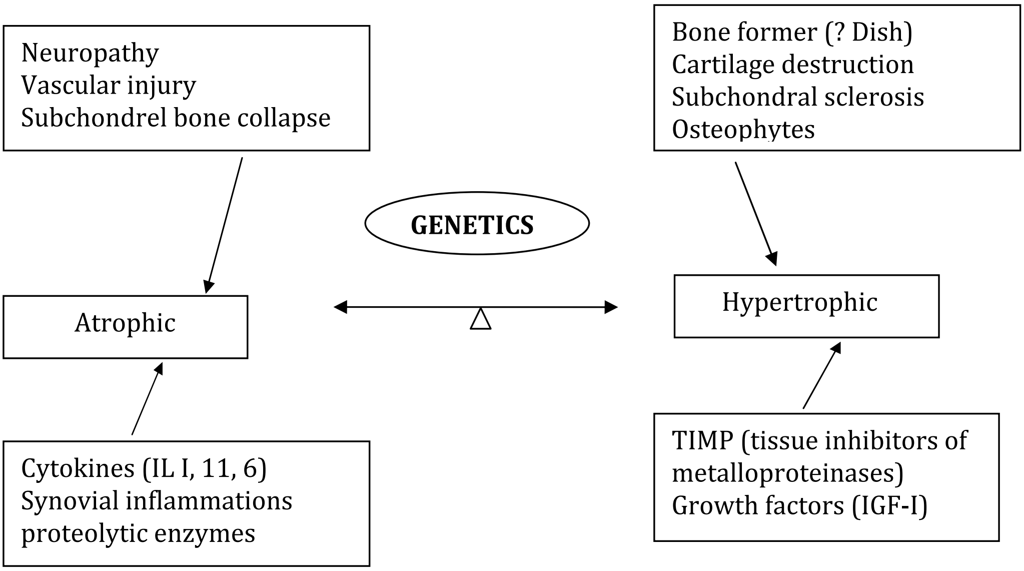

Neuropathy
Vascular injury
Subchondrel bone collapse
Bone former (? Dish)
Cartilage destruction
Subchondral sclerosis
Osteophytes
GENETICS
Atrophic
Hypertrophic
Cytokines (IL I, 11, 6)
Synovial inflammations
proteolytic enzymes
TIMP (tissue inhibitors of metalloproteinases)
Growth factors (IGF-I)

Fig. 2. Different factors that influence OA process.

Stage I	There is proteolytic breakdown of cartilage matrix
Stage II	There is fibrillation and erosion of cartilage surface, accompanied by the release of breakdown products into the synovial fluid
Stage III	Synovial inflammation begins when synovial cells ingest a breakdown product through phagocytosisand produce proteases and proinflammatory cytokines

Table 1. Stages of OA (Martel-Pelletier, 2004)

5.1 Classification of OA

Primary OA	Has no known cause. Common. Related to aging and hereditary. May be localised or generalised. Commonly affects the distal interphalangeal joints of the hands, hip and the knee. The cervical and lumbar spine may be affected.
Secondary OA	Causes include articular injury, obesity, Paget's disease, or inflammatory arthritis and aging process. May be localised or generalised. May affect any joint and can occur at any age.

5.2 Clinical features

Pain and functional restriction are the main symptoms in OA. The pain is characteristically made worse by movement and relieved by rest.

Signs	Symptoms
Joint tenderness	Joint pain
Crepitus on movement	Joint gelling (stiffening and pain after mobility)
Limitation of range of movement	Joint instability
Joint instability	Loss of function
Joint effusion and variable levels of inflammation	
Bone swelling	
Wasting of muscles	

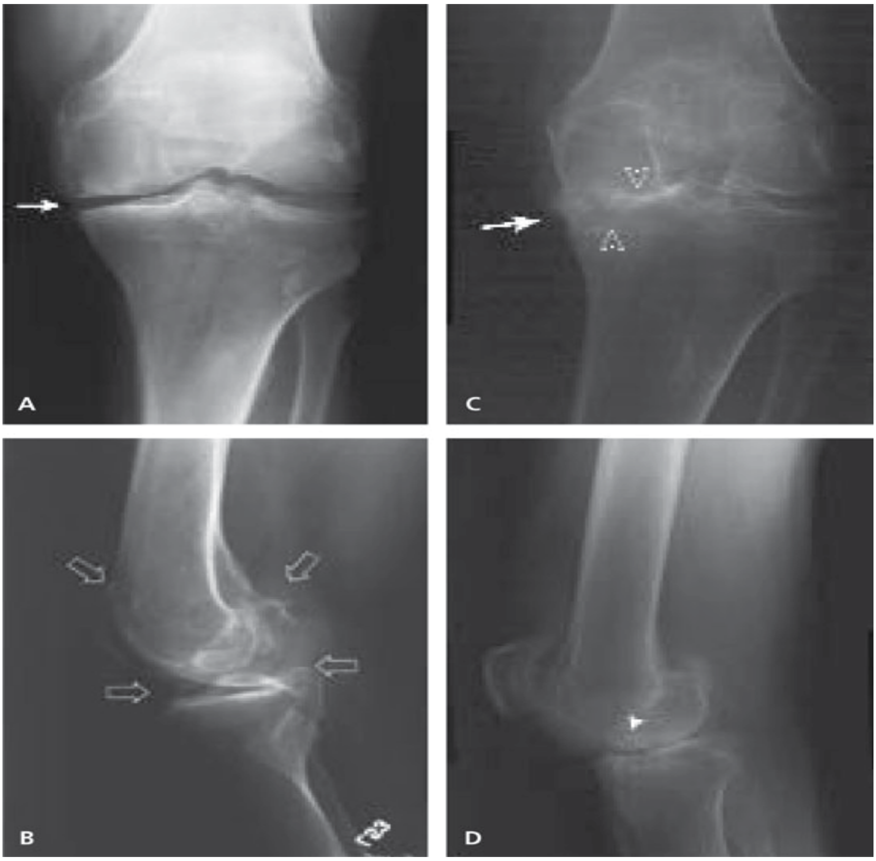


Fig. 3. Radiological changes in Osteoarthritis of the knees. (A) AP view of the left knee shows medial joint space narrowing (arrow). (B) Lateral view shows sclerosis with marked osteophyte formation (arrows). (C) Medial joint space narrowing (white arrow) causing a varus deformity of the knee and collapse of the joint space with destruction of the medial cartilage and the subchondral cortex (open arrow heads). (D) Subchondral cysts are noted (solid arrow head).

5.3 Radiological changes in OA

- Joint space narrowing
- Osteophytes
- Bony cysts
- Subchondral sclerosis

5.4 Mechanism of pain in OA

Summary of process of pain perception

- A noxious stimulus causes stimulation of nociceptors (pain receptors) in the receptor organ (e.g. joint).
- This firing of primary afferent fibres at the site of tissue injury causes axonal release of substance P (SP). This stimulation leads to activation of cells in the dorsal horn of the spinal cord and transmission of the nerve impulse to the midbrain and cortex. Thus, impulses travelling along first order neuron synapse on second-order neuron in the dorsal horn of the spinal cord. The axon crosses to the contralateral side and ascends to synapse on the third-order neurons. The third-order neurons send fibres to the cerebral cortex where conscious perception of the sensation occurs.
- Transmission of sensory information is modulated (inhibited or potentiated) throughout the nervous system by neurons from the midbrain and spinal cord that release endogenous opioids, catecholamines and other neurotransmitters.
- Peripheral nociceptor sensitisation, which is the transmission of impulses at subnormal threshold, occurs following the release of chemical mediators such as prostagladins and leukotrienes at the site of injury or damage. Continued stimulation by peripheral nociceptors then leads to sensitisation of neurons in the spinal cord. This is known as central sensitisation.

Tissue injury results in the release of inflammatory mediators such as serotonin, bradykinin, calcitonin gene-related peptide (CGRP) and SP, which lead to nociceptor nerve fibre sensitisation in peripheral tissue. These damaged fibres release inflammatory agents causing a spread of increased sensitivity around the area of tissue damage. This is called primary hyperalgesia. The repeated depolarisation of primary afferent fibres leads to a continuous release of neurotransmitters onto the secondary neurons in the spinal cord, resulting in central sensitisation and secondary hyperalgesia. Peripheral pain sensitisation is a feature of osteoarthritis in the joint.

In addition to peripheral pain sensitisation pain in OA, could also be due to local and central sensitisation of pain, pathways resulting in normal stimuli becoming painful with inflammation being an important feature in the process of OA.

Most of the substances involved in inflammation such as proinflammatory cytokines and bradykinins interact with the nociceptive fibres present within the joint and induce hyperalgesia and allodynia seen in patients with chronic inflammatory joint disease like OA. These mechanisms acting in concert could participate in the progression of hyperalgesia to chronicity.

5.5 Progression of OA to chronicity

Chronic pain (CP) is pain that persists for a month beyond the usual course of an acute disease or a reasonable time period for an injury to heal.

CP differs from the acute process not only in the duration of its course but also the different receptors involved in the mechanisms of action for acute pain and CP.

Those most involved in the acute process are a-amino-3-hydroxy- 5-methyl-isoxazole-4-propionic acid (AMPA) receptors, while those of primary importance in the sensation of CP are N-methyl- D-aspartate (NMDA) receptors. Activation of NMDA receptors causes the release of peptide neurotransmitter SP, which amplifies the pain by causing the spinal neurons carrying the pain to be easily stimulated.

Elevated levels of SP in spinal fluids have been documented in patients with OA and fibromyalgia. The progression of nociception from an acute to a chronic process has yet to be fully understood. However, recent evidence from animal experiments as well as human research suggests that peripheral mechanisms in acute pain and long-term potentiation (LTP) of neuronal sensitivity to nociceptive inputs in the dorsal horn of the spinal cord may underline the transition from acute to a chronic process.

LTP in spinal nociceptive systems has been suggested as one of the mechanisms underpinning the transition of acute pain to CP. It seems possible that LTP may underlie some forms of afferent induced hyperalgesia and that simultaneous activation of NMDA; SP neurokinin-I (NK-I) and glutamate receptors are required for the induction of spinal LTP. Therefore, it is likely that the conditioning stimuli that induce synaptic LTP in the superficial spinal dorsal horn are similar to those that trigger hyperalgesia. LTP is likely to occur in both the sensory and the affective pain pathways. Additionally, spinal LTP and injury-induced hyperalgesia share signal transduction pathways, which make use-dependent LTP an attractive model of injury-induced central sensitisation and hyperalgesia.

Summary of progression to chronic pain state

1. Rapid, intense stimulation of CA1 neurons in the hippocampus depolarizes them.
2. Binding of Glu and D-serine to their NMDA receptors opens them.
3. Ca^{2+} ions flow into the cell through the NMDA receptors and bind to calmodulin.
4. This activates calcium-calmodulin-dependent kinase II (CaMKII).
    - CaMKII phosphorylates AMPA receptors making them more permeable to the inflow of Na^{+} ions and thus increasing the sensitivity of the cell to depolarization.
    - In time CaMKII also increases the number of AMPA receptors at the synapse.
5. Increased gene expression (i.e., protein synthesis — perhaps of AMPA receptors) also occurs during the development of LTP.
6. Enlargement of the synaptic connections and perhaps the formation of additional synapses occur during the formation of LTP.

6. Treatment: emerging techniques and technologies

The modes of treatment for OA have always focus on decreasing pain and improving function ranging from information, education, physical therapy and aids, through analgesics, non-steroidal anti-inflammatory drugs and joint injections, and to surgery in which all or part of the joint is replaced with plastic, metal or ceramic implants.

OA is complex in genetics, pathogenesis, monitoring and treatment however, the principal goals of management are:

- Education of the patient about OA
- Pain relief
- Achieving and maintaining optimal joint and limb function
- Reducing adverse factors to beneficially modify the OA process and its outcome.

Despite huge laboratory and clinical research, there are no proven diseases modifying therapies for OA. However, emerging orthopaedic surgical procedures may help to alleviate the attendant pain and functional loss resulting from joint damage in OA.

Some of the surgical approaches in the management of OA include

- Arthroscopic approach
- Osteotomies
- Total joint replacements and arthrodesis
- Tissue engineering and biologic therapies
    - Autologous Chondrocyte Implantation (ACI)
    - Meniscal Transplantation (MT)

6.1 Arthroscopic procedures for OA

Arthroscopic surgery is a routine surgical procedure for joint debridement and lavage in the management of OA since the 1980s. The advent of this technique has permitted less invasive access to joints and the opportunity to intervene earlier in the course of joint destruction, potentially to delay and/or prevent a predictably progressive degenerative pathway. However, in recent times the only indication where this technique is thought to be of benefit is in the management of OA with a superimposed structural lesion such as a meniscal tear in which arthroscopic partial meniscectomy (APM) is performed simultaneously.

There is a strong research and clinical evidence that patients with symptoms attributable to knee OA per se, and not meniscal tear, do not improve following arthroscopic lavage and debridement. Whether APM is useful in patients with symptomatic meniscal tear and concomitant OA is unclear at this stage. This is an area of investigation at the moment.

6.2 Osteotomies

Osteotomies are performed to restore a more anatomic biomechanical environment and prevent or delay the onset of OA or slow its progression.

In symptomatic patients with OA, osteotomy is performed to realign joints with the aims of relieving pain and delaying the onset or progression of OA. Osteotomy and joint preserving surgical procedures should be considered in young adults with symptomatic OA, especially in the presence of dysplasia or varus/valgus deformity.

6.2.1 Indications

- As an adjunct in younger patients with predominantly unicompartmental OA
- Age less than 60 years
- 10 to 15 degrees of varus deformity on weight bearing radiographs
- Preoperative motion arc of at least 90 degrees
- Flexion contracture less than 15 degrees
- Ability and motivation to effectively and safely perform rehabilitation

6.2.2 Contraindications to osteotomy

- Lateral compartment loss of joint space
- Lateral tibial subluxation greater than 1 centimeter
- Medial bone loss greater than 2 to 3 millimeters
- Ligamentous instability
- Inflammatory arthritis

6.2.3 Surgical types and methods

Surgical Types	***Methods***
Medial compartment knee OA and varus deformity	High tibial osteotomy is performed either by removing a wedge of bone from the lateral proximal tibia or more commonly by opening wedge space in the a medial proximal tibia.
• Advantages	Permits the knee to adapt a more valgus alignment. Transfers load from the damaged medial compartment to the more normal cartilage of the lateral compartment.
Lateral compartment knee OA and varus deformity	This is a distal femoral osteotomy in which a wedge of bone is removed from the medial distal femur or a wedge is opened in the lateral aspect.
• Advantage	Shifts the load to the healthier medial compartment.

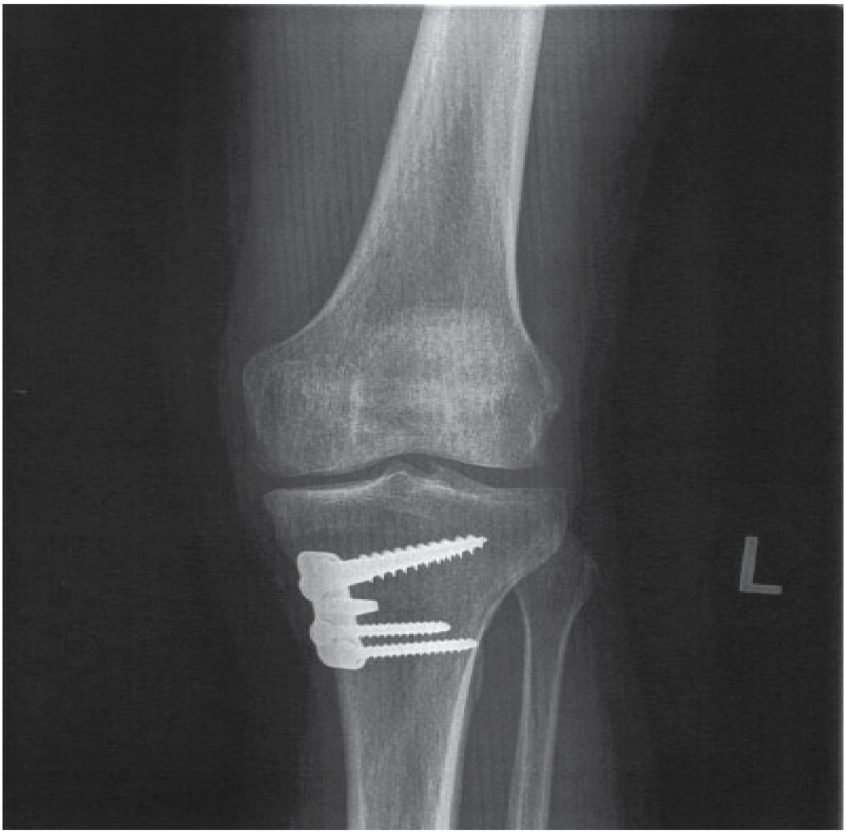


Fig. 4. AP radiograph of medial opening wedge high tibial osteotomy performed for medial compartment osteoarthritis.

6.3 Total joint replacements and arthrodesis

Total joint replacement (TJR) has to be considered in patients with radiographic evidence of hip/ knee OA who have refractory pain and disability. Principally, OA occurs less commonly at ankle, elbow, and wrist and thus total joint replacements are less frequent at these sites than at the hip or knee. In the last few years, interest in total knee arthroplasty has resulted in a proliferation of prosthetic designs, and many different types are now aaialble.

The indications for TJR have evolved and are expanding. Currently TJR are offered to patients earlier in the course of the disease as the risks of complications associated with TJR have reduced dramatically.

The prostheses available are:

1. *Condvlar replacements:* The joint surfaces alone are replaced. Ligaments then are needed to provide stability.
2. *Hinge-type prostheses:* In this type the ligaments are sacrificed and stability is provided by the design of the prosthesis itself.

The selection of a suitable prosthesis is dependent on the type and the indications

Types of prostheses

1. Unicondylar

This is an anatomically designed replacement for either the medial or the lateral femoral tibial articulation. It is designed to allow 120 degrees of flexion. The *unicondvlar* prosthesis is used only for cornpartmental OA.

2. Duocondylar

The femoral component of the duocondylar prosthesis is similar in shape to that of the unicondylar model except that there is no anterior flange and instead the halves are connected by an anterior cross bar which is countersunk during insertion. Because of its anatomical shape, it is most suitable when deformity, instability, and flexion contracture are not too severe.

3. Geometric

The prosthesis is non-anatomical in that the curvature of the femoral component is of constant radius. The plastic tibial component is in one piece, with two halves connected by an anterior bar. The prosthesis is designed to allow a 90- degree arc of motion. The cruciate ligaments are preserved. Two sizes are available.

4. Guepar

The Guepar is a Vitallium hinge prosthesis (improved over the Young model) which is fully constrained, providing motion in a fixed axis without rotation. Guepar prosthesis was used in knees with extrerne deformity or instability due to rheumatoid arthritis and OA.

Innovation continues to characterize the TJR field. This clinical dilemma has stimulated a search for biomaterials that produce less wear debris and, in turn, cause less osteolysis, attendant bone loss, and implant failure. This is the rationale for several developments, including highly cross-linked polyethylene and ceramic- on-ceramic and metal-on-metal bearing surfaces.

6.4 Surgical and biologic procedures

Advances in tissue engineering and biologic therapy have led to a few limited successes. Perhaps the most notable is autologous chondrocyte implantation (ACI).

Indications

1. Age <50 years
2. Isolated cartilage defects typically greater than 3 cm^2 in size

This procedure attempts to repair a symptomatic cartilage defect (Figure 3A) through implantation of chondrocytes grown ex vivo from a small cartilage biopsy sample obtained from the patient in a staging arthroscopy. After debridement of any degenerated tissue in the defect, a patch material, either periosteum from the patient or a synthetic collagen membrane, is sutured over the defect to create a watertight chamber into which the chondrocyte suspension is injected (Figure 3B). The chondrocytes attach to the subchondral bone and produce cartilage matrix, eventually filling the defect with hyaline-like cartilage.

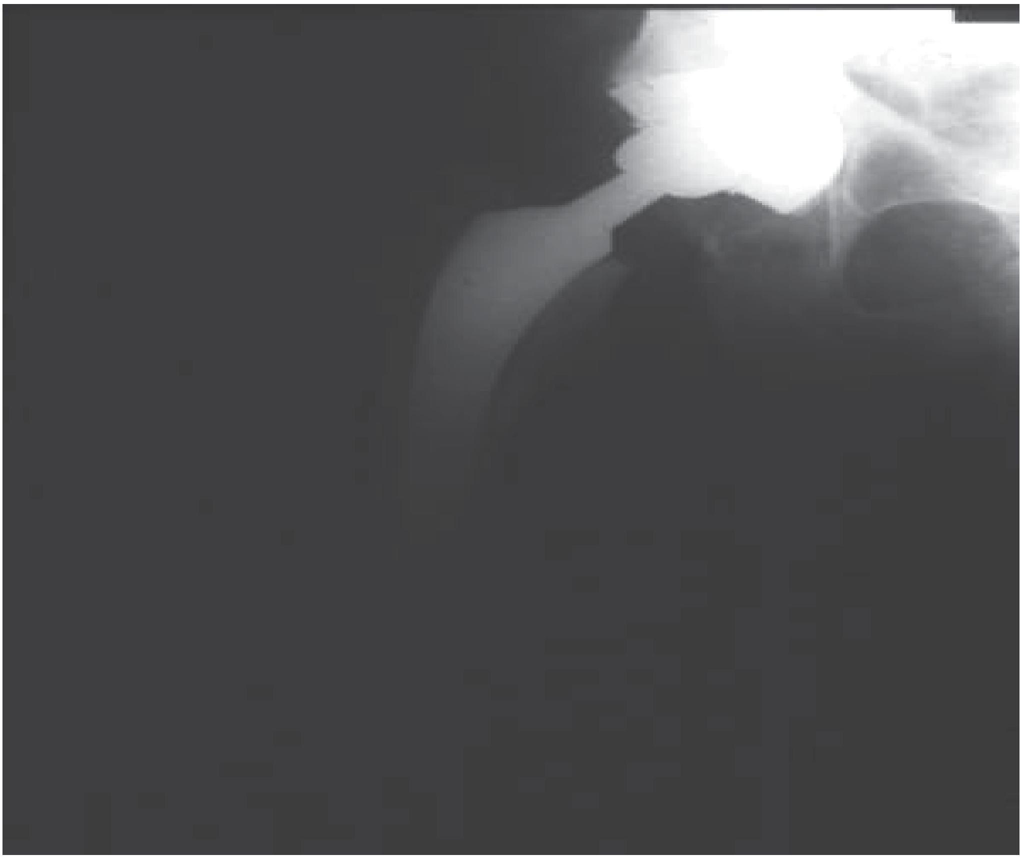


Fig. 5. Total hip replacement

6.5 Conclusion

It is clear from the foregoing that any simple unitary concept about the link between joint damage and symptoms in OA is untenable. We are faced with a complex interaction between local events in the joint, pain sensitisation, the cortical experience of pain, and what people are doing in their everyday lives.

In the absence of effective disease-modifying therapy, many patients with OA progress to advanced joint destruction. Therefore, surgery plays an important role in the management of OA. Advances in biomaterials and tissue engineering will continue to create exciting new opportunities to integrate surgical approaches in OA care.

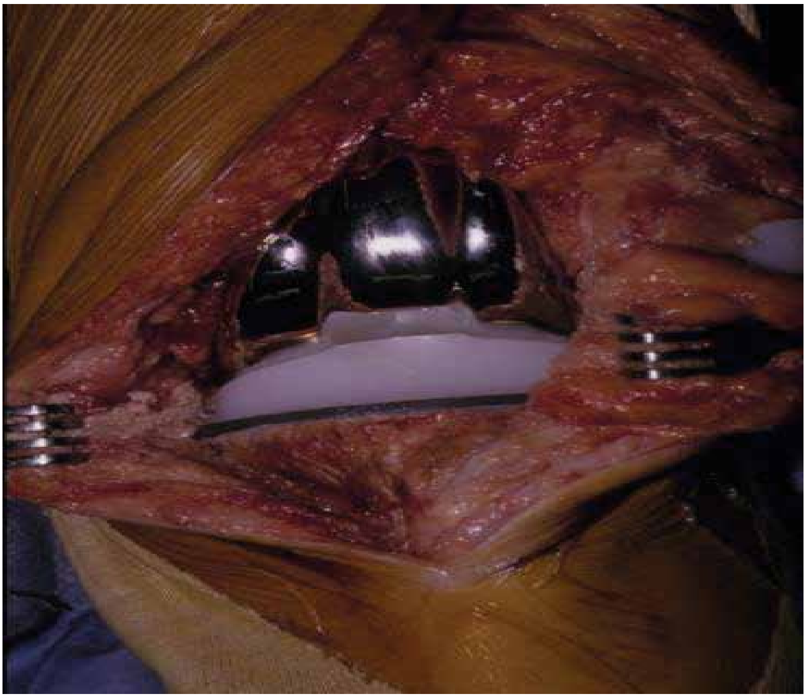


Fig. 6. Total knee replacement

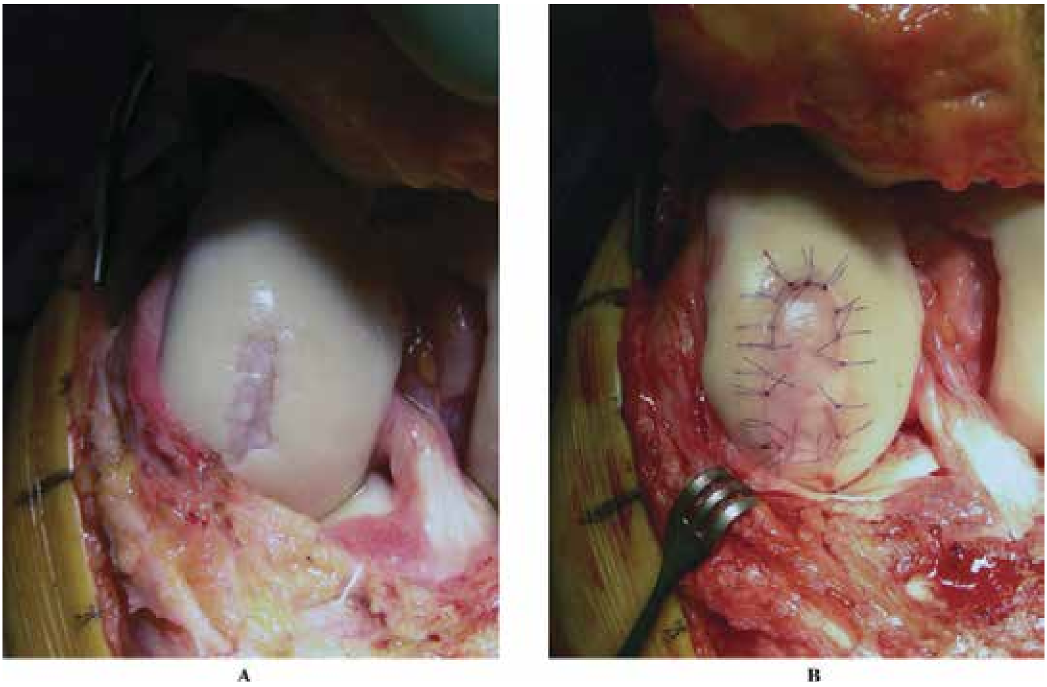


Fig. 7. **A**, Cartilage defect on femoral condyle.
B, Cartilage defect treated with autologous cartilage implantation

7. References

[1] Katz JN, Earp BE and Gomoll AH. Surgical management of Osteoarthritis. Arthritis Care & Research Vol. 62, No. 9, 2010, pp 1220–1228

[2] Haq I, Murphy E and Dacre J. Osteoarthritis. Postgrad Med J 2003; 79:377-383

[3] Kirkley A, Birmingham TB, Litchfield RB, Giffin JR, Willits KR, Wong CJ, et al. A randomized trial of arthroscopic surgery for osteoarthritis of the knee. N Engl J Med 2008; 359:1097–107.

[4] Felson DT, Lawrence RC, Dieppe PA, Hirsch R, Helmick CG, Jordan JM, et al. Osteoarthritis: new insights. Part 1: the disease and its risk factors. Ann Intern Med 2000; 133:635–46.

[5] W-Dahl A, Robertsson O, Lidgren L. Surgery for knee osteoarthritis in younger patients. Acta Orthop 2010; 81:161–4.

[6] Katz JN, Barrett J, Mahomed NN, Baron JA, Wright RJ, Losina E. Association between hospital and surgeon procedure volume and the outcomes of total knee replacement. J Bone Joint Surg Am 2004; 86A: 1909–16.

[7] Jacobs JJ, Urban RM, Hallab NJ, Skipor AK, Fischer A, Wimmer MA. Metal on-metal bearing surfaces. J Am Acad Orthop Surg 2009; 17:69–76.

[8] Knutsen G, Drogset JO, Engebretsen L, Grontvedt T, Isaksen V, Ludvigsen TC, et al. A randomized trial comparing autologous chondrocyte implantation with microfracture: findings at five years. J Bone Joint Surg Am 2007; 89:2105–12.

[9] Basad E, Ishaque B, Bachmann G, Sturz H, Steinmeyer J. Matrix-induced autologous chondrocyte implantation versus microfracture in the treatment of cartilage defects of the knee: a 2-year randomised study. Knee Surg Sports Traumatol Arthrosc 2010; 18:519–27.

[10] Agel J, et al. The burden of musculoskeletal conditions at the start of the new millennium. Geneva: World Health Organization; 2000.

[11] Rehman Q, Lane NE. Getting control of osteoarthritis pain: an update on treatment options. Postgrad Med 1999; 106:127–34.

[12] Murray CJL. The global burden of disease. Geneva: World Health Organization; 1996.

[13] Kenneth D.Osteoarthritis. Harrison's Principles of Internal Medicine 16th ed.2005.

[14] Romanes GJ. General introductions: Joints. In: Cunningham's Manual of practical anatomy. 15th ed. Vol (1). Oxford: Oxford University Press; 1986.

[15] Martini FH. Articulations: In: Fundamentals of anatomy and physiology. 4th ed. New Jersey: Prentice Hall; 1998.

[16] Snell RS. Basic anatomy. In: Clinical Anatomy. 7th ed. Place: Lippincott Williams & Wilkins; 2004

[17] Lim KKT, Shahid M and Sharif M. Recent advances in Osteoarthritis. Singapore Med J 1996; Vol 37: 189-193.

[18] Merskey H, Bogduk N. Classification of chronic pain: Description of chronic pain syndromes and definitions of pain terms. Seattle, WA: International Association for the Study of Pain Press; 1994.

[19] Sherrington CS. The integrative action of the nervous system. 2nd ed. New Haven: Yale University Press; 1947.

[20] Berne RM, Levy MN, editors. Physiology. 4th ed. St. Louis, MO: Mosby; 1998.

[21] Ganong WF. Review of medical physiology. 20th ed. Place: McGraw-Hill; 2001.

[22] Wyke B. The neurology of joints: a review of general principles. Clin Rheum Dis 1981; 57:233–9.

[23] Alvarez FJ, Fyffe RE. Nociceptors for the 21st century. Curr Rev Pain 2000; 4:451–8.

[24] Lawson SN. Phenotypes and function of somatic afferent nociceptive neurons with C, Adelta- or Aalpha/beta fibres. Exp Physiol 2002; 87:239–44.

[25] Wojtys EM, Beamann DN, Glover RA, et al. Innervation of the human knee joint by substance-P fibers. Arthroscopy 1990; 6:254–63.

[26] Kellgren JH, Samuel EP. The sensitivity and innervation of the articular capsule. J Bone Joint Surg Br 1950; 32:84–92.

[27] Dye SF, Chew MH. The use of scintigraphy to detect increased osseous metabolic activity about the knee. J Bone Joint Surg Am 1993; 75:1388–406.

[28] Creamer P, Hunts M, Dieppe P. Pain mechanisms in osteoarthritis of the knee: effect of intraarticular anesthetic. J Rheumatol 1996; 23:1031–36.

[29] Felson DT, Mclaughlin S, Goggins J, et al. Bone marrow edema and its relation to progression of knee osteoarthritis. Ann Intern Med 2003; 139:330–36

[30] Townes AS. Osteoarthritis. In: Barker LR, Burton JR, Zieve PD, editors. Principles of ambulatory medicine. 5th ed. Baltimore: Williams & Wilkins; 1999: 960–73.

[31] Jenkins BJ. Physiology of pain: Module 2 Lecture. In: PGCert/PGDip/MSc in Pain Management. Cardiff University, Wales; 1998.

[32] Farrel M, Gibson S, McMeeken J, et al. Pain and hyperalgesia in osteoarthritis of the hands. J Rheumatol 2000; 27:441–47

[33] Melzack R, et al. Central neuroplasticity and pathological pain. Ann NY Acad sci 2001; 933:157–74.

[34] Bonica JJ. History of pain concepts and pain therapy. Seminars in Anaesthesia 1985; 4:189–219.

[35] Brookoff D. Chronic pain: 1. A new disease? Hosp Pract 2000; 35:45–59.

[36] Brookoff D. Chronic pain: 1. A new disease? Hosp Pract 2000; 35:45–59.

[37] Bliss TV, Lømo T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J Physiol 1973; 232:331–56.

[38] Lømo T. The discovery of long-term potentiation. Philos Trans R Soc Lond B Biol Sci 2003; 358:617–20.

[39] Klein T, Magerl W, Hopf HC, et al. Perceptual correlates of nociceptive longterm potentiation and long-term depression in humans. J Neurosci 2004; 24:964–71.

[40] Sandku□hler J. Learning and memory in pain pathways. Pain 2000; 88:113–8.

[41] Sandku□hler J, Benrath J, Brechtel C, et al. Synaptic mechanisms of hyperalgesia. Prog Brain Res 2000; 129:81–100.

[42] Liu XG, Sandku□hler J. Activation of spinal NMDA or neurokinin receptors induces long-term potentiation of spinal C-fibre-evoked potentials. Neurosci 1998; 86:1209–16.

[43] Rosenberg JM, Harrell C, Ristic H, et al. The effect of gabapentin on neuropathic pain. Clin J Pain 1997; 13:251–5.

Part 2

Biomaterials and Implants

10

Non-Thermal Plasma Surface Modification of Biodegradable Polymers

N. De Geyter and R. Morent
Research Unit Plasma Technology – Department of Applied Physics,
Faculty of Engineering and Architecture – Ghent University,
Belgium

1. Introduction

Biodegradable polymers are often used as packaging materials. However, these polymers can also play an important role in tissue engineering as so called scaffolds (i.e. three-dimensional porous structures). The success of these biodegradable scaffolds strongly depends on the reaction of them with their surrounding biological environment. This reaction is mainly governed by the surface features of the scaffold and different approaches have already been tested to change the surface properties of biodegradable polymers. In particular, the research field on the use of non-thermal plasmas for a selective surface modification has known a steep rise. Therefore, this chapter will give an introductory and critical overview on recent achievements in plasma-assisted surface modification of biodegradable polymers. Firstly, we will discuss in short the most commonly biodegradable polymers. Secondly, we will go into more detail about surface modification by a non-thermal plasma and finally we will focus on some examples of plasma-treated biodegradable polymers.

2. Biodegradable polymers

2.1 Biomedical applications

A biodegradable polymer is defined as a polymer that preserves its mechanical strength and other material performances during its practical application, but that is finally degraded to low molecular weight compounds such as H_2O, CO_2 and other non-toxic by-products (Ikada & Tsuji, 2000). Next to their use as packaging material, which is not in the scope of this chapter, biodegradable polymers could play a key role in biomedical engineering for a variety of reasons (Ikada & Tsuji, 2000). Firstly, since the polymer degrades, it is clear that a device made of such a polymer can be implanted in the human body without necessitating a second surgery to remove the device (Athanasiou et al., 1998, Middleton & Tipton, 2000). Moreover, this prevented second operation makes the use of biodegradable polymers even more beneficial in other ways. For example, a fractured bone, fixated with a rigid, non-biodegradable stainless steel implant, has an inclination to fracture again when the implant is taken away because during the healing process the bone does not carry sufficient load, since the load is entirely intercepted by the rigid steel implant. This is in contrast to a biodegradable implant which degrades little by little and transfers by degrees the load from the implant to the fractured bone. This gradually movement of the load results in less bone re-fracture (Middleton & Tipton, 2000, Athanasiou et al., 1998).

Secondly, another interesting application field for biodegradable polymers is tissue engineering. This research branch aims to produce completely biocompatible tissues which could be employed to replace damaged or diseased tissues in reconstructive surgery (Djordjevic et al., 2008). Today's focus in this field is the use of so called scaffolds. These scaffolds are 3D artificial matrices that guarantee optimal support and conditions for growth of tissue (Djordjevic et al., 2008). Optimally, these scaffolds should fulfil the following two requirements (Ryu et al., 2005):

- being capable of supporting initial cell growth and further proliferation
- having the ability to degrade over time while leaving behind a reproduced functional tissue.

Finally, biodegradable polymers can also contribute in controlled drug delivery (Amass et al., 1998). A gradual delivery of antibiotics can be beneficial for the treatment of deep skeletal infections after a surgery, while the healing process of a fractured bone can be enhanced by a delivery in stages of bone morphogenetic proteins (Agrawal et al., 1995, Wang et al., 1990, Ramchandani & Robinson, 1998).

2.2 Overview of the most commonly used biodegradable polymers

Both natural and synthetic polymers have been extensively studied as biodegradable biomaterials (Nair & Laurencin, 2007). At first, natural polymers were considered as promising candidates. Nevertheless, several studies quickly demonstrated that some of these materials involve different drawbacks: possibility of disease transmission, strong immunogenic reactions and major difficulties with the purification process (Nair & Laurencin, 2007). In a later stage, synthetic biodegradable polymers were developed by the synthesis of polymers with hydrostatically unstable linkages in their backbones (Middleton & Tipton, 2000). These hydrostatically unstable groups are esters, ortho-esters, anhydrides and amines (Middleton & Tipton, 2000). The major advantages of synthetic over natural polymers are (Nair & Laurencin, 2007):

- synthetic biodegradable polymers are biologically inert
- they have more predictable properties
- their characteristics can be tailored with a specific application in mind

Various synthetic biodegradable polymers were invented of which the aliphatic polyesters appear to be the most attractive for biomedical applications. For this reason, this chapter will only discuss this type of polymers. An aliphatic polyester is a thermoplastic polymer which contains hydrolysable aliphatic ester linkage in its backbone (Donglu, 2010). In theory all polyesters are degradable, however only aliphatic polyesters with reasonably short aliphatic chains between the ester bonds will degrade within a time interval suitable for biomedical applications (Nair & Laurencin, 2007). Such aliphatic polyesters can be derived from a wide range of monomers through ring-opening and condensation polymerisation routes, but in some cases bacterial processes can also be used. In what follows we will briefly discuss the most common biodegradable aliphatic polyesters. The structural chemical formula of each discussed polymer is given in Table 1.

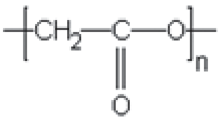
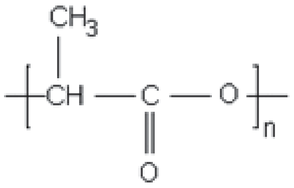
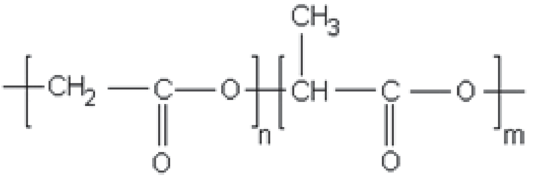
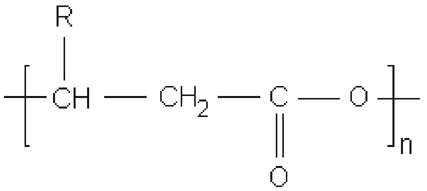
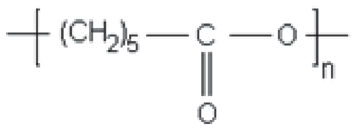
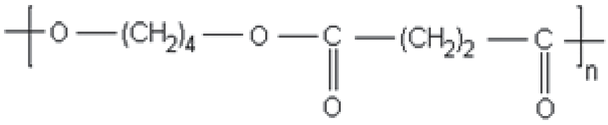
 $\left[CH_2-\underset{\underset{O}{\Vert}}{C}-O\right]_n$ polyglycolic acid (PGA)	 $\left[\overset{\overset{CH_3}{\vert}}{CH}-\underset{\underset{O}{\Vert}}{C}-O\right]_n$ polylactic acid (PLA)
 $\left[CH_2-\underset{\underset{O}{\Vert}}{C}-O\right]_n\left[\overset{\overset{CH_3}{\vert}}{CH}-\underset{\underset{O}{\Vert}}{C}-O\right]_m$ poly(lactic-co-glycolic acid)	
 $\left[\overset{\overset{R}{\vert}}{CH}-CH_2-\underset{\underset{O}{\Vert}}{C}-O\right]_n$ polyhydroxyalkanoates (PHA) R = CH_3: poly-3-hydroxybutyrate (PHB) R = CH_2CH_3: poly-3-hydroxyvalerate (PHV)	 $\left[(CH_2)_5-\underset{\underset{O}{\Vert}}{C}-O\right]_n$ polycaprolactone (PCL)
 $\left[O-(CH_2)_4-O-\underset{\underset{O}{\Vert}}{C}-(CH_2)_2-\underset{\underset{O}{\Vert}}{C}\right]_n$ polybutylene succinate (PBS)	

Table 1. Structural chemical formula of the most common biodegradable polymers.

2.2.1 Polyglycolic acid (PGA)

Polyglycolic acid (PGA) or polygycolide is the most simple linear biodegradable polymer and known as a rigid thermoplastic polymer (Vroman & Tighzert, 2009). Due to its high crystallinity (44-55 %) PGA provides excellent mechanical properties and it exhibits a low solubility in most organic solvents. Alow solubility (Nair & Laurencin, 2007). Nevertheless, it is soluble in highly-fluorinated solvents, such as hexafluoroisopropanol (Donglu, 2010). PGA has initially been studied for use as resorbable synthetic sutures and was for the first time commercialised in 1962 under the name DEXON® (Gilding & Reed, 1979). Today, due to their excellent biodegradability, good cell viability and good initial properties, PGA non-woven fabrics are also extensively utilized as scaffolds for tissue engineering (Nair & Laurencin, 2007).

PGA loses its strength in 1 to 2 months when hydrolyzed and its mass within 6 to 12 months (Nair & Laurencin, 2007). When inserted in the body, PGA breaks down into glycolic acid. Glycolic acid is not toxic and can be excreted in the urine or converted into H_2O and CO_2 and subsequently removed from the body via the respiratory system (Maurus & Kaeding, 2004). Despite the above-mentioned non-toxicity of glycolic acid, it may result in an increased and localized acid concentration leading to tissue dammage (Gunatillake & Adhikari, 2003, Taylor et al., 1994). This presents in particular problems for orthopaedic applications where implants with substantial dimensions are needed (Gunatillake & Adhikari, 2003). Together with the high degradation rate and low solubility, these acidic degradation products have hampered the use of PGA for biomedical engineering applications.

2.2.2 Polylactic acid (PLA)

The monomer building block of polylactic acid, lactic acid, is formed by converting sugar or starch from vegetable origin (e.g. wheat, corn, rice, etc.) via either bacterial fermentation or via a petrochemical process (Rasal et al., 2010). If PLA is implanted, it hydrolyses to its building block lactic acid which is a normal human metabolic by-product (Gunatillake & Adhikari, 2003). Lactic acid is degraded into H_2O and CO_2 which can be further removed by the respiratory system (Nair & Laurencin, 2007, Maurus & Kaeding, 2004). As can be seen in Table 1, lactic acid is a chiral molecule and therefore different forms of PLA occur. The two most important forms are poly(L-lactic acid) (PLLA) and poly(DL-lactic acid) (PDLLA).

Similar to PGA, PLLA has a high degree of crystallinity (± 37% depending on molecular weight and production processes) (Nair & Laurencin, 2007). Compared to PGA, PLLA slowly degrades: when PLLA is hydrolyzed, it loses its strength in circa 6 months. However, no mass loss is observed for a very long time and total degradation amount up to several years (Middleton & Tipton, 2000, Nair & Laurencin, 2007, Bergsma et al., 1995). Next to this slow degradation, PLLA offers good tensile strength, a high tensile modulus and low extension and can therefore be applied in load bearing applications like in orthopaedic fixation devices (Nair & Laurencin, 2007). PLLA fibres are also often used as surgical sutures, while PLLA composites, porous membranes or sponges can be employed as scaffolding matrices for tissue regeneration (Hu & Huang, 2010, Heino et al., 1996, Lam et al., 1995, Vaquette et al., 2008, Ma et al., 2006, Chen & Ma, 2004). For some other applications, the long degradation time of PLLA however presents a major concern.

PDLLA has an amorphous nature resulting in a substantial lower strength compared to PLLA (Nair & Laurencin, 2007). Moreover, PDLLA loses its strength in 1 to 2 months and its mass within 12 to 16 months (Maurus & Kaeding, 2004). Taking into account this low strength and its fast degradation rate, PDLLA can be employed as drug delivery system or as low strength scaffolding matrix for tissue engineering (Nair & Laurencin, 2007, Xie & Buschle-Diller, 2010).

2.2.3 Poly(lactic-co-glycolic acid) (PLGA)

A lot of research has been carried out on the development of a full range of poly(lactic-co-glycolic acid) (PLGA) polymers. This research has indicated that the degradation rate of PLGA strongly depends on the lactic acid/glycolic acid ratio (Gilding & Reed, 1979, Reed & Gilding, 1981, Miller et al., 1977). It is common knowledge that the intermediate co-polymers are much more unstable than the homo-polymers: a 50/50 PLGA and an 85/15 PLGA degrade in 1-2 months and 5-6 months respectively (Middleton & Tipton, 2000). This opportunity to tune the degradation rate of the polymer by varying the monomer ratio has made PLGA an ideal candidate for biomedical applications in the drug delivery and tissue engineering domain. However, the first commercial use of the co-polymer PLGA was as suture material under the name Vicryl ®(Nair & Laurencin, 2007, Gunatillake & Adhikari, 2003).

2.2.4 Polycaprolactone (PCL)

Polycaprolactone (PCL) is of great interest since it can be obtained from the relatively cheap monomer unit ε-caprolactone (Storey & Taylor, 1998). PCL degrades very slowly and complete degradation can take several years. Due to this slow degradation, its non-toxicity and its high permeability to small drug molecules, PCL has in the beginning been studied as a polymer for long-term drug delivery systems. PCL also offers excellent biocompatibility. Therefore, recently extensive research has been done on the use of PCL as scaffold matrices in tissue regeneration (Chiari et al., 2006, Mondrinos et al., 2006). Also several co-polymers have been developed to increase the degradation rate compared to pure PCL (Li et al., 2002, Li et al., 2003, Qian et al., 2000, Wang et al., 2001). For co-polymers synthesized from L-lactide and ε-caprolactone, the degradation rate was again strongly influenced by the L-lactide/ε-caprolactone ratio.

2.2.5 Polyhydroxyalkanoates (PHA)

Polyhydroxyalkanoates (PHA) are structurally related to PLA and are a polyester class derived from hydroxyalkanoic acids which can vary in chain length and in the hydroxyl group positions (Breulmann et al., 2009). As is the case for PLA, PHA can be obtained from renewable resources like starch, sugars or fatty acids, however, chemical transformation is not needed. The most widespread PHA is poly-3-hydroxybutyrate (PHB) which was discovered in 1920 as produced by the bacteria "Bacillus megaterium" (Nair & Laurencin, 2007). Subsequent research showed that PHB could also be synthesized via other bacterial strains and via chemical routes. **Subsequent research showed that PHB could also be synthesized via other bacterial strains and via chemical routes** (Shelton et al., 1971).

PHB degrades into D-3-hydroxybutyrate which is a normal element of human blood (Wang et al., 2001). To be used directly as biopolymer, PHB has the disadvantage of a very low degradation rate in the body compared with other biodegradable polyesters and is often considered too brittle for many applications (Nair & Laurencin, 2007, Pompe et al., 2007). Therefore, co-polymers of 3-hydroxybutyrate with other monomers such as 3-hydroxyvalerate have been synthesized (Nair & Laurencin, 2007). This poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHBV) is far less brittle and thus offers more potential as biomaterial (Nair & Laurencin, 2007, Ojumu et al., 2004). Moreover, PHBV is piezoelectric which enables electrical stimulation – known for promoting bone healing – of the implant (Nair & Laurencin, 2007). Although the faster degradation rate of PHBV compared to PHB, it has been observed that the in vivo degradation of both polymers remains slow. Therefore, these polymers may be potential candidates for long term implants.

2.2.6 Polybutylene succinate (PBS)

Polybutylene succinate (PBS) was discovered in 1990 and commercialized under the trade name Bionolle® (Fujimaki, 1998). PBS degrades via naturally occurring enzymes and micro-organisms into H_2O and CO_2 (Tserki et al., 2006). PBS can be easily produced in a wide variety of forms and structures, such as yarns, non-wovens, films, mono-filaments and it offers excellent mechanical properties comparable with polyethylene or polypropylene (Li et al., 2005, Vroman & Tighzert, 2009). These characteristics makes PBS an excellent choice for use as scaffolds in tissue regeneration.

3. Plasma-assisted surface modification of biodegradable polyesters

3.1 Introduction

Biodegradable polymers are non-toxic, possess low immunogenicity and good mechanical properties. Moreover, their degradation rate can be adjusted and therefore recently they have been extensively studied as scaffold matrices for tissue engineering (Shen et al., 2007). This research indicated that due to their hydrophobicity and their low surface energy cells only poorly attach, spread and proliferate on these biodegradable polyesters. Therefore, the surface of these polyesters should usually be modified and already several approaches have been presented to increase their cell affinity (Desmet et al., 2009). Typically the polyesters are chemically modified by introducing specific functional groups on their surface. Two possible wet-chemical routes are surface aminolysis and surface hydrolysis. Surface aminolysis in for example 1,6-hexanediamine leads to the production of free amino groups on the surface of the polyester which improves cell adhesion (Zhu et al., 2002, Zhu et al., 2004). By applying surface hydrolysis with the use of a NaOH solution, the ester group is hydrolyzed by the hydroxide anion leading to a rupture of the polymer chain and the formation of carboxylic acid and hydroxyl groups on the tail ends of the two new chains.

The presence of these groups results in an enhanced hydrophilicity and in improved cell-material interactions (Zhu et al., 2002, Zhu et al., 2004). Although these wet-chemical processes have their merit, some disadvantages cannot be neglected. Such surface modifications are quite rough and can thus possibly lead to unwanted side-effects such as a faster degradation rate and a reduction of mechanical performance (Chong et al., 2007, Desmet et al., 2009). In addition, research indicated that these techniques can lead to irregular surface etching and that the degree of modification strongly depends on molecular weight, crystallinity or tacticity and may therefore not be reproducible (Goddard & Hotchkiss, 2007, Desai & Singh, 2004). Moreover, it is clear that these wet-chemical techniques use substantial amounts of water or other liquids and consequentially generate hazardous chemical waste. Other approaches like peroxide oxidation, ozone oxidation, UV- and γ-radiaton can also introduce reactive chemical groups on the polyester surface, however, most of these techniques also lead to degradation of the polyesters (Ho et al., 2007, Koo & Jang, 2008, Loo et al., 2004, Place et al., 2009, Montanari et al., 1998).

Opposite those above-mentioned techniques, plasma-assisted surface modification offers a very suitable strategy to incorporate reactive functional groups on the polyester surface. Without the use of a solvent, these groups are efficiently introduced on the surface without altering the bulk properties of the polymer (Ho et al., 2006, Desmet et al., 2009). In addition, complex shaped scaffolds can be uniformly treated (Shen et al., 2007). Next to the incorporation of functional groups, plasma treatment can also be employed for the deposition of polymer coatings or for the immobilization of proteins or other biomolecules (Yang et al., 2002, Cheng & Teoh, 2004, Barry et al., 2005, Barry et al., 2006, Guerrouani et al., 2007, Zelzer et al., 2009).

Due to these numerous advantages, surface modification of biodegradable polymers by plasma treatment offers several excellent prospects. Therefore, in this section 3, we will give a more general introduction on plasma-surface interactions, while in section 4 we will focus on some successful examples of plasma modification of biodegradable aliphatic polyesters.

3.2 Plasma-surface interactions and surface modification strategies

Plasma is sometimes referred to as the fourth state of matter as introduced by Langmuir (Langmuir, 1928). Plasma is a partly ionized, but quasi-neutral gas in the form of gaseous or fluid-like mixtures of free electrons, ions and radicals, generally also containing neutral particles (atoms, molecules) (Denes & Manolache, 2004). Some of these particles may be excited and can return to their ground state by emission of a photon. The latter process is at least partially responsible for the luminosity of a typical plasma. In plasma several electrons are not bound to molecules or atoms, but free. Therefore, positive and negative charges can move somewhat independently from each other.

Plasmas are frequently subdivided into equilibrium (or non-thermal/low-temperature/cold) and non-equilibrium (or thermal/high-temperature/hot) plasmas (Denes & Manolache, 2004, Bogaerts et al., 2002, Fridman et al., 2008). Thermal equilibrium implies that the temperature of all particles (electrons, ions, neutrals and excited species) is the same. This is, for example true for stars, as well as for fusion plasmas. High temperatures are required to form these type of plasmas (Bogaerts et al., 2002, Lieberman & Lichtenberg, 2005). In contrast, plasmas with strong deflection from kinetic equilibrium have electron temperatures that are a lot more elevated than the temperature of the ions and neutrals. Such plasmas are classified as non-equilibrium or non-thermal plasmas. It is clear that the high temperatures used in thermal plasmas are destructive for heat-sensitive polymers and most applications for surface modification of polymers will make use of non-thermal or cold plasmas. Since a non-thermal plasma contains a mixture of reactive species, different interactions between the plasma and a surface are possible, including plasma treatment, plasma polymerization and plasma etching (Denes & Manolache, 2004, Rausher et al., 2010, Gomathi et al., 2008). These different interactions between a plasma and the surface can be divided into 4 different approaches to modify the biodegradable polymer. These 4 approaches will briefly be introduced in the following paragraphs, while section 4 will give some practical examples.

3.2.1 Plasma treatment

Plasma treatment is mostly used to enhance the surface energy of a polymer. Figure 1 shows the decrease in contact angle of a PLA surface after treatment in different discharge atmospheres. Oxygen or nitrogen containing groups are introduced on the surface of a (biodegradable) polymer when the material is exposed to a cold plasma generated in O_2, N_2, air or NH_3 (Morent et al., 2008a, Morent et al., 2008b). These functionalities are polar hydrophilic groups which are formed during the interaction of the plasma active species with the polymer molecules. Next to oxygen- and nitrogen-containing discharges, plasmas generated in pure helium or argon will lead to the creation of free radicals that can be used for cross-linking or grafting of oxygen-containing groups when the surface is exposed to oxygen or air after the treatment (Desmet et al., 2009, De Geyter et al., 2007, Ding et al., 2004). Finally, it should be mentioned that the induced surface characteristics are not permanent. The treated surfaces will tend to partially recover to their untreated state during storage in e.g. air (so-called hydrophobic recovery) and they will also undergo post-plasma oxidation reactions (De Geyter et al., 2008, Morent et al., 2010, Siow et al., 2006).

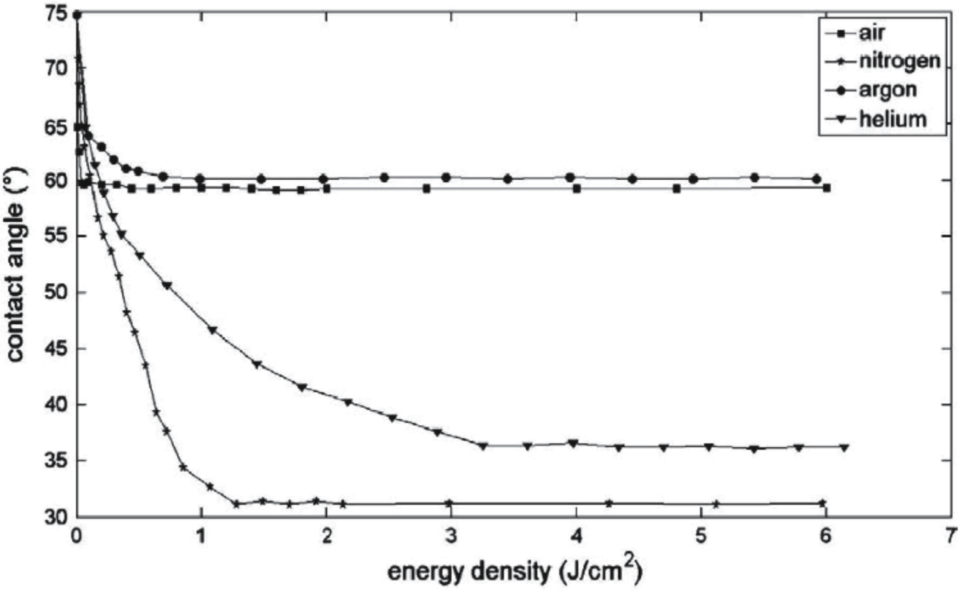


Fig. 1. Water contact angles as a function of energy density for air, nitrogen, argon and helium plasma-treated PLA samples. (Reprinted from (De Geyter et al., 2010) with permission of Elsevier).

3.2.2 Plasma post-irradiation grafting

Plasma post-irradiation grafting is a two-step process of which the first step is a plasma treatment as described in the previous paragraph. The induced functionalities can then be applied to initiate polymerization reactions (Desmet et al., 2009). In contrast to plasma treatment, this technique results in a permanent effect. In the second step, the activated polymer surface is brought into direct contact with a monomer. The monomer can be in the gas phase or the substrate can be immersed into a monomer solution (Vasilets et al., 1997, Zhu et al., 2007). It is important to notice that in both cases the monomer is not subject to the reactive plasma environment. Therefore, the grafted polymers will have similar characteristics as polymers synthesized by conventional polymerization processes (Desmet et al., 2009).

3.2.3 Plasma syn-irradiation

Firstly, a monomer is adsorbed to a material, after which the substrate is exposed to a plasma (Ding et al., 2004). This plasma will generate radicals in the adsorbed monomer layer and the surface of the substrate. This approach will lead to a cross-linked polymer top-layer (Desmet et al., 2009). Opposite to the plasma post-irradiation grafting described in the previous paragraph, the monomer is in plasma syn-irradiation directly subjected to the plasma.

3.2.4 Plasma polymerization

Thin films with unique chemical and physical properties can be developed by plasma polymerization and are called plasma polymers (Gomathi et al., 2008). During plasma polymerization, gaseous or liquid monomers are typically via a carrier gas inserted into the discharge zone in which they are converted into reactive fragments (Morent et al., 2009). These reactive fragments recombine to polymers and a polymer film is deposited on the substrate exposed to the plasma. The formed plasma polymers will not necessarily have the same chemical structure and composition as polymers obtained via conventional polymerization processes (Desmet et al., 2009). In general, plasma polymers are pinhole-free and highly cross-linked and are therefore insoluble, thermally stable, chemically inert and mechanically tough. Furthermore, such films are often highly coherent and adherent to a variety of substrates including conventional polymer, glass and metal surfaces (Morent et al., 2011, Morent et al., 2009).

4. Examples of plasma-assisted surface modification of biodegradable polyesters

In this last section, we will give some examples of the above-mentioned approaches of plasma-assisted surface modification of biodegradable aliphatic polyesters. Due to the introductory nature of this chapter, not all literature will be discussed in detail. For a more complete overview of literature, the reader could consult the review paper on the same subject of our research group (Morent et al., 2011). Plasma treatment as described in section 3.2.1 is by far the most occurring approach used to modify the surface of biodegradable polyesters and numerous examples can be found in literature. In this chapter, we will try to give some examples for all the biodegradable polyesters discussed in section 2. Therefore, section 4.1 will be much more extensive than the other sections (4.2 up to 4.4) since the availability of literature on these latter approaches is less pronounced.

4.1 Plasma treatment of biodegradable polymers

Hirotsu et al. published in 1997 one of the first studies on plasma modification of biodegradable polymers and treated PLA fabrics with a low pressure radio frequent (RF) discharge generated in pure oxygen and nitrogen (Hirotsu et al., 1997). The same group reported in 2002 about an enhancement of the wettability of PLLA sheets and showed a strong decrease in water contact angle from 80° to approximately 55° after 30 seconds of oxygen and helium plasma treatment (Hirotsu et al., 2002). They suggested that this increased wettability was only due to chemical changes of the surface, since pronounced etching is not likely to happen after such short treatment times. However, they were not able to determine the groups incorporated at the surface.

	C (at%)	O (at%)	N (at%)
no treatment	68	32	0
air plasma treatment	62	38	0

Table 2. Atomic composition of untreated and air plasma-treated PLA films (De Geyter et al., 2010).

To identify the functionalities incorporated at the surface, De Geyter et al. did detailed XPS studies on PLA sheets plasma-treated with a medium pressure dielectric barrier discharge (DBD) sustained in air. Table 2 shows the atomic composition of the PLA films plasma-treated in air. This table suggests that air plasma mainly adds oxygen atoms to the PLA surfaces. From high-resolution XPS scans, the authors have concluded that after plasma treatment in air, the concentration of C-O and O-C=O groups increases, while the C-C and C-H functional groups decrease. Hirtosu et al. observed a gradual increase in water contact angle when PLA samples were kept in dry air (Hirotsu et al., 2002). A similar hydrophobic recovery has recently been examined in detail by Morent et al., who employed a medium pressure DBD in different atmospheres for the surface modification of PLA. They concluded that during storage in air, the induced polar chemical groups reorientate or migrate to the bulk of the material (Morent et al., 2010). The introduction of specific functional groups on the surface of PLA samples and the accompanying increase in wettability often have the aim to improve the cell-material interactions. These interactions between B65 nervous tissue cells and oxygen plasma-treated PLLA films were studied by Khorasani et al. (Khorasani et al., 2008). Figure 2 shows optical photomicrographs of B65 cell attachment and growth on untreated and plasma-modified PLLA surfaces and it can clearly be observed that this oxygen plasma treatment substantially improves cell attachment and growth. The authors concluded that plasma-modified PLLA surfaces are very suitable for nervous tissue engineering purposes.

The majority of the above-discussed research is on flat 2D PLLA surfaces. However, from biomedical point of view, 3D porous polymer scaffolds are needed in the field of tissue engineering in order to offer sufficient support for tissue growth (Djordjevic et al., 2008). Only few authors have worked with 3D structures because of two reasons: (1) the insufficient knowledge on the penetration of plasma into porous structures and (2) the difficulty of characterisation of the interior surface with classical surface analytical tools. Wan et al. have modified 4 mm thick PLLA scaffolds with an ammonia plasma (Wan et al., 2006). To examine the plasma effect, they have immersed the scaffolds in blue ink after treatment to demonstrate the influence of treatment time on the modifying depth. Figure 3 shows that due to the poor hydrophilicity of the internal surface of the untreated PLLA sample only the most outside layer is dyed. However, with increasing treatment time, the ink increasingly penetrates the PLLA scaffold and after a treatment of half an hour the interior part of the scaffold is fully dyed.

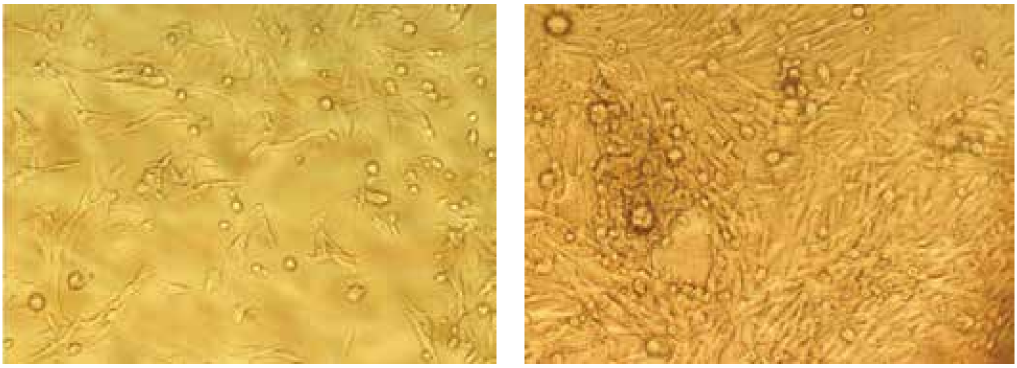

(a) (b)

Fig. 2. B65 attachment on (a) untreated PLLA and (b) oxygen plasma-treated PLLA (magnification 400x) (Reprinted from (Khorasani et al., 2008) with permission of Elsevier).

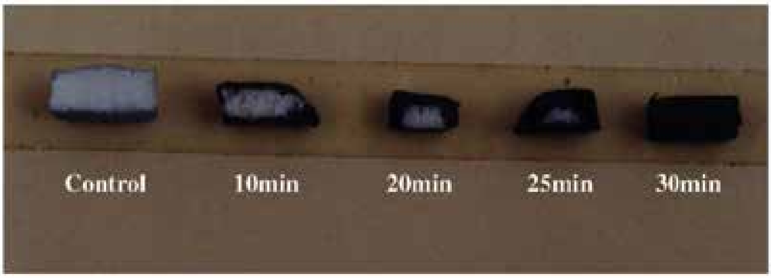


Fig. 3. Effect of plasma treatment time on the modifying depth of PLLA scaffolds (Reprinted from (Wan et al., 2006) with permission of Elsevier).

To our knowledge, no literature on plasma treatment of polyglycolic acid has been published so far. However, quite a few research articles deal with plasma modification of the copolymer PLGA (Khorasani et al., 2008, Hasirci et al., 2010, Khang et al., 2002, Park et al., 2007, Park et al., 2010, Safinia et al., 2007, Safinia et al., 2008, Shen et al., 2008, Wang et al., 2004). 50/50 PLGA films were modified in an oxygen plasma at low pressure and a decrease of the contact angle from 67° to below 40° after plasma treatment was observed. XPS revealed that oxygen containing functionalities are introduced and cell culture tests (3T3 fibroblasts) showed a higher cell attachment and proliferation on oxygen plasma-treated PLGA surfaces. As discussed in the previous paragraph on PLLA treated surfaces, Khorasani et al. also investigated in the same paper the interaction between nervous tissue cells and plasma modified PLGA samples (Khorasani et al., 2008). Figure 4 shows that oxygen plasma treatment clearly improves attachment and growth of B65 cells, however, the effect of oxygen plasma treatment seems less pronounced as was the case for PLLA surfaces (see Figure 2). Khang et al. studied and compared several modification methods including chemical methods (sulphuric acid, chloric acid, sodium hydroxide) as well as physical methods (atmospheric pressure air discharge) for the surface treatment of PLGA (Khang et al., 2002). Their results clearly evidenced that both chemical methods and plasma treatment could enhance cell attachment and growth. The high potential of non-thermal plasma for the surface modification of biodegradable polymers was clearly demonstrated since plasma treatment showed to be almost as efficient in increasing cell-material interactions as a chloric acid treatment and more efficient than sulphuric acid and sodium hydroxide treatments.

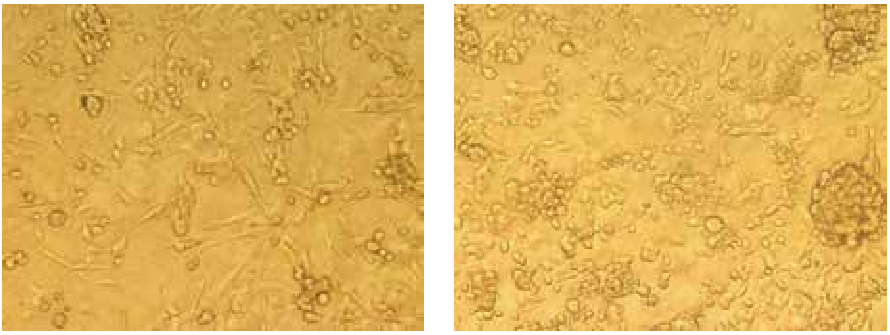
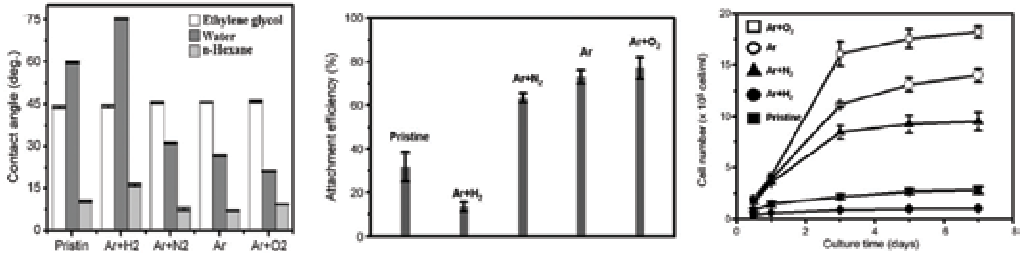

(a) (b)

Fig. 4. B65 cell attachment on (a) untreated PLGA and (b) oxygen plasma-treated PLGA (magnification 400x) (Reprinted from (Khorasani et al., 2008) with permission of Elsevier).

Surface modification of PCL with oxygen, helium and air plasmas has resulted into similar effects as on PLA and PLGA: an increased hydrophilicity, a higher oxygen amount and consequently an enhanced cell attachment and proliferation (Yildirim et al., 2008, Hirotsu et al., 2000a, Lee et al., 2009, Prabhakaran et al., 2008, Little et al., 2009). Lee and co-workers treated PCL with atmospheric pressure plasmas with different discharge gases (Lee et al., 2008). Figure 5 shows the enhancement in hydrophilicity (Figure 5 (a)), the increased cell attachment (Figure 5 (b)) and the increased cell proliferation (Figure 5 (c)).



(a) (b) (c)

Fig. 5. Plasma treatment of PCL: (a) contact angle (b) cell attachment and (c) cell proliferation of human epithelial cells (Reprinted from (Lee et al., 2008) with permission of Elsevier).

The most common polyhydroxyalkanoate subjected to plasma treatments is the co-polymer poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHBV) and oxygen plasmas have been widely employed to modify this co-polymer (Wang et al., 2006, Hasirci et al., 2003, Tezcaner et al., 2003, Kose et al., 2003b, Kose et al., 2003a, Ferreira et al., 2009). A low pressure oxygen plasma was employed to PHBV films containing 8% hydroxyvalerate in its structure by Hasirci et al. (Hasirci et al., 2003). A decrease in water contact angle upon oxygen plasma treatment was observed which was attributed to the incorporation of oxygen-containing functional groups on the PHBV surface. A subsequent study showed that O_2 plasma treatment significantly enhanced the interaction between retinal pigment epithelium (RPE) cells and PHBV (Tezcaner et al., 2003).

Hirotsu et al. published an interesting article on the plasma modification of self-made PBS sheets in different discharge atmospheres (O_2, N_2 and helium) (Hirotsu et al., 2000b). Contact angle measurements on the plasma-modified samples clearly showed that plasmas are able to increase the hydrophilicity, however, it was not stated which chemical groups contributed to this increased wettability.

4.2 Plasma post-irradiation grafting of biodegradable polyesters

Section 4.1 clearly focussed on the observation that plasma treatment can easily induce desired functionalities onto the surface of biodegradable polymers resulting in an improved cell affinity. However, hydrophobic recovery acts as a brake on practical applications of plasma-treated polyesters. Nevertheless, this drawback can be solved by covalently immobilizing bioactive molecules on plasma-treated surfaces (Gupta et al., 2002). Typically extracellular matrix (ECM) proteins such as gelatine, collagen or fibrin have been grafted on the surface of biodegradable polyesters since these proteins are known to enhance cell adhesion and proliferation (Ma et al., 2007). Different authors have studied the immobilization of collagen on PCL films (Ma et al., 2007, Chong et al., 2007, Cheng & Teoh, 2004). Firstly, an argon plasma is applied to a PCL film to generate radicals on the polyester surface. Exposure to the atmosphere for several minutes leads to the formation of functionalities such as surface peroxides and hydroperoxides that will be employed as initiator sites for UV-induced graft polymerization of acrylic acid. To pre-activate the carboxyl groups, the grafted substrates are immersed into a carbodiimide solution. In a final step, the material is immersed into a collagen solution leading to the production of a collagen-immobilized biodegradable polyester. These collagen-modified PCL surfaces have been tested with a diversity of cells including human dermal fibroblasts, human myoblasts, human endothelial cells and human smooth muscle cells and all demonstrated favourable response from these cells (Ma et al., 2007, Chong et al., 2007, Cheng & Teoh, 2004). Next to collagen, Kang et al. also immobilized insulin on the surface of a PHBV co-polymer and observed that the proliferation of human fibroblasts was significantly accelerated on these films compared to the untreated samples (Kang et al., 2001).

4.3 Plasma syn-irradiation of biodegradable polyesters

As described in section 3.2.3, a polymer can also be grafted on the surface of a biodegradable polyester by pre-adsorption of the monomer followed by a plasma treatment. However for the specific case of biodegradable polymers, we were able to track only one research paper using this plasma approach (Ding et al., 2004). In this paper, Ding et al. tried to modify the surface of PLLA films with a chitosan layer. However, results indicated poor cell adhesion, but acceptable cell proliferation.

4.4 Plasma polymerization on biodegradable polyesters

Plasma polymerization differs from plasma grafting in that respect that it coats the substrate rather than covalently binds species to a plasma-modified polymer surface (Barry et al., 2005). Allylamine is one of the most frequently used monomers to plasma polymerize on biodegradable polymers such as PLLA, PCL and PHBV (Barry et al., 2005, Guerrouani et al., 2007, Carlisle et al., 2000). Plasma polymerized allylamine films on biodegradable polymers resulted in highly hydrophilic surfaces with contact angles of 20° or lower due to the amine groups on the surface.

As there is a great interest in the surface modification of 3D implants, plasma polymerization has also been performed on the surface of 3D PLA scaffolds. Plasma grafting has been compared with plasma polymerization using allylamine (Barry et al., 2005). In the case of plasma grafting, the scaffolds were first pre-treated with an oxygen plasma and afterwards exposed to allylamine vapour, while plasma polymerization was carried out by exposing the scaffolds to an allylamine vapour plasma after an oxygen plasma pre-treatment. XPS measurements of the scaffolds at different points across the scaffold diameter demonstrated that the grafting process resulted in a more homogeneous nitrogen concentration through the scaffold while the concentration of nitrogen on the internal surface of the scaffold on which the plasma deposit was formed decreased from the edge to the core of the scaffold, as can be seen in Figure 6. However, at the lowest nitrogen concentration, the nitrogen concentration on the internal surface of the plasma-polymerized scaffold was still greater than that of the grafted surface. The plasma-coated scaffolds also showed a higher metabolic activity than the plasma-grafted samples. Moreover, fibroblasts were detected in the centre of the plasma-coated samples, which was not the case for the grafted scaffolds (Barry et al., 2005).

5. Conclusion

The growing research fields of tissue engineering and regenerative medicine are a leverage for surface engineering of biodegradable polymers. Next to chemical surface modification techniques which encounter problems with the use of hazardous organic solvents in relation to cell viability, non-thermal plasma technology knows a steep growth as solvent-free technique. Plasma treatments are already commonly performed on biodegradable polymers such as PLA and PLGA, while treatment of more advanced biodegradable polymers (such as PCL, PHBV, PBS and composites) and other plasma-based techniques (such as plasma grafting and plasma polymerization) are only at the verge of breaking through. Non-thermal plasma technology can greatly enhance cell-material interactions, however, a better understanding of these interactions is of crucial importance. This knowledge can provide us information on which plasma-based strategies should exactly be pursued.

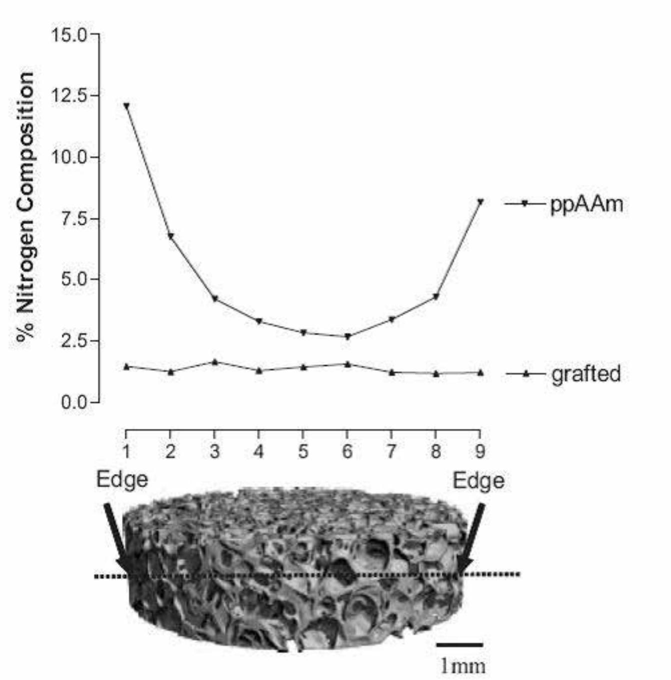


Fig. 6. Nitrogen concentration as determined by XPS at set points across the internal diameter of grafted and plasma-polymerized allylamine (ppAAm) scaffolds (Reprinted from (Barry et al., 2005) with permission from Wiley-VCH Verlag GmbH & Co. KGaA).

6. References

Agrawal, C. M., Best, J., Heckman, J. D. & Boyan, B. D. 1995. Protein Release Kinetics of A Biodegradable Implant for Fracture Non-Unions. *Biomaterials*, 16, 1255-1260.

Amass, W., Amass, A. & Tighe, B. 1998. A review of biodegradable polymers: Uses, current developments in the synthesis and characterization of biodegradable polyesters, blends of biodegradable polymers and recent advances in biodegradation studies. *Polymer International,* 47, 89-144.

Athanasiou, K. A., Agrawal, C. M., Barber, F. A. & Burkhart, S. S. 1998. Orthopaedic applications for PLA-PGA biodegradable polymers. *Arthroscopy-the Journal of Arthroscopic and Related Surgery,* 14, 726-737.

Barry, J. J. A., Howard, D., Shakesheff, K. M., Howdle, S. M. & Alexander, M. R. 2006. Using a core-sheath distribution of surface chemistry through 3D tissue engineering scaffolds to control cell ingress. *Advanced Materials,* 18, 1406-+.

Barry, J. J. A., Silva, M. M. C. G., Shakesheff, K. M., Howdle, S. M. & Alexander, M. R. 2005. Using plasma deposits to promote cell population of the porous interior of three-dimensional poly(D,L-lactic acid) tissue-engineering scaffolds. *Advanced Functional Materials,* 15, 1134-1140.

Bergsma, J. E., Rozema, F. R., Bos, R. R. M., Boering, G., Debruijn, W. C. & Pennings, A. J. 1995. In-Vivo Degradation and Biocompatibility Study of In-Vitro Pre-Degraded As-Polymerized Polylactide Particles. *Biomaterials,* 16, 267-274.

Bogaerts, A., Neyts, E., Gijbels, R. & van der Mullen, J. 2002. Gas discharge plasmas and their applications. *Spectrochimica Acta Part B-Atomic Spectroscopy,* 57, 609-658.

Breulmann, M., Künkel, A., Philipp, S., Reimer, V., Siegenthaler, K. O., Skupin, G. & Yamamoto, M. 2009. Biodegradable polymers. *Ullmann's Encyclopedia of Industrial Chemistry.*

Carlisle, E. S., Mariappan, M. R., Nelson, K. D., Thomes, B. E., Timmons, R. B., Constantinescu, A., Eberhart, R. C. & Bankey, P. E. 2000. Enhancing hepatocyte adhesion by pulsed plasma deposition and polyethylene glycol coupling. *Tissue Engineering,* 6, 45-52.

Chen, V. J. & Ma, P. X. 2004. Nano-fibrous poly(L-lactic acid) scaffolds with interconnected spherical macropores. *Biomaterials,* 25, 2065-2073.

Cheng, Z. Y. & Teoh, S. H. 2004. Surface modification of ultra thin poly (epsilon-caprolactone) films using acrylic acid and collagen. *Biomaterials,* 25, 1991-2001.

Chiari, C., Koller, U., Dorotka, R., Eder, C., Plasenzotti, R., Lang, S., Ambrosio, L., Tognana, E., Kon, E., Salter, D. & Nehrer, S. 2006. A tissue engineering approach to meniscus regeneration in a sheep model. *Osteoarthritis and Cartilage,* 14, 1056-1065.

Chong, M. S. K., Lee, C. N. & Teoh, S. H. 2007. Characterization of smooth muscle cells on poly(epsilon-caprolactone) films. *Materials Science & Engineering C-Biomimetic and Supramolecular Systems,* 27, 309-312.

De Geyter, N., Morent, R., Desmet, T., Trentesaux, M., Gengembre, L., Dubruel, P., Leys, C. & Payen, E. 2010. Plasma modification of polylactic acid in a medium pressure DBD. *Surface & Coatings Technology,* 204, 3272-3279.

De Geyter, N., Morent, R. & Leys, C. 2008. Influence of ambient conditions on the ageing behaviour of plasma-treated PET surfaces. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms,* 266, 3086-3090.

De Geyter, N., Morent, R., Leys, C., Gengembre, L. & Payen, E. 2007. Treatment of polymer films with a dielectric barrier discharge in air, helium and argon at medium pressure. *Surface & Coatings Technology,* 201**,** 7066-7075.

Denes, F. S. & Manolache, S. 2004. Macromolecular plasma-chemistry: an emerging field of polymer science. *Progress in Polymer Science,* 29**,** 815-885.

Desai, S. M. & Singh, R. P. 2004. *Surface modification of polyethylene*.

Desmet, T., Morent, R., De Geyter, N., Leys, C., Schacht, E. & Dubruel, P. 2009. Nonthermal Plasma Technology as a Versatile Strategy for Polymeric Biomaterials Surface Modification: A Review. *Biomacromolecules,* 10**,** 2351-2378.

Ding, Z., Chen, J. N., Gao, S. Y., Chang, J. B., Zhang, J. F. & Kang, E. T. 2004. Immobilization of chitosan onto poly-L-lactic acid film surface by plasma graft polymerization to control the morphology of fibroblast and liver cells. *Biomaterials,* 25**,** 1059-1067.

Djordjevic, I., Britcher, L. G. & Kumar, S. 2008. Morphological and surface compositional changes in poly(lactide-co-glycolide) tissue engineering scaffolds upon radio frequency glow discharge plasma treatment. *Applied Surface Science,* 254**,** 1929-1935.

Donglu, S. 2010. Synthetic biodegradable polymers. *Introduction to biomaterials.* World Scientific Publishing Co.

Ferreira, B. M. P., Pinheiro, L. M. P., Nascente, P. A. P., Ferreira, M. J. & Duek, E. A. R. 2009. Plasma surface treatments of poly(L-lactic acid) (PLLA) and poly(hydroxybutyrate-co-hydroxyvalerate) (PHBV). *Materials Science & Engineering C-Biomimetic and Supramolecular Systems,* 29**,** 806-813.

Fridman, G., Friedman, G., Gutsol, A., Shekhter, A. B., Vasilets, V. N. & Fridman, A. 2008. Applied plasma medicine. *Plasma Processes and Polymers,* 5**,** 503-533.

Fujimaki, T. 1998. Processability and properties of aliphatic polyesters, 'BIONOLLE', synthesized by polycondensation reaction. *Polymer Degradation and Stability,* 59**,** 209-214.

Gilding, D. K. & Reed, A. M. 1979. Biodegradable Polymers for Use in Surgery - Polyglycolic-Poly(Actic Acid) Homopolymers and Copolymers .1. *Polymer,* 20**,** 1459-1464.

Goddard, J. M. & Hotchkiss, J. H. 2007. Polymer surface modification for the attachment of bioactive compounds. *Progress in Polymer Science,* 32**,** 698-725.

Gomathi, N., Sureshkumar, A. & Neogi, S. 2008. RF plasma-treated polymers for biomedical applications. *Current Science,* 94**,** 1478-1486.

Guerrouani, N., Baldo, A., Bouffin, A., Drakides, C., Guimon, M. F. & Mas, A. 2007. Allylamine plasma-polymerization on PLLA surface evaluation of the biodegradation. *Journal of Applied Polymer Science,* 105**,** 1978-1986.

Gunatillake, P. A. & Adhikari, R. 2003. Biodegradable synthetic polymers for tissue engineering. *European Cells and Materials,* 5**,** 1-16.

Gupta, B., Hilborn, J., Plummer, C., Bisson, I. & Frey, P. 2002. Thermal crosslinking of collagen immobilized on poly(acrylic acid) grafted poly(ethylene terephthalate) films. *Journal of Applied Polymer Science,* 85**,** 1874-1880.

Hasirci, N., Endogan, T., Vardar, E., Kiziltay, A. & Hasirci, V. 2010. Effect of oxygen plasma on surface properties and biocompatibility of PLGA films. *Surface and Interface Analysis,* 42, 486-491.

Hasirci, V., Tezcaner, A., Hasirci, N. & Suzer, S. 2003. Oxygen plasma modification of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) film surfaces for tissue engineering purposes. *Journal of Applied Polymer Science,* 87, 1285-1289.

Heino, A., Naukkarinen, A., Kulju, T., Tormala, P., Pohjonen, T. & Makela, E. A. 1996. Characteristics of poly(L-)lactic acid suture applied to fascial closure in rats. *Journal of Biomedical Materials Research,* 30, 187-192.

Hirotsu, T., Ketelaars, A. A. J. & Nakayama, K. 2000a. Plasma surface treatment of PCL/PC blend sheets. *Polymer Engineering and Science,* 40, 2324-2331.

Hirotsu, T., Masuda, T., Matumura, Y. & Takahashi, M. 1997. Surface effects of plasma treatments on some biodegradable polyesters. *Journal of Photopolymer Science and Technology,* 10, 123-128.

Hirotsu, T., Nakayama, K., Tsujisaka, T., Mas, A. & Schue, F. 2002. Plasma surface treatments of melt-extruded sheets of poly(L-lactic acid). *Polymer Engineering and Science,* 42, 299-306.

Hirotsu, T., Tsujisaka, T., Masuda, T. & Nakayama, K. 2000b. Plasma surface treatments and biodegradation of poly(butylene succinate) sheets. *Journal of Applied Polymer Science,* 78, 1121-1129.

Ho, M. H., Hou, L. T., Tu, C. Y., Hsieh, H. J., Lai, J. Y., Chen, W. J. & Wang, D. M. 2006. Promotion of cell affinity of porous PLLA scaffolds by immobilization of RGD peptides via plasma treatment. *Macromolecular Bioscience,* 6, 90-98.

Ho, M. H., Lee, J. J., Fan, S. C., Wang, D. M., Hou, L. T., Hsieh, H. J. & Lai, J. Y. 2007. Efficient modification on PLLA by ozone treatment for biomedical applications. *Macromolecular Bioscience,* 7, 467-474.

Hu, W. & Huang, Z. M. 2010. Biocompatibility of braided poly(L-lactic acid) nanofiber wires applied as tissue sutures. *Polymer International,* 59, 92-99.

Ikada, Y. & Tsuji, H. 2000. Biodegradable polyesters for medical and ecological applications. *Macromolecular Rapid Communications,* 21, 117-132.

Kang, I. K., Choi, S. H., Shin, D. S. & Yoon, S. C. 2001. Surface modification of polyhydroxyalkanoate films and their interaction with human fibroblasts. *International Journal of Biological Macromolecules,* 28, 205-212.

Khang, G., Choee, J. H., Rhee, J. M. & Lee, H. B. 2002. Interaction of different types of cells on physicochemically treated poly(L-lactide-co-glycolide) surfaces. *Journal of Applied Polymer Science,* 85, 1253-1262.

Khorasani, M. T., Mirzadeh, H. & Irani, S. 2008. Plasma surface modification of poly (L-lactic acid) and poly (lactic-co-glycolic acid) films for improvement of nerve cells adhesion. *Radiation Physics and Chemistry,* 77, 280-287.

Koo, G. H. & Jang, J. 2008. Surface modification of poly(lactic acid) by UV/Ozone irradiation. *Fibers and Polymers,* 9, 674-678.

Kose, G. T., Ber, S., Korkusuz, F. & Hasirci, V. 2003a. Poly(3-hydroxybutyric acid-co-3-hydroxyvaleric acid) based tissue engineering matrices. *Journal of Materials Science-Materials in Medicine,* 14, 121-126.

Kose, G. T., Kenar, H., Hasirci, N. & Hasirci, V. 2003b. Macroporous poly(3-hydroxybutyrate-co-3-hydroxyvalerate) matrices for bone tissue engineering. *Biomaterials,* 24, 1949-1958.

Lam, K. H., Nijenhuis, A. J., Bartels, H., Postema, A. R., Jonkman, M. F., Pennings, A. J. & Nieuwenhuis, P. 1995. Reinforced Poly(L-Lactic Acid) Fibers As Suture Material. *Journal of Applied Biomaterials,* 6, 191-197.

Langmuir, I. 1928. Oscillations in ionized gases. *Proceedings of the National Academy of Sciences of the United States of America,* 14, 627-637.

Lee, H. U., Jeong, Y. S., Jeong, S. Y., Park, S. Y., Bae, J. S., Kim, H. G. & Cho, C. R. 2008. Role of reactive gas in atmospheric plasma for cell attachment and proliferation on biocompatible poly epsilon-caprolactone film. *Applied Surface Science,* 254, 5700-5705.

Lee, H. U., Jeong, Y. S., Koh, K. N., Jeong, S. Y., Kim, H. G., Bae, J. S. & Cho, C. R. 2009. Contribution of power on cell adhesion using atmospheric dielectric barrier discharge (DBD) plasma system. *Current Applied Physics,* 9, 219-223.

Li, G. M., Cai, Q., Bei, J. Z. & Wang, S. G. 2002. Relationship between morphology structure and composition of polycaprolactone/poly (ethylene oxide)/polylactide copolymeric microspheres. *Polymers for Advanced Technologies,* 13, 636-643.

Li, G. M., Cai, Q., Bei, J. Z. & Wang, S. G. 2003. Morphology and levonorgestrel release behavior of polycaprolactone/poly(ethylene oxide)/polylactide tri-component copolymeric microspheres. *Polymers for Advanced Technologies,* 14, 239-244.

Li, H. Y., Chang, J., Cao, A. M. & Wang, J. Y. 2005. In vitro evaluation of biodegradable poly(butylene succinate) as a novel biomaterial. *Macromolecular Bioscience,* 5, 433-440.

Lieberman, M. A. & Lichtenberg, A. J. 2005. *Principles of plasma discharges and materials processing - Second edition,* Hoboken, New Jersey, John Wiley & Sons, Inc.

Little, U., Buchanan, F., Harkin-Jones, E., Graham, B., Fox, B., Boyd, A., Meenan, B. & Dickson, G. 2009. Surface modification of poly(epsilon-caprolactone) using a dielectric barrier discharge in atmospheric pressure glow discharge mode. *Acta Biomaterialia,* 5, 2025-2032.

Loo, S. C. J., Ooi, C. P. & Boey, Y. C. F. 2004. Radiation effects on poly(lactide-co-glycolide) (PLGA) and poly(L-lactide) (PLLA). *Polymer Degradation and Stability,* 83, 259-265.

Ma, P. X., Xiaohua, L. & Youngjun, W. 2006. Porogen-induced surface modification of nano-fibrous poly(l-lactic acid) scaffolds for tissue engineering. *Biomaterials,* vol.27, no.21, 3980-3987.

Ma, Z. W., Mao, Z. W. & Gao, C. Y. 2007. Surface modification and property analysis of biomedical polymers used for tissue engineering. *Colloids and Surfaces B-Biointerfaces,* 60, 137-157.

Maurus, P. B. & Kaeding, C. C. 2004. Bioabsorbable implant material review. *Operative Techniques in Sports Medicine,* 12, 158-160.

Middleton, J. C. & Tipton, A. J. 2000. Synthetic biodegradable polymers as orthopedic devices. *Biomaterials,* 21, 2335-2346.

Miller, R. A., Brady, J. M. & Cutright, D. E. 1977. Degradation Rates of Oral Resorbable Implants (Polylactates and Polyglycolates) - Rate Modification with Changes in Pla-Pga Copolymer Ratios. *Journal of Biomedical Materials Research,* 11**,** 711-719.

Mondrinos, M. J., Dembzynski, R., Lu, L., Byrapogu, V. K. C., Wootton, D. M., Lelkes, P. I. & Zhou, J. 2006. Porogen-based solid freeform fabrication of polycaprolactone-calcium phosphate scaffolds for tissue engineering. *Biomaterials,* 27**,** 4399-4408.

Montanari, L., Costantini, M., Signoretti, E. C., Valvo, L., Santucci, M., Bartolomei, M., Fattibene, P., Onori, S., Faucitano, A., Conti, B. & Genta, I. 1998. Gamma irradiation effects on poly(DL-lactictide-co-glycolide) microspheres. *Journal of Controlled Release,* 56**,** 219-229.

Morent, R., De Geyter, N., Desmet, T., Dubruel, P. & Leys, C. 2011. Plasma Surface Modification of Biodegradable Polymers: A Review. *Plasma Processes and Polymers,* 8**,** 171-190.

Morent, R., De Geyter, N., Gengembre, L., Leys, C., Payen, E., Van Vlierberghe, S. & Schacht, E. 2008a. Surface treatment of a polypropylene film with a nitrogen DBD at medium pressure. *European Physical Journal-Applied Physics,* 43**,** 289-294.

Morent, R., De Geyter, N. & Leys, C. 2008b. Effects of operating parameters on plasma-induced PET surface treatment. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms,* 266**,** 3081-3085.

Morent, R., De Geyter, N., Trentesaux, M., Gengembre, L., Dubruel, P., Leys, C. & Payen, E. 2010. Influence of Discharge Atmosphere on the Ageing Behaviour of Plasma-Treated Polylactic Acid. *Plasma Chemistry and Plasma Processing,* 30**,** 525-536.

Morent, R., De Geyter, N., Van Vlierberghe, S., Dubruel, P., Leys, C., Gengembre, L., Schacht, E. & Payen, E. 2009. Deposition of HMDSO-based coatings on PET substrates using an atmospheric pressure dielectric barrier discharge. *Progress in Organic Coatings,* 64**,** 304-310.

Nair, L. S. & Laurencin, C. T. 2007. Biodegradable polymers as biomaterials. *Progress in Polymer Science,* 32**,** 762-798.

Ojumu, T. V., Yu, J. & Solomon, B. O. 2004. Production of Polyhydroxyalkanoates, a bacterial biodegradable polymer. *African journal of Biotechnology,* 3**,** 18-24.

Park, H., Lee, J. W., Park, K. E., Park, W. H. & Lee, K. Y. 2010. Stress response of fibroblasts adherent to the surface of plasma-treated poly(lactic-co-glycolic acid) nanofiber matrices. *Colloids and Surfaces B-Biointerfaces,* 77**,** 90-95.

Park, K. E., Lee, K. Y., Lee, S. J. & Park, W. H. 2007. Surface characteristics of plasma-treated PLGA nanofibers. *Macromolecular Symposia,* 249**,** 103-108.

Place, E. S., George, J. H., Williams, C. K. & Stevens, M. M. 2009. Synthetic polymer scaffolds for tissue engineering. *Chemical Society Reviews,* 38**,** 1139-1151.

Pompe, T., Keller, K., Mothes, G., Nitschke, M., Teese, M., Zimmermann, R. & Werner, C. 2007. Surface modification of poly(hydroxybutyrate) films to control cell-matrix adhesion. *Biomaterials,* 28**,** 28-37.

Prabhakaran, M. P., Venugopal, J., Chan, C. K. & Ramakrishna, S. 2008. Surface modified electrospun nanofibrous scaffolds for nerve tissue engineering. *Nanotechnology,* 19.

Qian, H. T., Bei, J. Z. & Wang, S. G. 2000. Synthesis, characterization and degradation of ABA block copolymer of L-lactide and epsilon-caprolactone. *Polymer Degradation and Stability,* 68, 423-429.

Ramchandani, M. & Robinson, D. 1998. In vitro and in vivo release of ciprofloxacin from PLGA 50 : 50 implants. *Journal of Controlled Release,* 54, 167-175.

Rasal, R. M., Janorkar, A. V. & Hirt, D. E. 2010. Poly(lactic acid) modifications. *Progress in Polymer Science,* 35, 338-356.

Rausher, H., Perucca, M. & Buyle, G. 2010. *Plasma Technology for hyperfunctional surfaces,* Weinheim, Wiley-VCH Verlag GmbH & Co.

Reed, A. M. & Gilding, D. K. 1981. Biodegradable Polymers for Use in Surgery - Poly(Glycolic)-Poly(Lactic Acid) Homo and Co-Polymers .2. Invitro Degradation. *Polymer,* 22, 494-498.

Ryu, G. H., Yang, W. S., Roh, H. W., Lee, I. S., Kim, J. K., Lee, G. H., Lee, D. H., Park, B. J., Lee, M. S. & Park, J. C. 2005. Plasma surface modification of poly(D,L-lactic-co-glycolic acid)(65/35) film for tissue engineering. *Surface & Coatings Technology,* 193, 60-64.

Safinia, L., Wilson, K., Mantalaris, A. & Bismarck, A. 2007. Atmospheric plasma treatment of porous polymer constructs for tissue engineering applications. *Macromolecular Bioscience,* 7, 315-327.

Safinia, L., Wilson, K., Mantalaris, A. & Bismarck, A. 2008. Through-thickness plasma modification of biodegradable and nonbiodegradable porous polymer constructs. *Journal of Biomedical Materials Research Part A,* 87A, 632-642.

Shelton, J. R., Lando, J. B. & Agostini, D. E. 1971. Synthesis and Characterization of Poly (Beta-Hydroxybutyrate). *Journal of Polymer Science Part B-Polymer Letters,* 9, 173- & .

Shen, H., Hu, X. X., Bei, J. Z. & Wang, S. G. 2008. The immobilization of basic fibroblast growth factor on plasma-treated poly(lactide-co-glycolide). *Biomaterials,* 29, 2388-2399.

Shen, H., Hu, X. X., Yang, F., Bel, J. Z. & Wang, S. G. 2007. Combining oxygen plasma treatment with anchorage of cationized gelatin for enhancing cell affinity of poly(lactide-co-glycolide). *Biomaterials,* 28, 4219-4230.

Siow, K. S., Britcher, L., Kumar, S. & Griesser, H. J. 2006. Plasma methods for the generation of chemically reactive surfaces for biomolecule immobilization and cell colonization - A review. *Plasma Processes and Polymers,* 3, 392-418.

Storey, R. F. & Taylor, A. E. 1998. Effect of stannous octoate on the composition, molecular weight, and molecular weight distribution of ethylene glycol-initiated poly(epsilon-caprolaccone). *Journal of Macromolecular Science-Pure and Applied Chemistry,* A35, 723-750.

Taylor, M. S., Daniels, A. U., Andriano, K. P. & Heller, J. 1994. 6 Bioabsorbable Polymers - In-Vitro Acute Toxicity of Accumulated Degradation Products. *Journal of Applied Biomaterials,* 5, 151-157.

Tezcaner, A., Bugra, K. & Hasirci, V. 2003. Retinal pigment epithelium cell culture on surface modified poly(hydroxybutyrate-co-hydroxyvalerate) thin films. *Biomaterials,* 24, 4573-4583.

Tserki, V., Matzinos, P., Pavlidou, E., Vachliotis, D. & Panayiotou, C. 2006. Biodegradable aliphatic polyesters. Part I. Properties and biodegradation of poly(butylene succinate-co-butylene adipate). *Polymer Degradation and Stability,* 91**,** 367-376.

Vaquette, C., Frochot, C., Rahouadj, R. & Wang, X. 2008. An innovative method to obtain porous PLLA scaffolds with highly spherical and interconnected pores. *Journal of Biomedical Materials Research Part B-Applied Biomaterials,* 86B**,** 9-17.

Vasilets, V. N., Hermel, G., Konig, U., Werner, C., Muller, M., Simon, F., Grundke, K., Ikada, Y. & Jacobasch, H. J. 1997. Microwave CO2 plasma-initiated vapour phase graft polymerization of acrylic acid onto polytetrafluoroethylene for immobilization of human thrombomodulin. *Biomaterials,* 18**,** 1139-1145.

Vroman, I. & Tighzert, L. 2009. Biodegradable polymers. *Materials,* 2**,** 307-344.

Wan, Y. Q., Tu, C. F., Yang, J. A., Bei, J. Z. & Wang, S. G. 2006. Influences of ammonia plasma treatment on modifying depth and degradation of poly(L-lactide) scaffolds. *Biomaterials,* 27**,** 2699-2704.

Wang, E. A., Rosen, V., Dalessandro, J. S., Bauduy, M., Cordes, P., Harada, T., Israel, D. I., Hewick, R. M., Kerns, K. M., Lapan, P., Luxenberg, D. P., McQuaid, D., Moutsatsos, I. K., Nove, J. & Wozney, J. M. 1990. Recombinant Human Bone Morphogenetic Protein Induces Bone-Formation. *Proceedings of the National Academy of Sciences of the United States of America,* 87**,** 2220-2224.

Wang, S. G., Chen, H. L., Cai, Q. & Bei, J. Z. 2001. Degradation and 5-fluorouracil release behavior in vitro of polycaprolactone/poly(ethylene oxide)/polylactide tri-component copolymer. *Polymers for Advanced Technologies,* 12**,** 253-258.

Wang, Y. J., Lu, L., Zheng, Y. D. & Chen, X. F. 2006. Improvement in hydrophilicity of PHBV films by plasma treatment. *Journal of Biomedical Materials Research Part A,* 76A**,** 589-595.

Wang, Y. Q., Qu, X., Lu, J., Zhu, C. F., Wan, L. J., Yang, J. L., Bei, J. Z. & Wang, S. G. 2004. Characterization of surface property of poly(lactide-co-glycolide) after oxygen plasma treatment. *Biomaterials,* 25**,** 4777-4783.

Xie, Z. W. & Buschle-Diller, G. 2010. Electrospun Poly(D,L-lactide) Fibers for Drug Delivery: The Influence of Cosolvent and the Mechanism of Drug Release. *Journal of Applied Polymer Science,* 115**,** 1-8.

Yang, J., Bei, J. Z. & Wang, S. G. 2002. Enhanced cell affinity of poly (D,L-lactide) by combining plasma treatment with collagen anchorage. *Biomaterials,* 23**,** 2607-2614.

Yildirim, E. D., Ayan, H., Vasilets, V. N., Fridman, A., Guceri, S. & Sun, W. 2008. Effect of dielectric barrier discharge plasma on the attachment and proliferation of osteoblasts cultured over poly(epsilon-caprolactone) scaffolds. *Plasma Processes and Polymers,* 5**,** 58-66.

Zelzer, M., Scurr, D., Abdullah, B., Urquhart, A. J., Gadegaard, N., Bradley, J. W. & Alexander, M. R. 2009. Influence of the Plasma Sheath on Plasma Polymer Deposition in Advance of a Mask and down Pores. *Journal of Physical Chemistry B,* 113**,** 8487-8494.

Zhu, L. P., Zhu, B. K., Xu, L., Feng, Y. X., Liu, F. & Xu, Y. Y. 2007. Corona-induced graft polymerization for surface modification of porous polyethersulfone membranes. *Applied Surface Science,* 253**,** 6052-6059.

Zhu, Y. B., Gao, C. Y., Liu, X. Y. & Shen, J. C. 2002. Surface modification of polycaprolactone membrane via aminolysis and biomacromolecule immobilization for promoting cytocompatibility of human endothelial cells. *Biomacromolecules,* 3**,** 1312-1319.

Zhu, Y. B., Gao, C. Y., Liu, Y. X. & Shen, J. C. 2004. Endothelial cell functions in vitro cultured on poly(L-lactic acid) membranes modified with different methods. *Journal of Biomedical Materials Research Part A,* 69A**,** 436-443.

11

Poly(Lactic Acid)-Based Biomaterials: Synthesis, Modification and Applications

Lin Xiao[1,2], Bo Wang[1,2], Guang Yang*[1,2] and Mario Gauthier[3]
[1]*College of Life Science and Technology, Huazhong University of Science and Technology,*
[2]*National Engineering Research Center for Nano-Medicine,*
Huazhong University of Science and Technology,
[3]*Department of Chemistry, University of Waterloo,*
[1,2]*China*
[3]*Canada*

1. Introduction

Social and economic development has driven considerable scientific and engineering efforts on the discovery, development, and utilization of polymers. Widespread reliance in everyday life on conventional polymeric materials such as polyolefins has resulted in serious pollution which cannot be resolved in a straightforward fashion. Sustainable development and a green economy both require brand new materials which can avoid the occurrence of these problems.

Poly(lactic acid) (PLA), an aliphatic polyester, has outstanding advantages over other polymers, and may thus be part of the solution. As early as the 1970's, PLA products have been approved by the US Food and Drug Administration (FDA) for direct contact with biological fluids. Four of its most attractive advantages are renewability, biocompatibility, processability, and energy saving (Rasal, 2010). First of all, PLA is derived from renewable and degradable resources such as corn and rice, which can help alleviate the energy crisis as well as reduce the dependence on fossil fuels of our society; PLA and its degradation products, namely H_2O and CO_2, are neither toxic nor carcinogenic to the human body, hence making it an excellent material for biomedical applications including sutures, clips, and drug delivery systems (DDS). Furthermore, PLA can be processed by film casting, extrusion, blow molding, and fiber spinning due to its greater thermal processability in comparison to other biomaterials such as poly(ethylene glycol) (PEG), poly(hydroxyalkanoates) (PHAs), and poly(ε-caprolactone) (PCL) (Rhim et al., 2006). These thermal properties contribute to the application of PLA in industry in fields such as textiles and food packaging. Last but not least, PLA production consumes 25-55% less fossil energy than petroleum-based polymers. Cargill Dow has even targeted a reduction in fossil energy consumption by more than 90% as compared to any of the petroleum-based polymers for the near future, which will surely also lead to significant reductions in air and water pollutant emissions. It is also noteworthy that the total amount of water required for PLA production is competitive with the best performing petroleum-based polymers. This energy-saving feature perfectly caters to the

---

[*] Corresponding author

new concept of “low-carbon economy” which emerged recently in response to the global warming and energy crisis concerns, and makes investment in PLA a necessary and wise strategy in the future (Vink et al., 2003). Fig. 1 shows the cycle of PLA in nature.

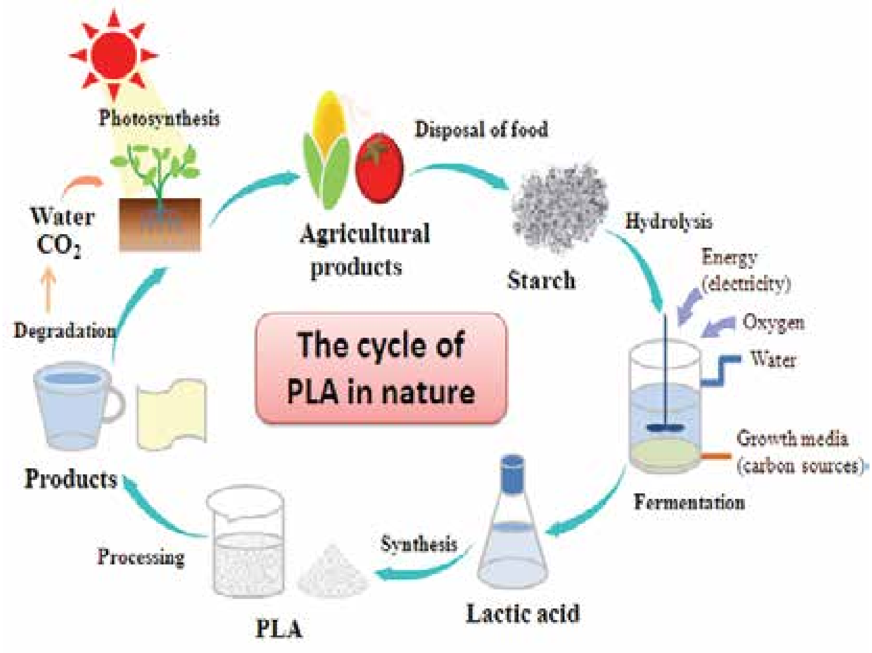

Photosynthesis
Disposal of food
Water
CO_2
Agricultural products
Starch
Hydrolysis
Energy (electricity)
Oxygen
Water
Degradation
The cycle of PLA in nature
Growth media (carbon sources)
Products
Fermentation
Processing
Synthesis
PLA
Lactic acid

Fig. 1. The cycle of PLA in nature.

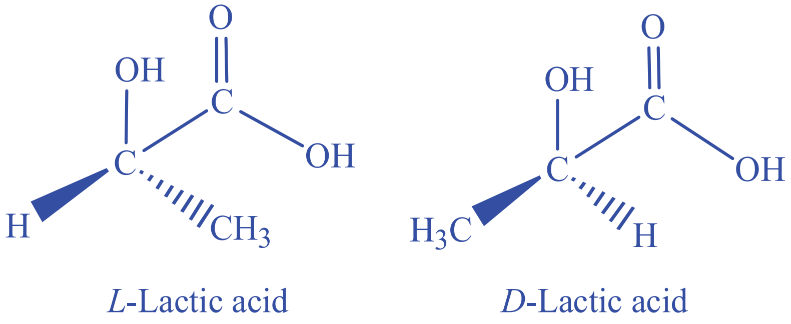
While PLA can be considered an eco-friendly biomaterial with excellent properties, it also has many obvious drawbacks when confronted with requirements for certain applications: 1) Its degradation rate through hydrolysis of the backbone ester groups is too slow. This process sometimes takes several years, which can impede its biomedical and food packaging applications (Bergsma et al., 1995). 2) PLA is very brittle, with less than 10% elongation at break, thus it is not suitable for demanding mechanical performance applications unless it is suitably modified (Rasal & Hirt, 2009). 3) PLA is strongly hydrophobic and can elicit an inflammatory response from the tissues of living hosts, because of its low affinity with cells when it is used as a tissue engineering material. 4) Another limitation of PLA towards its wider industrial application is its limited gas barrier properties which prevent its complete access to industrial sectors such as packaging (Singh et al., 2003). Considering the disadvantages of PLA stated above and its high cost (another shortcoming of that material), it is not surprising that PLA has not received the attention it deserves. Nevertheless, researchers have examined different methods for the bulk or surface modification of PLA, the introduction of other components, or the control of its surface energy, surface charge and surface roughness, depending on the requirements of specific applications.

Previous reviews have examined different aspects of PLA chemistry and engineering. Thus Maharana et al. (Maharana et al., 2009) presented a review on the melt-solid polycondensation of lactic acid (LA). Gupta et al. (Gupta et al., 2007) presented an overview of the production of PLA fibers by various methods, along with correlations between the structure and the properties of the fibers. Butterwick et al. (Butterwick et al., 2009) discussed the applications of PLA in Europe and the United States with respect to practitioner experiences and techniques to optimize the outcomes. Rasal et al. (Rasal et al., 2010) examined the chemical modification of PLA, while Graupner et al. (Graupner et al., 2009) assessed the production and the mechanical characteristics of composites prepared from PLA and renewable raw materials including cotton, hemp, kenaf, and man-made cellulose fibres (Lyocell) by compression molding.

In this chapter we will underline novel ideas or technologies introduced over the last 5-10 years, emphasizing some ambitious work which, even though it appears less successful than other mature methods, introduces concepts that may prove extremely positive in the near future. We will also attempt to foretell developmental trends on the basis of social demands and the progress achieved so far. More traditional topics including the synthesis, modification, and applications of PLA in biomedical field will be introduced mainly to provide a more comprehensive picture of PLA as a biomaterial.

2. Physical and chemical properties of PLA

L-lactic acid and D-lactic acid, the two isomers of lactic acid, are shown in Scheme 1. Pure L-lactic acid or D-lactic acid, or mixtures of both components are needed for the synthesis of PLA.



Scheme 1. The stereoisomers of lactic acid.

The homopolymer of LA is a white powder at room temperature with T_g and T_m values of about 55°C and 175°C, respectively. High molecular weight PLA is a colorless, glossy, rigid thermoplastic material with properties similar to polystyrene. The two isomers of LA can produce four distinct materials: Poly(*D*-lactic acid) (PDLA), a crystalline material with a regular chain structure; poly(*L*-lactic acid) (PLLA), which is hemicrystalline, and likewise with a regular chain structure; poly(*D*,*L*-lactic acid) (PDLLA) which is amorphous; and *meso*-PLA, obtained by the polymerization of *meso*-lactide. PDLA, PLLA and PDLLA are soluble in common solvents including benzene, chloroform, dioxane, etc. and degrade by simple hydrolysis of the ester bond even in the absence of a hydrolase. PLA has a degradation half-life in the environment ranging from 6 months to 2 years, depending on the size and shape of the article, its isomer ratio, and the temperature. The tensile properties of PLA can vary widely depending on whether it is annealed or oriented, or its degree of crystallinity (Garlotta et al., 2001). Some of the physical and chemical properties of PLA are summarized in Table 1.

Properties	PDLA	PLLA	PDLLA
Solubility	All are soluble in benzene, chloroform, acetonitrile, tetrahydrofuran (THF), dioxane etc., but insoluble in ethanol, methanol, and aliphatic hydrocarbons		
Crystalline structure	Crystalline	Hemicrystalline	Amorphous
Melting temperature (T_m)/ °C	~180	~180	Variable
Glass transition temperature (T_g)/ °C	50-60	55-60	Variable
Decomposition temperature/°C	~200	~200	185-200
Elongation at break/ (%)	20-30	20-30	Variable
Breaking strength/ (g/d)	4.0-5.0	5.0-6.0	Variable
Half-life in 37°C normal saline	4-6 months	4-6 months	2-3 months

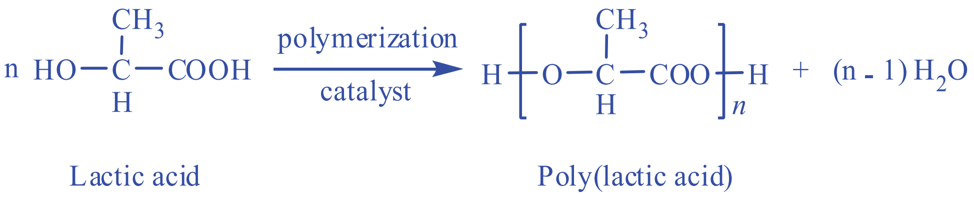
Table 1. Selected physical and chemical properties of PLA.

3. Synthesis of PLA

Two main synthetic methods are used to obtain PLA: Direct polycondensation (including solution polycondensation and melt polycondensation), and ring-opening polymerization (ROP).

3.1 Direct polymerization

Since the LA monomer has both -OH and -COOH groups, necessary for polymerization, the reaction can take place directly by self-condensation (Scheme 2):



Scheme 2. Direct polymerization.

Direct polymerization includes solution and melt polycondensation, depending on whether a solvent is used in the reaction to dissolve the PLA or not.

3.1.1 Solution polycondensation

In this case an organic solvent capable of dissolving the PLA without interfering with the reaction is added, and the mixture is refluxed with removal of the water generated in the polycondensation process, which is beneficial to achieve a high molecular weight. Many procedures yield PLA with a weight-average molecular weight (M_w) of over 200,000 by this method (Ohta et al., 1995; Ichikawa et al., 1995). This approach was developed by Carothers and is still used by Mitsui Chemicals. The resultant polymer can be coupled with isocyanates, epoxides or peroxides to produce a range of molecular weights (Lunt et al., 1998). The reaction proceeds smoothly, however solution polymerization suffers from certain disadvantages such as being susceptible to impurities from the solvent and various side reactions including racemization and trans-esterification. It also consumes large volumes of organic solvents, which are potential pollutants to the environment.

Under optimized conditions, Ajioka et al. obtained PLA with M_w > 300,000 by this method (Ajioka et al., 1995). Characterization data have shown that the glass transition temperatures (T_g) of PLA and polylactide synthesized by the conventional lactide process are essentially identical (T_g = 58°C and 59 °C, respectively), but PLA has a lower melting point (T_m = 163 °C) than polylactide (T_m = 178°C). The mechanical properties of the two polymers are also very similar.

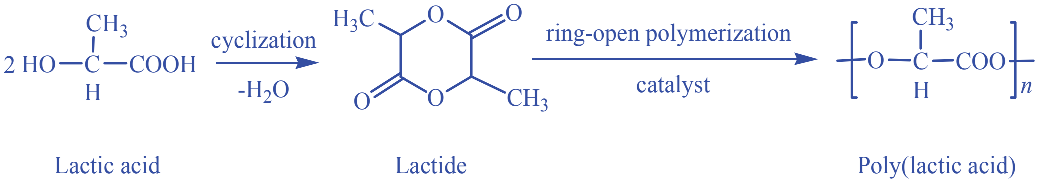
3.1.2 Melt polycondensation

In contrast to solution polycondensation, the melt polycondensation of monomers can proceed without any organic solvent, but only if the temperature of the reaction remains above the T_m of the polymer (Gao et al., 2002). Moon et al. discovered that high M_w PLLA [$M_w \geq 100{,}000$] could be produced in this way in a relatively short reaction time (≤ 15 h) (Moon et al., 2000). This method can lower the cost of the synthesis significantly due to the simplified procedure, but major problems still need to be solved before it can be applied industrially because of its sensitivity to reaction conditions (Maharana et al., 2009). Thus Moon et al. worked to develop a melt/solid polycondensation technique using a binary catalyst system (tin dichloride hydrate and *p*-toluenesulfonic acid) (Moon et al., 2001). Simply put, thermal oligocondensates of LA were first subjected to melt polycondensation to obtain a melt polycondensate, which was then subjected to solid state polycondensation at 105°C. As a consequence, the molecular weight of the PLA was as high as 600,000 after a short reaction time under optimized conditions.

In summary, these one-step polymerization processes are relatively economical and easy to control, but they are equilibrium reactions affected by numerous parameters such as the temperature, the reaction time, catalysts, pressure, and so on. These factors can strongly influence the molecular weight of the products obtained. Besides, the water generated in this process can cause high molecular weight PLA to break down at high reaction temperatures. Thus the polymer resulting from these reactions usually has an unsatisfactorily low molecular weight. Attention must be paid to three aspects of the reaction to obtain a high molecular weight, namely controlling the reaction kinetics, removing the water formed, and preventing the degradation of the PLA chains.

3.2 Ring-opening polymerization

Considering the drawbacks of direct polymerization, PLA is typically synthesized by ring-opening polymerization (ROP) (Scheme 3), an important and effective method to manufacture high molecular weight PLA. This reaction requires strict purity of the lactide monomer, obtained by dimerization of the lactic acid monomer. PLA is obtained by using a catalyst with the monomer under vacuum or an inert atmosphere. By controlling the residence time and the temperatures in combination with the catalyst type and concentration, it is possible to control the ratio and sequence of *D*- and *L*-lactic acid (LA) units in the final polymer (Lunt et al., 1998). The polymerization mechanism involved can be ionic, coordination, or free-radical, depending on type of catalyst employed (Penczek et al., 2000). Most of the researchers are now exploring new and effective catalysts. Köhn et al. (Köhn et al., 2003) first reported that the ROP of *D,L*-lactide could be catalyzed by bis(trimethyl triazacyclohexane) praseodymium triflate [$(Me_3TAC)_2$ $Pr(OTf)_3$] (Cat), while $Pr(OTf)_3$ by itself had a poor catalytic activity. Cat was found to catalyze the polymerization of *D,L*-lactide in various solvents (THF, dichloromethane, ethyl acetate, and toluene) without any additional reagents. The optimal polymer yield (95%) and molecular weight (18,000) were obtained after 18 h at 170°C, with a ratio of [LA]:[Cat] = 1000. John et al. (John et al., 2007) produced one of the few reports on lactide polymerization with a Cu-based catalyst. ROP of *L*-lactide catalyzed by $\{2\text{-}[1\text{-}(2,6\text{-diethylphenylimino})\text{ethyl}]\text{phenoxy}\}_2\text{Cu(II)}$ yielded the highest M_w (26.3×10^3) with a monomer conversion of 57%. Two other copper complexes, $\{2\text{-}[1\text{-}(2,6\text{-dimethylphenylimino})\text{ethyl}]\text{phenoxy}\}_2\text{Cu(II)}$ and $\{2\text{-}[1\text{-}(2\text{-methylphenylimino})\text{ethyl}]\text{phenoxy}\}_2\text{Cu(II)}$, also catalyzed the reaction under solvent-free melt conditions (160°C) but produced polylactides of moderate molecular weights (M_w = 12.0×10^3 and 15.9×10^3, respectively).



Scheme 3. Ring-opening polymerization.

Numerous studies have examined the influence of different factors such as the concentration and type of catalyst, monomer purity, and temperature on the polymerization of lactide. Special attention has been paid to the catalyst. Currently tin octoate is the most widely used catalyst for the ring-opening polymerization of lactides, but numerous novel efficient metal-free catalytic systems are emerging as valuable alternatives (Jérôme & Lecomte, 2008). The heavy metal-based catalysts are indeed very likely to contaminate the product, which complicates the purification of the PLA obtained and also limits the applications of PLA in the fields of food packaging and biomedicine. Some of the means developed to solve this problem will be addressed later.

3.3 New approaches

The inherent disadvantages of the traditional synthetic methodology have led some researchers to explore solutions such as the development of nontoxic catalysts, unusual polymerization conditions, or other polymerization pathways (Lassalle & Ferreira, 2008).

To solve the potential pollution problems caused by heavy metal catalysts, many nontoxic catalysts derived from magnesium (Wu et al., 2005), calcium (Zhong et al., 2003), zinc (Sarazin et al., 2004), alkali metals (Chisholm et al., 2003), and aluminum (Nomura et al., 2002) have been developed for the ROP of lactides (Deng et al., 2000; Ejfler et al., 2005). For example, Chen et al. (Chen et al., 2007) tested a series of β-diketiminate zinc complexes as initiators for the ROP of lactide and they were all highly active, however the M_w attained was unsatisfactory. It is worth noting that a variety of rare earth derivatives are usually highly reactive, which entitles them to be very promising initiators for the ROP of lactide (Agarwal et al., 2000).

With respect to unusual reaction conditions, supercritical CO_2 ($scCO_2$) technology has attracted much attention because this environmentally friendly, chemically inert, inexpensive, non-toxic, and nonflammable solvent can be substituted for organic solvents (Nalawade et al., 2006). Yoda et al. (Yoda et al., 2004) thus carried out the synthesis of PLLA from an *L*-lactic acid oligomer in $scCO_2$ with dicyclohexyldimethylcarbodiimide (DCC) as an esterification promoter and 4-dimethylaminopyridine (DMAP) as a catalyst. PLLA with a number-average molecular weight M_n reaching 13,500 was obtained in 95% yield after 24 h at 3500 psi and 80°C. The molecular weight distribution of the products was also narrower than for PLLA prepared by melt–solid phase polymerization under conventional conditions. Not only can $scCO_2$ be used as a medium to synthesize polymers, but it can also serve in the purification and processing of the polymer microparticles obtained (Kang et al., 2008).

The direct polycondensation of lactic acid has been considered to have a promising future due to its low cost; however it is hard to increase the molecule weight due to the difficulty in removing the water from the system under these conditions. One way to solve this problem is a chain-extension method, although the properties of the PLA obtained in this way can be somewhat affected by the procedure. Simply put, hydroxyl- or carboxyl-terminated low molecular weight PLA obtained by direct polymerization can be linked together through a chain extender, which is a bifunctional compound carrying highly reactive functional groups. Many achievements have been reported in this area, hexamethylene diisocyanate (HDI) being the most widely used chain extender for hydroxyl-terminated prepolymers since the work done by Woo and coworkers (Woo et al., 1995). Finding new and satisfactory chain extenders will remain a major goal in the near future, since HDI is toxic and subject to side reaction in this process.

In addition, LA-polymerizing enzymes functioning in replacement of metal catalysts should enable the biosynthesis of PLA, even though it is enormously challenging both in terms of research and industrial implementation. The best solution could be the development of a PLA-producing microorganism, but this has not been reported so far. Taguchi et al. (Taguchi et al. 2008) have nonetheless obtained encouraging results by developing a recombinant *Escherichia coli* strain allowing the synthesis of LA-based polyesters by introducing the gene encoding polyhydroxyalkanoate (PHA) synthase. This is illustrated in Fig. 2. They thus achieved the one-step biosynthesis of a copolymer with 6 mol% of lactate and 94 mol% of 3-hydroxybutyrate units, having a molecular weight of 1.9×10^5. This extremely important result represents a milestone towards the biological synthesis of PLA and confirms that the work is moving in the right direction. At present, the LA fraction in the copolyesters has been enriched up to 96 mol% (Shozui et al., 2011), so the synthesis of homopolymers of LA represents a major goal. To that end, the current microbial cell factory ought to be improved with further evolved LA-polymerizing enzymes (LPE) and metabolic engineering-based optimization (Taguchi, 2010). Matsumura et al. (Matsumura et al., 1997) likewise reported the lipase PC-catalyzed polymerization of cyclic diester-D,L-lactide at a temperature of 80-130°C to yield poly(lactic acid) with molecular masses of up to 12,600. Other novel methods (e.g. metal-free catalysts, non-catalytic systems) are also under development (Zhong et al., 2003; Achmad et al., 2009). The advantages and disadvantages of the PLA synthesis methods mentioned above are summarized in Table 2.

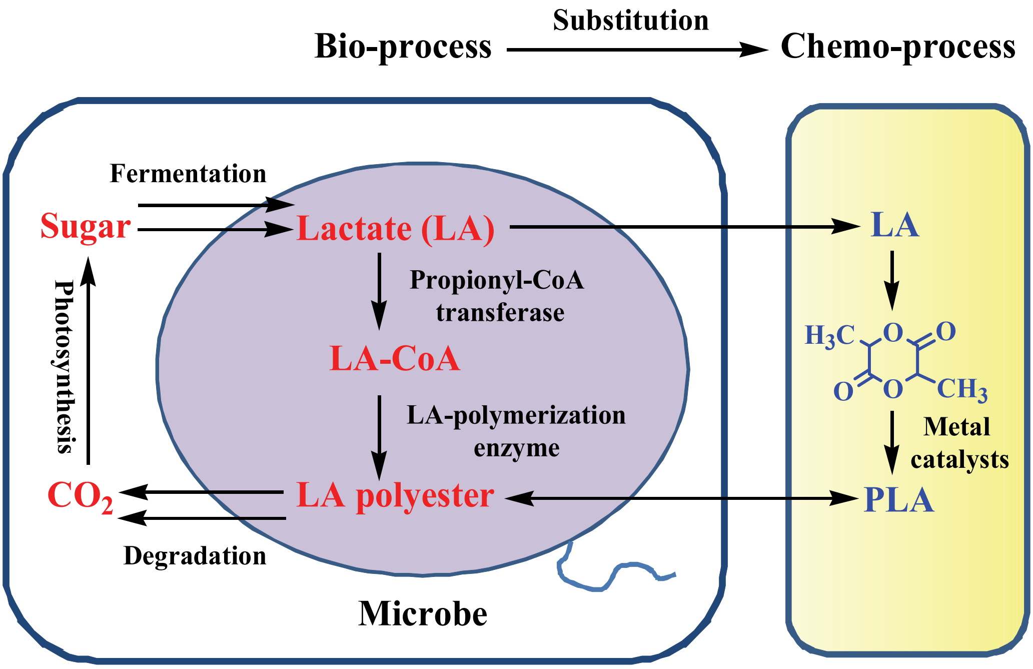

Bio-process
Substitution
Chemo-process
Fermentation
Sugar
Lactate (LA)
Photosynthesis
Propionyl-CoA transferase
LA-CoA
LA-polymerization enzyme
CO_2
LA polyester
Degradation
Microbe
LA
H_3C
O
O
O
O
CH_3
Metal catalysts
PLA

Fig. 2. Mechanism for the bio-synthesis of LA polyester. In the bio-process, the LA monomer is converted into LA-CoA ,which is recognized by the LA-polymerizing enzyme recruited from microbial PHA synthase (Tajima et al., 2009).

Synthesis methods	Advantages	Disadvantages
Solution polycondensation	One-step, economical and easy to control	Impurities, side reactions, pollution, low molecular weight PLA
Melt polycondensation		High reaction temperature, sensitivity to reaction conditions, low molecular weight PLA
Ring-opening polymerization	High molecular weight PLA	Requires strict purity of the lactide monomer, related high cost
New solutions (new catalysts, polymerization conditions, etc)	Efficient, non-toxic, no pollution , high molecular weight PLA, etc.	Under development
Biosynthesis	One-step, efficient, non-toxic, no pollution, low cost, etc.	Under development

Table 2. Comparison of PLA synthesis methods.

4. Modification of PLA

The major drawbacks of PLA limiting its applications are its poor chemical modifiability and mechanical ductility, slow degradation profile, and poor hydrophilicity. In order to be suitable for specific biomedical applications, the PLA has been modified mainly concerning two aspects: Bulk properties and surface chemistry. To achieve this, both chemical modification and physical modification have been tried, involving the incorporation of functional monomers with different molecular architectures and compositions, the tuning of crystallinity and processibility via blending and plasticization, etc., which are described in the following sections.

4.1 Bulk modification

Biomaterials must possess bulk properties, particularly hydrophilic and mechanical properties, meeting special requirements. Critical factors affecting these characteristics include chemical additives, composition, and morphological structure. At present considerable research focuses on a variety of hydrolytic groups, controlling the flexibility and crystallinity of the molecular chains, and the presence of hydrophilic groups.

4.1.1 Physical modification

Blending, plasticization, and composition variations belong to this category.

Blending

Polymer blending is an effective, simple, and versatile method to develop new materials with tailored properties without synthesizing new polymers (Peesan et al., 2005). The properties of different polymers (biodegradable and non-biodegradable) can be combined by blending with PLA, or even new properties can arise in the products due to interactions between the components. Biodegradable components blended with PLA include poly(ethylene glycol) (PEG), poly(β-hydroxybutyrate) (PHB), poly(ε-caprolactone) (PCL), poly(butylene adipate-*co*-terephthalate) (PBAT), chitosan, and starch (Sheth et al., 1997). While blends of PLA and non-biodegradable polymers have not been as extensively studied, low-density polyethylene (LDPE), poly(vinyl acetate) (PVA), and polypropylene (PP) have been examined. Reddy et al. (Reddy et al., 2008) found that PLA in blends obtained from five ratios of PLA/PP had substantially better resistance to biodegradation and hydrolysis, and improved dyeability with dispersed dyes. However most of these blends are immiscible (phase-separated) and display poor mechanical properties due to low interfacial adhesion between the polymer phases.

To improve the processing and mechanical properties of PLA without sacrificing its degradability and biocompatibility, Xu et al. (Xu et al., 2009) blended PLA with a new degradable thermoplastic derived from konjac glucomannan (TKGM), synthesized by graft copolymerization of vinyl acetate and methyl acrylate onto konjac glucomannan (KGM). Dynamic mechanical analysis (DMA) and scanning electron microscopy (SEM) measurements showed that the PLA/TKGM system was miscible due to specific interactions between PLA and TKGM. This led to a maximum elongation at break of 520% for the blend (20/80), as compared to 14% for neat PLA. The impact strength also increased from 11.9 kJ/m^2 for neat PLA to 26.9 kJ/m^2 for the 20/80 blend. The synthesis of new polymers, biodegradable or non-biodegradable, to be compatibly blended with PLA, will represent a major task in the future.

Plasticization

PLA is a glassy polymer with poor elongation at break (typically less than 10%). The modification of PLA with different biodegradable and non-biodegradable plasticizers, having a low molecular weight but a high boiling point and a low volatility, has been explored as a mean to lower the T_g and increase the ductility and softness of PLA. This has been achieved by varying the molecular weight, the polarity and functional groups of the plasticizers. Biocompatible molecules such as oligomeric lactic acid, oligomeric citrate ester, oligomeric PEG, and glycerol are all plasticizers of choice for PLA (Martin & Averous, 2001; Ljungberg et al., 2005). Ljungberg et al. (Ljungberg & Wesslén, 2002) have blended PLA with five plasticizers (triacetine, tributyl citrate, triethyl citrate, acetyl tributyl citrate, and acetyl triethyl citrate) and found that triacetine and tributyl citrate were more effective as plasticizers than the others to obtain a significant decrease in T_g for PLA.

Wang et al. (Wang et al., 2009) found that diisononyl cyclohexane-1,2-dicarboxylate (DINCH), a new plasticizer obtained by the hydrogenation of the benzene ring of o-phthalates, had limited compatibility with PLA when compared with tributyl citrate ester (TBC). PLA samples plasticized with 10 and 20 phr DINCH gave a constant T_g of 50°C. They were stiff materials displaying elevated values of elongation at break (129% and 200%, respectively) and impact strength (41.1 MPa and 30.1 MPa, respectively). On the other hand, TBC significantly decreased the T_g and increased the crystallinity of PLA, the PLA/TBC (20 phr TBC) blend being a soft material with a T_g of 24°C. Results from thermogravimetric and thermal analysis also indicated that PLA plasticized with DINCH had good mechanical properties and excellent water resistance (as reflected in time-dependent weight loss data in phosphate buffer) and aging resistance (characterized by the mechanical and thermal properties of specimens exposed to ambient conditions for 4 months).

Composition

Fibers can serve as fillers in the formation of PLA composites processable by compression or injection molding, to enhance the thermal stability, the hydrolysis resistance, or the mechanical properties of PLA. Several investigations on PLA composites prepared from natural and modified cellulose fibers have shown that their mechanical properties scale with the mass fraction of added fibers (Wan et al., 2001; Mathew et al., 2005). Optimization of the natural fiber-reinforced PLA composites, in terms of mechanical and other properties, is critical to minimize their cost, tailor their biodegradability, and broaden their areas of application. Inorganic fillers can also contribute to property modification. Table 3 provides a comparison of some of the organic and inorganic materials tested as PLA fillers.

Graupner et al. (Graupner et al., 2009) prepared composites from different types of natural fibers (cotton, hemp, kenaf) and modified cellulose fibers (Lyocell), with a fiber mass fraction of 40%, by compression molding. The mechanical properties of these composites are summarized in Table 4. Tomé et al. (Tomé et al., 2011) prepared composites from PLA and acetylated bacterial cellulose by mechanical compounding. The composites displayed significant increases in both elastic and Young moduli, as well as in tensile strength (increments of about 100, 40, and 25%, respectively, as compared with neat PLA) at 6% filler loading. Some surface modifiers can enhance adhesion between the fibers and the PLA matrix. For example, 3-aminopropyltriethoxysilane (APS) hydrolyzes in water or solvents to produce silanol groups that are capable of bonding to -OH groups on the kenaf fiber surface (Huda et al., 2008). The $-NH_2$ groups from APS can also bond with $-CO_2^-$ sites formed on the PLA surface by treatment with a sodium hydroxide solution. Thus APS effectively functions as a coupling agent. Yang et al. (Yang et al., 2011) produced a composite from PLA and microcrystalline cellulose modified by *L*-lactic acid. The tensile strength and the elongation at break of the composite were higher than for neat PLA. The surface modification of the cellulose substrates was considered a key element of the mechanical reinforcement.

Type	Filler	Result	Reference
Organic	Jute	Tensile stress and modulus increase with fiber volume fraction	Khondkeret al., 2006
	Flax fibers	Composite strength about 50% higher than for PP/flax composites	Oksman et al., 2003
	Kenaf fibers	Greatly improved crystallization rate, tensile and storage moduli	Pan et al., 2007
	Bamboo fibers	Increased bending strength and improved thermal properties	Tokoro et al., 2008
	Silkworm silk fibers	Good wettability, increased elasticity modulus and ductility	Cheung et al., 2008
	Microcrystalline cellulose	Poor mechanical properties and adhesion; increased storage modulus	Mathew et al., 2005
	LA-modified microcrystalline cellulose	Higher tensile strength and elongation at break than neat PLA	Yang et al., 2011
	Acetylated bacterial cellulose	Considerable improvement in thermal and mechanical properties	Tomé et al., 2011
Inorganic	Calcium metaphosphate	Narrow pore size distribution and high tensile strength	Jung et al., 2005
	Calcium carbonate	No brittle fracture behavior and comparably high bending strength	Kasuga et al., 2003
	Montmorillonite	Good affinity and improved thermal stability of the nanocomposites	Pluta et al., 2002
	HAP	Improved elastic modulus and unchanged bending strength	Kasuga et al., 2001
	Carbon nanotubes	Dramatic enhancement in thermal and mechanical properties	Wu & Liao, 2007
	Nano/Micro-silica	Increased tensile strength, thermal stability, and hydrolysis resistance	Huang et al., 2009

Table 3. Organic and inorganic fillers for the preparation of PLA composites.

	Tensile strength/ (N/mm^2)	Young's modulus/ (N/mm^2)	Elongation at break/ (%)	Charpy impact strength/ (kJ/mm^2)
Pure PLA sample	30.1	3820.2	0.83	24.4
Cottom-PLA	41.2	4242.3	3.07	28.7
Kenaf-PLA	52.9	7138.6	1.05	9.0
Hemp-PLA	57.5	8064.2	1.24	9.5
Lyocell-PLA	81.8	6783.8	4.09	39.7
Hemp/kenaf-PLA	61.0	7763.8	1.22	11.8
Hemp/Lyocell-PLA	71.5	7034.9	1.65	24.7

Table 4. Mechanical properties of composites and a pure PLA sample (mean values; all the specimens were tested at 0°C; adapted in part from (Graupner et al., 2009).

Kim et al. (Kim et al., 2010) prepared a series of PLA/exfoliated graphite (EG) nanocomposites and confirmed that the graphite nanoplatelets could be dispersed homogeneously within the PLA matrix. Thermogravimetric analysis also showed that the thermal stability of the nanocomposites was improved with incremental amounts of EG up to 3 wt %. For example, the temperature corresponding to a 3% weight loss for a composite with 3.0 wt % EG increased by 14 degrees to ~364 °C vs. pure PLA. Additionally, the Young's modulus of the composites increased with their graphite content and their electrical resistivity was dramatically lowered. Poly(lactic acid)/hydroxyapatite (PLA/HAP) composite scaffolds processed by foaming with supercritical CO_2 were shown to be promising for bone replacement, because their mechanical characteristics closely matched the properties of bone in terms of viscoelasticity and anisotropy (Mathieu et al., 2006).

4.1.2 Chemical modification

The chemical modification of PLA has been achieved mainly through copolymerization and cross-linking.

Copolymerization

The carboxyl and hydroxyl groups of LA make it possible to copolymerize it with other monomers through polycondensation with lactone-type monomers such as ε-caprolactone, which generally leads to low molecular weight copolymers, or alternately through the ring-opening copolymerization of lactide with other cyclic monomers including glycolide, δ-valerolactone, and trimethylene carbonate, as well as with monomers like ethylene oxide (EO) to produce high molecular weight copolymers. The hydrophobicity and crystallinity of the copolymers can be increased for low to moderate comonomer contents. Besides, poly(ethylene oxide) (PEO) and PEG have been most commonly copolymerized with PLA to prepare copolymers on account that it is highly biocompatible, hydrophilic and non-toxic, non-immunogenic and non-antigenic (Metters et al., 2000). Such properties reduce protein adsorption and enhance resistance to bacterial and animal cell adhesion.

Block copolymers are composed of long sequences (blocks) of the same monomer unit, covalently bound to sequences of a different type. The blocks can be connected in a variety of ways. Fig. 3 shows examples of block copolymer structures. Diblock PLA-PEG copolymers and triblock PLA-PEG-PLA copolymers allow modulation of the biodegradation rate, the hydrophilicity, and the mechanical properties of the copolymers, while phase separation can be tailored with PLA-PEG multi-block copolymers of predetermined block lengths (Wang et al., 2005). Star- and dendrimer-like PLA-PEG copolymers have also been synthesized to lower the T_g, T_m, and the crystallinity of the materials (Zhang et al., 2004).

Riley et al. (Riley et al., 2001) prepared a range of PLA-PEG copolymers incorporating a PEG block of constant molecular weight (Mn = 5,000) and varying PLA segment lengths (Mn = 2,000-110,000) by ROP of *D*,*L*-lactide catalyzed by stannous octoate; all the dispersions were stable under physiological conditions. In 2003, Li and Vert (Li & Vert, 2003) prepared series of diblock and triblock copolymers by ring-opening polymerization of *L*(*D*)-lactide from mono- and dihydroxyl PEO, using zinc metal as a catalyst under vacuum. The copolymers were semicrystalline, their composition and molar mass being determining factors affecting their solubility in water. Fu et al. (Fu et al., 2008) prepared series of LA-based polyurethanes modified by castor oil with controllable mechanical properties. In this work, hydroxyl-terminated prepolymers were synthesized by copolymerization of L-LA and 1, 4-butanediol.

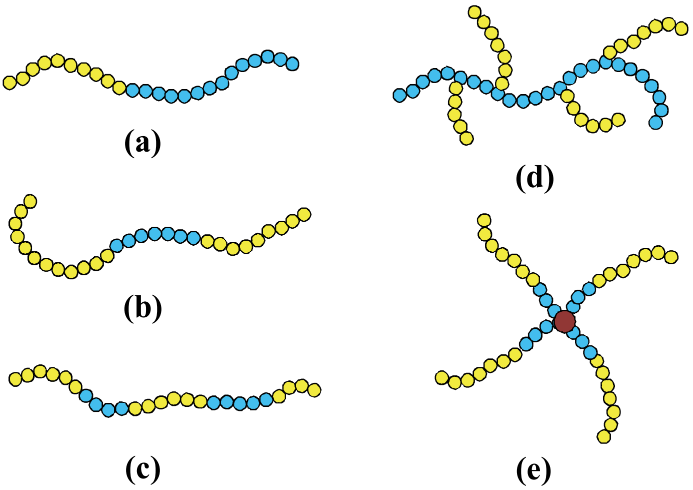


Fig. 3. Schematics of block copolymer structures: (a) diblock; (b) triblock; (c) alternating multiblock; (d); dendrimer-like copolymer; (e) star-like copolymer.

Cross-linking

Cross-linked PLA structures can be formed either by irradiation or through chemical reactions. Electron beam and γ-irradiation have been widely applied to cross-linking PLA in the presence of small amounts of cross-linking agents such as triallyl isocyanurate (TAIC) (Quynh et al., 2008; Phong et al., 2010). The thermal stability of PLA-based materials can be significantly improved in this way (Quynh et al., 2007). Quynh et al. (Quynh et al., 2009) obtained stereocomplexes by cross-linking blends of PLLA and low molecular weight PDLA. Alkaline hydrolysis and enzymatic degradation of the stereocomplex could be controlled by radiation cross-linking, because the alkaline solution as well as proteinase hardly attacked the cross-linked polymer network. Unfortunately, irradiation equipment is expensive and the PLA samples must be processed as thin plates to absorb enough energy from the radiation to initiate cross-linking reactions, which significantly limits its practical application.

Modified PLA with different gel fractions and cross-linking densities can also be obtained through chemical reactions between linking agents and the polymer chains without irradiation (Agrawal et al., 2010). Yang et al. (Yang et al., 2008) thus induced cross-linking via treatment of the PLA melt with small amounts of TAIC and dicumyl peroxide (DCP). The results obtained for samples with different gel fractions and cross-link densities showed that the cross-linking of PLA was initiated at low contents of either TAIC or DCP. The crystallinity of cross-linked PLA samples obtained with 0.5 wt% TAIC and 0.5 wt% DCP decreased from 32% for pure PLA to 24%. Significant increases in tensile modulus from 1.7 GPa to 1.9 GPa, and in tensile strength from 66 GPa to 75 GPa were also observed, and the thermal degradation initiation and completion temperatures were both increased relatively to neat PLA. Additional advantages of this method are that it requires neither extra purification steps nor specialized equipment, since the reaction is carried out in the molten state with only small amounts of cross-linking agent. It is thus economically very advantageous over irradiation, which requires expensive equipment. An increase in brittleness was nevertheless observed following the formation of highly cross-linked structures, which remains a problem to be solved.

4.2 Surface modification

The surface properties of materials play a key role in determining their applications. The presence of specific surface chemical functionalities, hydrophilicity, roughness, surface energy, and topography is crucial for biomedical applications of PLA and its interactions with biomacromolecules. Pure PLA causes a mild inflammatory response if it is implanted into human tissues. It is therefore important to design biomaterials with the required surface properties. The different surface modification strategies examined include physical methods, including surface coating, entrapment and plasma treatment, and chemical methods. Both types of approaches are reviewed.

4.2.1 Physical methods

Surface coating

This is one of the simplest surface modification methods and has been applied to various polymers, but particularly to PLA nanoparticles used for drug delivery. For instance, PEG coating delayed the phagocytosis of PLA nanoparticles and prolonged the circulation time of the nanoparticles in vivo (Gref et al., 1994). Unfortunately the PEG-coated PLA nanoparticles cannot provide specific targeting, which influences their delivery efficiency. One of the most promising alternatives to PEG in this respect is the use of polysaccharides. These materials provide steric protection to the nanoparticles against non-specific interactions with proteins and thereby insure particle stability in the blood circulation system (Ma et al., 2008). Additionally, ligands to achieve active targeting can be conjugated on the surface of these nanoparticles, because many reactive groups are available on the polysaccharides and their derivatives (Gu et al., 2007). Another option is coating of the surface with extracellular matrix (ECM) proteins such as fibronectin, laminin, vitronectin, and collagen, which are conducive to cell adhesion and can greatly improve biocompatibility as well (Lin et al., 2010).

Innovative work was accomplished by Cronin et al. (Cronin et al., 2004), who tested a PLLA fiber scaffold as a substrate for the differentiation of human skeletal muscle cells. Cell attachment (the number of cells attached to the films counted along the center, from one edge to the opposite edge of the film within the field of view) increased significantly on PLLA films coated with ECM gel, fibronectin, or laminin as compared to uncoated or gelatin-coated PLLA films. Myoblasts were able to differentiate into multinucleated myofibers on the ECM gel-coated PLLA fibers and expressed muscle markers such as myosin and α-actinin, as demonstrated by western blot and oligonucleotide microarray analysis.

Entrapment

The entrapment of modifying species (e.g. PEG, alginate, gelatin, etc.) can be achieved through reversible swelling of the PLA surface as illustrated in Fig. 4. This is a simple yet effective method for surface modification requiring no specific functional groups in the polymer chains, as the modifying molecules accumulate merely on the surface of the material without modifying its bulk properties (Lu et al., 2009). Additionally, entrapment can be used to generate different morphologies and thicknesses of 3D scaffolds, which cannot be achieved by other surface modification methods. Finally, entrapment allows the modification of the surface in a controlled fashion because various parameters (e.g. solvent ratio, gelatin concentration, immersion time, and chemical cross-linking) can be varied to tailor the process (Zhu et al., 2003).

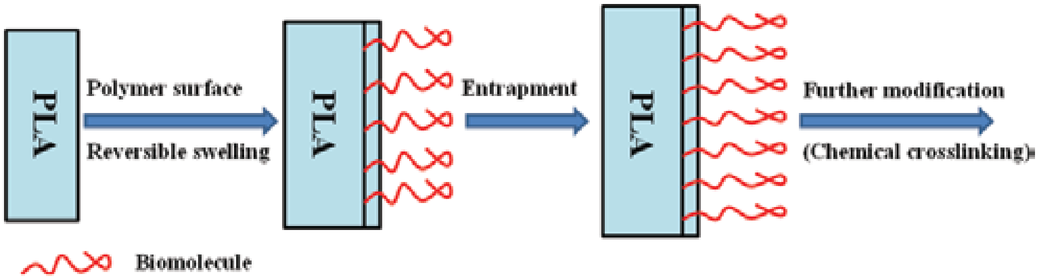


Fig. 4. Schematic illustration of entrapment process.

PEG (M_w = 18,500) and poly(*L*-lysine) (PLL) (M_w = 29,300) have been trapped on PLA surfaces using 2,2,2-trifluoroethanol (TFE)/water as solvent/nonsolvent mixtures (Quirk et al., 2002). A new entrapment process has also been reported by Liu et al. (Liu et al., 2005), through chemical cross-linking of gelatin with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) HCl and *N*-hydroxysuccinimide (NHS) (97%) in {2-[N-morpholino] ethanesulfonic acid} (MES) hydrate buffer, after the pretreated PLLA films were immersed in the gelatin solution for a set time. Results in comparison to the control scaffolds have shown that the surface hydrophilicity increased with the amount of entrapped gelatin and that cell attachment and proliferation, the deposition of collagen fibers, and other cell excretion (extracellular matrix, etc.) were also significantly improved.

Plasma treatment

Tests with plasma treatment were initiated in the 1960s and have been since then widely utilized to improve the hydrophilicity and cell affinity of PLA surfaces. The obvious advantages of plasma treatment as compared to other surface modification methods include its ability to control the surface structure, energy and charge, and to uniformly modify the surface without impacting bulk properties (Chu et al., 2002). Functional groups such as $-NH_2$, $-COOH$, and $-OH$, which are apt to form covalent bonds with other materials for further modification, are most frequently introduced by plasma treatment (Favia et al., 1998).

Liu et al. thus investigated the influence of the main operation parameters, namely the plasma power, the treatment duration (number of treatment cycles) and the electrode gap on a dielectric barrier discharge (DBD) plasma treatment of PLA films in terms of changes in surface wettability and chemistry (Liu et al., 2004). They further developed equations relating the surface properties (water contact angle and oxygen enrichment, as observed by XPS analysis) to these operational parameters. It was determined that the magnitude of the electrode gap played a dominant role in the treatment of PLA, and the observed wettability improvements were attributed to changes in both surface chemistry and microstructure. Chaiwong and coworkers (Chaiwong et al., 2010) investigated the influence of SF_6 plasma on the hydrophobicity and barrier properties of PLA. It was found that the SF_6 plasma enhanced the hydrophilicity and increased the water absorption time of PLA two-fold. Plasma treatment did not have any significant influence on the water vapor permeability of PLA, however, since the bulk structure controlling the transport properties are unaffected by the treatment. Other types of plasmas such as oxygen, helium, and nitrogen plasmas have also been investigated (Hirotsu et al., 2002).

While plasma treatment has been successfully applied to improving PLA wettability and cell affinity, its main disadvantage is that surface rearrangements caused by thermally activated macromolecular motions, to minimize its interfacial energy, can also influence the surface modification. Moreover, the potential influence of plasma on the degradation of PLA cannot be ignored.

4.2.2 Chemical modification

PLA does not carry reactive side-chain functional groups. Consequently, the first step of chemical modification is typically a simple surface hydrolysis (with an alkali) or an aminolysis treatment. The hydrophilic –COOH and –OH or reactive $–NH_2$ groups introduced by cleavage of the ester bonds can be used to bind bioactive molecules such as arginine-glycine-aspartic acid (RGD)-containing peptides, chitosan (CS), arginine and lysine, PEG, collagen, and so on to regulate cell adhesion or protein adsorption.

The synthetic RGD-containing peptides could be immobilized on PLA after treatment by hydrolysis or aminolysis (Stupack et al., 2001). Materials prepared by this method provide suitable recognition sites for cell adhesion receptors and biodegradation rates, making them suitable for various applications in fields such as tissue engineering and implant technology. It has also been determined that RGD-conjugated poly(lactic acid-*co*-lysine)(arginine-glycine-aspartic acid) nanoparticles (PLA-PLL-RGD NP) are non-toxic and bind more efficiently to human umbilical vein endothelial cells (HUVECs) as compared to bare PLA-PLL NP in vitro. Targeted imaging results obtained in vivo showed that PLA-PLL-RGD can selectively bind to BACP-37 breast cancer cells. Lieb et al. also demonstrated largely increased cell densities and cell proliferation on surfaces modified with RGD-anchored monoaminated poly(ethylene glycol)-*block*-poly(*D*,*L*-lactic acid) (H_2N-PEG-PLA), mediated through RGD–integrin interactions (Lieb et al., 2005).

Chitosan (CS) is a biopolymer displaying good biocompatibility, non-toxicity and biodegradability, produced by the alkaline *N*-deacetylation of chitin. The immobilization of this polymer on PLA has been accomplished by coating the surface with chitosan, modified with the photosensitive hetero-bifunctional cross-linking reagent 4-azidobenzoic acid, and irradiation with ultraviolet light to photolyze the azide groups and covalently link the two polymers (Zhu et al., 2002). The -OH and $-NH_2$ groups of chitosan provide further opportunities to introduce a wide range of functional groups on the surface. Thus CS molecules immobilized on PLA were modified with a heparin (Hp) solution to form a polyelectrolyte complex on the surface, which inhibited platelet adhesion and activation, and enhanced cell adhesion.

4.3 Outlook of PLA surface modification

The surface attributes of PLA can be tailored to enhance its hydrophilicity and biocompatiblity through various methods. Unfortunately, all these established methods for surface modification are inherently flawed to some extent. For example, a single plasma treatment can merely improve cell adhesion but cannot accelerate cell growth; non-covalent attachment of a functional material onto a PLA surface is not stable and permanent. An excellent method suggested to solve the second issue is the use of 1,6-hexanediamine for surface aminolysis, followed by conjugation with biocompatible macromolecules such as gelatin, chitosan, or collagen (Zhu et al., 2004). Hong et al. have shown that chondrocyte cells could attach, proliferate, and spread on PLA microspheres coated with collagen in the same way as described above, in particular those having high collagen contents (Hong et al., 2005). It appears that the surface modification of PLA would be best achieved with a combination of distinct approaches, to benefit from the advantages of all the methods. Polysaccharide polyelectrolyte multilayers, including chitosan and dextran sulfate-stabilized silver nano-sized colloids developed by Yu et al. (Yu et al., 2007), were successfully deposited on an aminolyzed PLA membrane in a layer-by-layer self-assembly manner. This seemingly easy process resulted in significant improvements in hydrophilicity, antibacterial activity, hemocompatibility, and cytocompatibility for the PLA membrane, thanks to the different attributes of $-NH_3^+$ (positive charge), chitosan (biocompatibility), and silver nanoparticles (antibacterial activity). The radiation-induced methods are emerging as powerful surface modification techniques, particularly when relying on PLA photoactivation to create reactive groups or moieties useful to graft specific chemical functionalities. The irradiation of PLA with UV (ozone can be generated from molecular oxygen irradiated with UV in this process), for example, is known to increase fiber adhesion to high surface energy components due to the introduction of photo-oxidized polar groups on the surface (Koo & Jang, 2008). Irradiation followed by grafting has also been used extensively to alter PLA surface characteristics, mainly due to the advantages it offers, namely a low operation cost, mild reaction conditions, selectivity to UV light, and the permanent surface chemistry changes induced (Ma et al., 2000).

5. Applications of PLA in the biomedical field

Due to its bioresorbability and biocompatibility in the human body, PLA has been employed to manufacture tissue engineering scaffolds, delivery system materials, or covering membranes, different bioabsorbable medical implants, as well as in dermatology and cosmetics.

5.1 Tissue engineering

Since the introduction of the concept in 1988, tissue engineering, a technique invented to reconstruct living tissues by associating the cells with biomaterials that provide a scaffold on which they can proliferate three-dimensionally and under physiological conditions, has emerged as a potential alternative to tissue or organ transplantation and has thus attracted great attention in science, engineering, and medicine. To meet the diverse needs of tissue engineering, scaffolds made from various materials have been tested in this field. Although certain metals are somewhat good choices for medical implants due to their superior mechanical properties, their lack of degradability in a biological environment makes them disadvantageous for scaffold applications (Liu & Ma, 2004). Inorganic/ceramic materials such as HAP or calcium phosphates, being studied for mineralized tissue engineering with good osteoconductivity, are also limited due to poor processability into porous structures (Ilan & Ladd, 2002). In contrast, polymers have great design flexibility because their composition and structure can be tailored to meet specific needs (Huang et al., 2007). Degradable polymers frequently used for tissue engineering applications are linear aliphatic polyesters such as PGA, PLA, and their copolymers (PLGA), which are fabricated into scaffolds. These polymers are among the few synthetic polymers approved by the FDA for human clinical applications. The drawbacks of these polyesters include their hydrophobicity and lack of functional groups, which limits cell adhesion, an important factor when constructing polymeric scaffolds. Another drawback is their slow hydrolytic degradation (Iwata & Doi, 1998).

An ideal scaffold used for tissue engineering should possess the following properties: 1) Be biocompatible, so that the scaffold can be well integrated into host tissues without resulting in any immune response; 2) It should be porous with appropriate pore size, size distribution and mechanical function, to allow cell or tissue growth and the removal of metabolic waste; 3) It must be mechanically able to withstand local stress and maintain the pore structure for tissue regeneration; 4) Very importantly, the scaffold should be biodegradable (Ma, 2004). Synthetic scaffolds are considered important components of a successful tissue engineering strategy (Wang, 2007). Hybrid three dimentional porous scaffolds of synthetic and naturally derived biodegradable polymers are particularly promising because they combine the advantages of the two types of materials. They should maintain sufficient mechanical strength while providing specific cell-surface receptors during the tissue remodeling process that stimulate both in vitro and in vivo cell growth (Chen et al., 2002). PLA-based hybrid materials have been successfully tested clinically for that purpose so far, and tests on other tissues including bladder (Engelhardt et al., 2011), cartilage (John et al., 2003), liver (Lv et al., 2007), adipose (Mauney et al., 2007), and bone tissues (Mathieu et al., 2006) have also been reported.

Jiang et al. (Jiang et al., 2010) functionalized chitosan/PLGA by heparin immobilization with controlled loading efficiency. One of the main benefits of introducing chitosan into PLGA microspheres is that chitosan imparts functionality due to its reactive amino groups, so that biomolecules such as heparin could be attached (Jiang et al., 2006). The compressive strength and modulus remained in the range of human trabecular bone after the heparinization process. More importantly, heparinized chitosan/PLGA scaffolds with a low heparin loading (1.71 g/scaffold) showed a stimulatory effect on cell differentiation, as indicated by enhanced osteocalcin expression as compared with a non-heparinized chitosan/PLGA scaffold. Based on these results, Jiang et al. (Jiang et al., 2006) continued to evaluate the novel scaffolds for bone regeneration in vivo. In the rabbit ulnar critical-sized-defect model created, successful bridging of the critical-sized defect on the sides both adjacent to and away from the radius occurred using chitosan/PLGA-based scaffolds. However, the addition of chitosan to PLGA led to somewhat higher inflammation and lower mineralization than for the PLGA counterpart, which is a major problem that remains to be solved.

Three-dimensional (3D) electrospun fibrous scaffolds have been suggested as a potential tissue engineering tool for bone regeneration. Shim et al. (Shim et al., 2010) thus reported a 3D microfibrous PLLA scaffold fabricated using electrospinning techniques with a subsequent mechanical expansion process. The use of these 3D scaffolds for the proliferation of osteoblasts was examined. The 3D scaffolds led to a 1.8-fold higher level of osteoblast proliferation than generally achieved for electrospun 2D nanofibrous membranes. In vivo results further showed that 3D electrospun microfibrous matrices provided a favorable substrate for cell infiltration and bone formation after 2 and 4 weeks when using a rabbit calvarial defect model.

3D printing technology has rapidly expanded in the tissue engineering field since it was first developed at the Massachusetts Institute of Technology. Ge et al. (Ge et al., 2009) developed 3D-printed poly(lactic acid-co-glycolic acid) (PLGA) scaffolds which could support the proliferation and osteogenic differentiation of osteoblasts. Based on their in vitro study, they also evaluated PLGA scaffolds for bone regeneration within a rabbit model (Ge et al., 2009). In both the intra-periosteum and the iliac bone defect models, the implanted scaffolds facilitated new bone tissue formation and maturation over a time period of 24 weeks.

The current clinical use of PLA-based scaffolds nevertheless remains very limited (Iwasa et al., 2009), mainly because of the risk of disease transmission and immune response. This can be illustrated by taking cartilage tissue engineering as an example. Traditional autologous chondrocyte implantation (ACI), first introduced by Brittberg et al. in 1994 (Brittberg et al., 1994), has yielded good clinical results (Bentley et al., 2003). To date, none of the short- or mid-term clinical and histological results using scaffolds were reported to be better than ACI. As for the scaffolds, collagen and hyaluronan-based matrices are among the most popular scaffolds in clinical use nowadays, since they offer substrates which are normally essential elements in native articular cartilage (Iwasa et al., 2009). Among the very few cases of scaffolds in clinical use is the copolymer of PGA/PLA (polyglactin, vicryl) and polydioxanone, which is used for cartilage repair under the trade name of BioSeed®-B and BioSeed®-C (Biotissue Technologies AG, Freiburg, Germany).

In summary, tissue engineering is one of the most exciting interdisciplinary fields today and is growing rapidly with time. The inclusive criteria for studies on scaffolds capable of clinical application were in vivo or clinical studies and thus certain artificially designed scaffold features (such as pore size, interpore connectivity, etc.) are necessary for optimal tissue engineering applications (accelerated tissue regeneration). Suggestions for future directions include the use of designer scaffolds with in vivo experimentation, and coupling scaffold design with cell printing to create designer material/biofactor hybrids to optimize tissue engineering treatments (Hollister, 2005).

5.2 Delivery systems

There has long been a desire to achieve the targeted delivery of bioactive compounds to areas in the body to maximize therapeutic potential and minimize side-effects. Many types of particles have been tested as delivery tools for biomedical applications such as liposomes, solid lipid nanoparticles, and biodegradable polyesters like PLA and PLGA (Torchilin, 2006). With its excellent biocompatibility, biodegradability, mechanical strength, heat processability, and solubility in organic solvents, PLA can be used to produce dosage forms such as pellets, microcapsules, microparticles (MP), nanoparticles (NP), etc. MP and NP of PLA, modified or unmodified, are increasingly investigated for sustained release and targeted drug, peptide/protein, and RNA/DNA delivery applications because of their small size enabling their permeation through biological barriers such as the blood-brain barrier (Roney et al., 2007).

Although PLA-based materials such as PLGA have been FDA-approved and are clinically available, they lack chemical functionalities to facilitate specific cell interactions. Furthermore, their potential for the sustained release of hydrophilic molecules (e.g. proteins) is often limited (Fahmy et al., 2005). Frequent undesired effects include low encapsulation efficiency and high burst release of the encapsulated biomolecule within the first few hours or days, which is mainly due to the desorption of surface-associated hydrophilic molecules having weak interactions with the polymer (Fahmy et al., 2005). To circumvent these limitations and establish therapeutic efficacy, large doses or site-specific administration are often required for devices comprised of polyester biomaterials. In an attempt to address these problems, numerous groups have introduced functional groups (such as amine functionalities) on these materials, either through direct conjugation or device fabrication with additives (Betram et al., 2009).

As for drug release from MPs or NPs, it is generally controlled by both drug diffusion and polymer degradation. To ensure the efficacy of drug delivery, control over the particle size and particle size distribution is critical, since smaller particles and narrower size distributions facilitate the design of targeted drug delivery systems. These involve binding fragments specific to a tumor-associated surface antigen, with a ligand binding to its corresponding receptor on the tumor cell surface, which can be attached on the surface of the PLA-based materials. Furthermore, polymers that display a physicochemical response to changes in their environment are being intensively explored as potential drug and gene delivery systems. The use of stimuli-responsive nanocarriers offers an attractive opportunity for targeted delivery, in which the delivery system becomes an active participant rather than a passive vehicle. The advantage of stimuli-responsive nanocarriers becomes obvious when the stimuli are unique to disease pathology, allowing the nanocarrier to respond specifically to the pathological characteristics. For instance, in solid tumors, the extracellular pH decreases significantly from 7.4 (the pH value under normal physiological conditions) to about 6.5 (Vaupel et al., 1989; Haag, 2004). In addition, the pH ranges from 4.5 in lysosomes to about 8.0 in mitochondria. Given these pH shifts, therapeutic compounds with a pK_a between 5.0 and 8.0 are able to exhibit dramatic changes in physicochemical properties (Ganta et al., 2008). Another option is thermo-sensitive polymeric micelles, containing a hydrophobic core and a thermo-sensitive shell, the later changing from a hydrophilic nature at body temperature to a collapsed hydrophobic polymer at a hyperthermic condition of 42°C (Na et al., 2006). Investigations concerned with this theme include responses induced by chemical substances, changes in temperature (Tyagi et al., 2004) or pH (Sethuraman & Bae, 2007), electric signals (Sawahata et al., 1990) or other environmental conditions (Qiu & Park, 2001).

The use of nucleic acids as therapeutic agents for genetic diseases has been extensively studied (Torchilin, 2008). However, a major limitation of this technique lies in the low delivery efficiency of the therapeutic DNA to the diseased site. To address this issue, various strategies have been explored including vectors engineered from viruses (Brun et al., 2008) and PLGA in NP formulations. PLGA NP have shown particular promise in improving the delivery efficacy (Kocbek et al., 2007). Besides, the physical characteristics of the nanoparticles can be manipulated to escape the degradative endosomal lumen, resulting in cytosolic localization. To develop novel administration paths, hybrid versions of research have been conducted on this subject, yet the results are mostly based on animal models or in vitro results, making it difficult to draw final conclusions. From clinical trials, substantial obstacles to their use, such as immunogenicity and inflammatory potential, have also been demonstrated (Nafee et al., 2007). Therefore, there is still a long way to go before real clinical applications come through.

Some examples of delivery systems incorporating PLA are provided in Table 5 and in Fig. 5 (Chen et al., 2007; Sethuraman & Bae, 2007). Sethuraman et al. (Sethuraman & Bae, 2007) developed a novel drug delivery system for acidic tumors consisting of two components: 1) A polymeric micelle with a hydrophobic core of PLLA and a hydrophilic shell of PEG conjugated to TAT (a cell-penetrating peptide in HIV), and 2) a highly pH-sensitive copolymer of poly(methacryloyl sulfadimethoxine) (PSD) and PEG (PSD-b-PEG). The final carrier, which was able to shield the micelles and expose them at slightly more acidic tumor pH levels, was achieved by complexing PSD with the TAT of the micelles. The results obtained showed significantly higher uptake of TAT micelles at pH 6.6 in comparison with pH 7.4, and that TAT not only translocated into the cells but it was also traced to the surface of the nucleus [see Fig. 5].

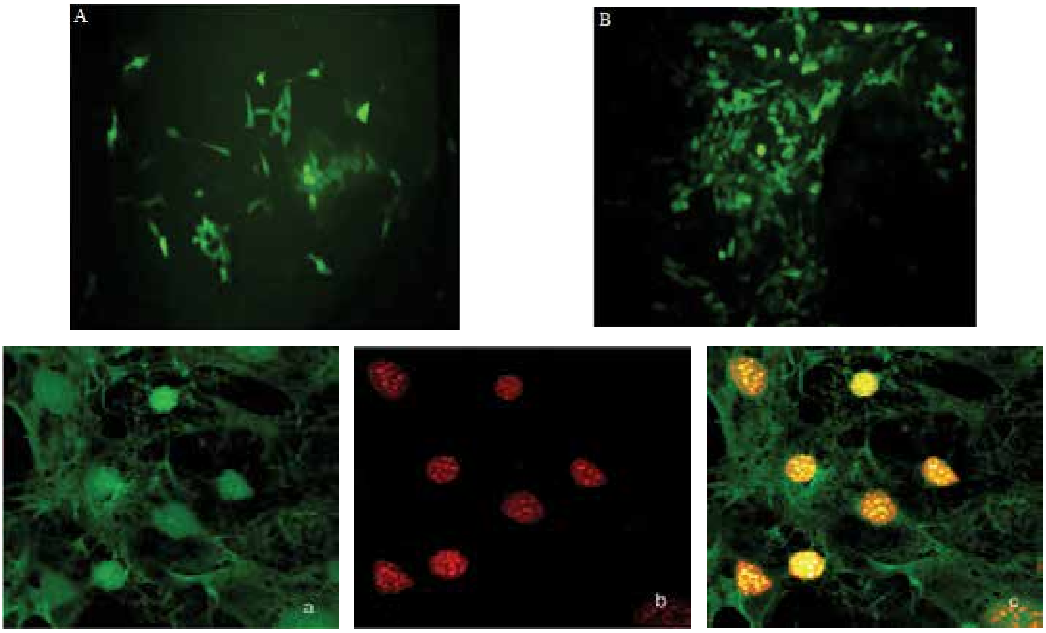


Fig. 5. Test results for PLA-based drug delivery materials. Fluorescent microscopy images are shown on top for COS7 cells transfected by plasmid encoding enhanced green fluorescence protein (EGFP) with different carriers: (A) lipofectamine, (B) methoxypolyethyleneglycol-PLA-chitosan nanoparticles (MePEG-PLA-CS NP); the transfection efficiency, as detected by flow cytometry, is higher in (B) than in (A) (Reproduced with the permission from Chen, J. et al. (2007). Preparation, characterization and transfection efficiency of cationic PEGylated PLA nanoparticles as gene delivery systems, Journal of Biotechnology, Vol.130, No.2, pp.107. Copyright (2007) Elsevier) At the bottom are dual label confocal micrographs for MCF-7 cells incubated with TAT micelles: (a) Cells stained with fluorescein isothiocyanate (FITC) attached to TAT in the micelles; (b) the same nuclei as in (a) were stained with TOPRO-3; (c) superimposed images of (a) and (b); the yellow color shows the localization of TAT within the nuclei (Reproduced with the permission from Sethuraman, V. A. & Bae, Y. H. (2007). TAT peptide-based micelle system for potential active targeting of anti-cancer agents to acidic solid tumors. Journal of Controlled Release, Vol.118, No.2, pp.216. Copyright (2007) Elsevier).

Material	Application	Results	Reference
PLA-PEG particles	Carrier for tetanus toxoid	Enhanced transport across the rat nasal mucosa	Vila et al., 2005
PEG-PLA NP	Conjugated with lactoferrin (Lf)	Increased uptake of the Lf-NP by bEnd.3 cells	Hu et al., 2009
PLA-b-Pluronic-b-PLA	Carrier for oral insulin	Good control over blood glucose concentration	Xiong et al., 2007
PLA NP	Carrier for HIV p24 proteins	Induced seric and mucosal antibody production	Aline et al., 2009
Surfactant-free PLA NP	Carrier for HIV p24 proteins	Elicited strong CTL response and cytokine release	Liggins et al., 2004
PLA microspheres	Carrier for paclitaxel	Reduced inflammation of arthritis rabbit model	Jie et al., 2005
PEO-PLA copolymers	Carrier for 5-FU and paclitaxel	Complete drug release	Zhang & Feng, 2006
PLA-TPGS copolymers	Carrier for paclitaxel	Initial burst followed by sustained release	Freitas et al., 2005
PLA microspheres	Carrier for nimesulide	Initial burst followed by an exponential decrease	Chen et al., 2007
PEGylated PLA NP	Gene delivery systems	Improved transfection activity	Ataman-Önal et al., 2006
PLA-PEG-PLA copolymer	Carrier for 5-FU and paclitaxel	Good control over the release	Venkatraman et al., 2005
AP-PEG-PLA/MPEG-PAE	Drug carrier for cancer therapy	Presented high tumor-specific targeting ability	Wu et al., 2010
PLGA/PEI NP	Carrier for luciferase siRNA	Effective silencing of the gene in cells	Patil & Panyam, 2009
cNGR-PEG-PLA NP	Carrier for DNA	Rapid and efficient nanoparticle internalization	Liu et al., 2011
DMAB coated PLGA NP	Loaded with plasmid DNA	Significantly improved transfection efficiencies	Fay et al., 2010

Table 5. Investigations on PLA-based material as drug delivery systems. AP: peptide, CRKRLDRN; MPEG: methyl ether poly(ethylene glycol); PAE: poly(β-amino ester); PEI: polyethylenimine; cNGR: Cyclic Asn-Gly-Arg; DMAB: dimethyldidodecylammonium bromide.

In summary, some problems still remain to be tackled for this promising novela dministration method. A major issue is the presence of surfactants such as SDS or stabilizers such as PVA in the microparticles, necessary to achieve antigen binding and colloidal stability. Although present only at low concentrations, the acceptability of such components in human vaccines depends on the results of extensive and costly toxicological studies. Biodegradable polymers used for drug delivery to date have mostly been in the form of injectable microspheres or implant systems requiring complicated fabrication processes with organic solvents. In such systems, the organic solvents can denature components such as protein drugs being encapsulated. Besides, these delivery systems have relatively low transfection efficiencies in vitro as compared with reagents commercialized for cell transfection. The last problem concerns the lack of test results for these delivery systems using animal models or in clinical trials, which is of fundamental importance for real applications in biomedical therapy.

5.3 Other fields

Due to its versatility, PLA has been investigated for membrane applications (e.g. wound covers), implants and medical devices (fixation rods, plates, pins, screws, sutures, etc.), and dermatological treatments (e.g. facial lipoatrophy and scar rejuvenation).

With respect to wound treatment, bacterial infections are one of the main factors impacting the healing process. One of the best approaches to treat wound infections is by the immobilization of drugs or antibacterial agents within the nanofibers by electrospinning, or the electrospinning of polymers with intrinsic antibacterial and wound-healing properties. Dozens of patents have been issued on that topic so far (Ghosh et al., 2007; Robinson et al., 2009). Silver nanoparticles (nAg) and the natural polysaccharide chitosan (as well as its quaternized derivatives) are most commonly used as antibacterial agents with a high intrinsic activity against a broad spectrum of bacteria (Rujitanaroj et al., 2008; Ignatova et al., 2009).

Metals are still the most popular materials for fracture fixation, but their disadvantages include stress shielding, accumulation in tissues, hypersensitivity, growth restriction, pain, corrosion, and interference with imaging techniques. Consequently, the focus of research is increasingly shifted to biomaterials like PLA, which offers satisfactory strength during the healing of bone tissue and then degrades over time (Mavrogenis et al., 2009). A commercial product with a proven track record in clinical applications is the VICRYL™ suture material, based on PGA/PLA copolymers (Mehta et al., 2005). The number of applications of PLA as fixation rods, plates, pins, screws, sutures, etc. in orthopaedics and dentistry is also increasing (Raghoebar et al., 2006).

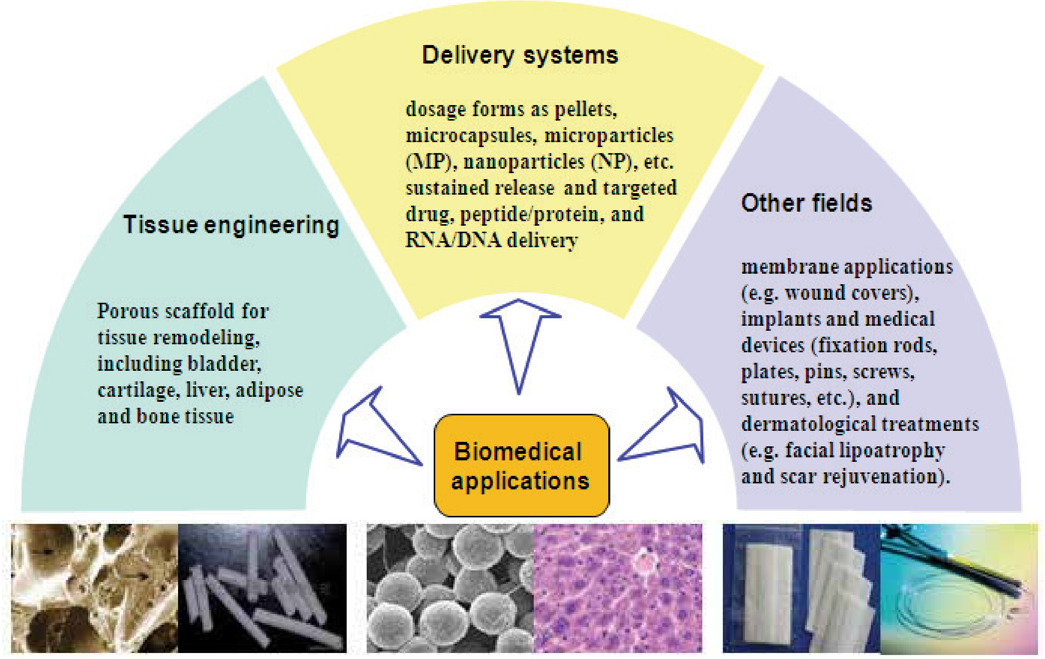

Delivery systems
dosage forms as pellets, microcapsules, microparticles (MP), nanoparticles (NP), etc. sustained release and targeted drug, peptide/protein, and RNA/DNA delivery
Tissue engineering
Porous scaffold for tissue remodeling, including bladder, cartilage, liver, adipose and bone tissue
Other fields
membrane applications (e.g. wound covers), implants and medical devices (fixation rods, plates, pins, screws, sutures, etc.), and dermatological treatments (e.g. facial lipoatrophy and scar rejuvenation).
Biomedical applications

Fig. 6. Applications of PLA in the biomedical field.

In 2004, Sculptra™ [poly(*L*-lactic acid)] was approved by the FDA as the first injectable facial "volumizer" in the treatment of lipoatrophy due to its significant therapeutic effectiveness (Burgess & Quiroga, 2005). The lipodystrophy syndrome is associated with the usage of highly active antiretroviral therapy (HAART) containing protease inhibitors or nucleoside reverse transcriptase inhibitors for HIV patients. The action mechanism of Sculptra™ is via stimulation of the fibroblastic activity with generation of collagen and other connective tissue fibers. In addition, it acts as dermal matrix adding support by thickening the dermis (Vleggaar & Bauer, 2004). Moreover, PLA can help improve the appearance of scars due to acne, surgery, trauma, or suture (Lowe & Beer, 2005, as cited in Beer & Rendon, 2006).

6. Conclusions

Due to the multiple desirable characteristics of PLA including renewability, biocompatibility, transparency, and thermoplasticity, it is being used or is a potential candidate for many consumer and biomedical applications (Jamshidian et al., 2010). Ever increasing environmental concerns associated with conventional polymers derived from petrochemicals lead to constantly expanding applications for PLA since its discovery in 1932 by Carothers at DuPont.

In previous years, the most negative point of PLA was its higher price as compared with petrochemical-based polymers. Today, by optimizing the LA and PLA production processes, and with increasing PLA demand, a reduction in its price can be achieved. The price of PLA is currently much lower than in previous years. Meanwhile, PLA is mainly synthesized in the industry by ROP employing tin(II) bis(2-ethylhexanoate) ($SnOct_2$) as a catalyst, which has been approved as a food additive by the FDA, but the potential toxicity associated with most tin compounds cannot be ignored for biomedical applications. Scientists all over the world are now exploring novel, well-defined catalysts with good biocompatibility, high catalytic activity, low toxicity, and excellent stereoselectivity. This should remain an everlasting interest area. Finally, the possibility of obtaining 100% bio-sourced opens the way for PLA to become more independent from petrochemical-based polymers, free of environmental and health concerns.

However, the major disadvantages of PLA such as its poor ductility, slow degradation rate, and poor hydrophilicity somewhat limit its applications. The modification of PLA bulk and surface properties has thus become crucial to increase its applicability. Many of the bulk and surface modification strategies discussed above have been designed to tune the PLA surface properties according to the demands of biomedical applications. Unfortunately, all these established methods for surface modification are somewhat deficient and while they provide control over the wettability, degradation rate, and functionality, it is still compulsory to minimize their negative impact on PLA bulk properties. Thus a combined modification strategy (e.g. irradiation followed by grafting) or a better balance of PLA surface and bulk properties should be sought. Ideally, with respect to a better balance of properties and shorter modification times, one-step approaches need to be developed because it is time-consuming to carry out surface and bulk modifications separately, and the solvents and reagents involved in multiple modification steps tend to affect PLA properties significantly.

All these modification strategies aim at tailoring the properties of PLA-based materials for certain applications. Fortunately, more and more encouraging results have been reported, but the present conclusions from most of these reports cannot be directly generalized to truly biomedical applications since most of the experiments were carried out in vitro. Nevertheless, these findings offer some clues for further improvements. The increasing number of functional PLA-based polyesters provides the opportunity to study the relationships between structure and functionality of these polymers such as cell adhesion and degradability in vitro and in vivo, as well as to develop applications of these materials for delivery system in the form of micro- and nanoparticles or scaffolds for tissue engineering. Finally, cancers and acquired or inherited genetic diseases represent one of the most serious threats to the health of human beings, but no effective therapies are available so far. It is suggested that the development of effective and ideal tools for drug, peptide/protein, and RNA/DNA delivery will represent a good alternative in drug development. It therefore appears that it would be best to focus future research work on the rational design of novel carriers for biomedical uses and targeted delivery systems. Obviously, this requires plenty of relevant experiments on animal model and enough clinical trials before they are widely utilized.

Even though countless studies have focused on the synthesis and the modification of PLA and remarkable progresses has been achieved over the last two decades, vast opportunities as well as challenges remain in terms of exploring the characteristics of PLA-based materials and expanding their domains of applications.

7. Acknowledgement

It was supported by the Natural Science of Hubei Province for Distinguished Young Scholars (No. 2008CDB279) and the Key Technologies R & D Program of Hubei Province (20091933), the Fundamental Research Funds for the Central Universities, Huazhong University of Science and Technology (2010JC016) and the National High Technology Research and Development Program ("863" Program) of China (No. 2008AA10Z339). The authors are also grateful to Dr. Wen-bin Zhang from the University of Akron for his valuable suggestions during the preparation of this manuscript.

8. References

Achmad, F. et al., (2009). Synthesis of polylactic acid by direct polycondensation under vacuum without catalysts, solvents and initiators. *Chemical Engineering Journal,* Vol.151, No.1, (August 2009), pp. 342-350, ISSN 1385-8947

Agrawal, S. K. et al., (2000). Rare earth metal initiated ring-opening polymerization of lactones. *Macromolecular Rapid Communications*, Vol.21, No.5, (March 2000), pp. 195-212, ISSN 1022-1336

Agrawal, S. K. et al., (2010). Energetics of association in poly(lactic acid)-based hydrogels with crystalline and nanoparticle-polymer junctions. *Langmuir,* Vol.26, No.22, (October 2010), pp. 17330-17338, ISSN 0743-7463

Ajioka, M. et al. (1995). The basic properties of poly(lactic acid) produced by the direct condensation polymerization of lactic acid. *Journal of Polymers and the Environment,* Vol.3, No.4, (October 1995), pp. 225-234, ISSN 1566-2543

Aline, F. et al., (2009). Dendritic cells loaded with HIV-1 p24 proteins adsorbed on surfactant-free anionic PLA nanoparticles induce enhanced cellular immune responses against HIV-1 after vaccination. *Vaccine,* Vol.27, No.38, (August 2009), pp. 5284-5291, ISSN 0264-410X

Ataman-Önal, Y. et al., (2006). Surfactant-free anionic PLA nanoparticles coated with HIV-1 p24 protein induced enhanced cellular and humoral immune responses in various animal models. *Journal of Controlled Release,* Vol.112, No.2, (May 2006), pp. 175-185, ISSN 0168-3659

Bentley, G. et al., (2003). A prospective, randomized comparison of autologous chondrocyte implantation versus mosaicplasty for ostechondral defects of the knee. *Journal of Bone and Joint Surgery-British Volume,* Vol.85, No.2, (March 2003), pp. 223-230, ISSN 0301-620X

Bergsma, J. E. et al., (1995). Late degradation tissue response to poly(L-lactide) bone plates and screws. *Biomaterials,* Vol.16, No.1, (1995), pp. 25-31, ISSN 0142-9612

Bertram, J. P. et al., (2009). Functionalized poly(lactic-co-glycolic acid) enhances drug delivery and provides chemical moieties for surface engineering while preserving biocompatibility. *Acta Biomaterialia,* Vol.5, No.8, (October 2009), pp. 2860-2871, ISSN 1742-7061

Brittberg, M. et al., (1994). Treatment of deep cartilage defects in the knee with autologous chondrocyte transplantation. *New England Journal of Medicine,* Vol.331, No.14, (October 1994), pp. 889-895, ISSN 0028-4793

Brun, A. et al., (2008). Antigen delivery systems for veterinary vaccine development. Viral-vector based delivery systems. *Vaccine,* Vol.26, No.51, (December 2008), pp. 6508-528, ISSN 0264-410X

Burgess, C. M. & Quiroga, R. M. (2005). Assessment of the safety and efficacy of poly-L-lactic acid for the treatment of HIV-associated facial lipoatrophy. *Journal of the American Academy of Dermatology,* Vol.52, No.2, (February 2005), pp. 233-239, ISSN 0190-9622

Butterwick, K. & Lowe, N. J. (2009) Injectable poly-L-lactic acid for cosmetic enhancement: Learning from the European experience. *Journal of the American Academy of Dermatology*, Vol.61, No.2, (August 2009), pp. 281-293, ISSN 0190-9622

Chaiwong, C. et al., (2010). Effect of plasma treatment on hydrophobicity and barrier property of polylactic acid. *Surface and Coatings Technology,* Vol.204, No.18-19, (June 2010), pp. 2933-2939, ISSN 0257-8972

Chen, G.; Ushida, T. & Tateishi, T. (2002). Scaffold design for tissue engineering. *Macromolecular Bioscience,* Vol.2, No.2, (February 2002), pp. 67-77, ISSN 1616-5187

Chen, H. Y. et al., (2011). Comparative study of lactide polymerization by zinc alkoxide complexes with a β-diketiminato ligand bearing different substituents. *Journal of Molecular Catalysis A: Chemical,* Vol.339, No.1-2, (April 2011), pp. 61-71, ISSN 1381-1169

Chen, J. et al., (2007). Preparation, characterization and transfection efficiency of cationic PEGylated PLA nanoparticles as gene delivery systems. *Journal of Biotechnology,* Vol.130, No.2, (June 2007), pp. 107-113, ISSN 0168-1656

Cheung, H. Y. et al., (2008) A potential material for tissue engineering: Silkworm silk/PLA biocomposite. *Composites Part B: Engineering,* Vol.39, No.6, (September 2008), pp. 1026-1033, ISSN 1359-8368

Chisholm, M. H. et al., (2003). Binolate complexes of lithium, zinc, aluminium, and titanium; preparations, structures, and studies of lactide polymerization. *Dalton Transactions,* No.3, (July 2003), pp. 406-412, ISSN 1477-9226

Chu, P. K. et al., (2002). Plasma-surface modification of biomaterials. *Materials Science and Engineering: R: Reports,* Vol.36, No.5-6, (March 2002), pp. 143-206, ISSN 0927-796X

Cronin, E. M. et al., (2004).Protein-coated poly(L-lactic acid) fibers provide a substrate for differentiation of human skeletal muscle cells. *Journal of Biomedical Materials Research Part A,* Vol.69A, No.3, (June 2004), pp. 373-381, ISSN 1549-3296

Deng, X. et al., (2000). Polymerization of lactides and lactones VII. Ring-opening polymerization of lactide by rare earth phenyl compounds. *European Polymer Journal,* Vol.36, No.6, (June 2000), pp. 1151-1156, ISSN 0014-3057

Ejfler, J. et al., (2005). Highly efficient magnesium initiators for lactide polymerization. *Dalton Transactions,* No.11, (June 2005), pp. 2047-2050, ISSN 1477-9226

Engelhardt, E. M. et al., (2011). A collagen-poly(lactic acid-co-caprolactone) hybrid scaffold for bladder tissue regeneration. *Biomaterials,* Vol.32, No.16, (June 2011), pp. 3969-3976, ISSN 0142-9612

Fahmy, T. M. et al., (2005). Surface modification of biodegradable polyesters with fatty acid conjugates for improved drug targeting. *Biomaterials*, Vol.26, No.28, (October 2005), pp. 5727–5736, ISSN 0142-9612

Favia, P. & d'Agostino, R. (1998). Plasma treatments and plasma deposition of polymers for biomedical applications. *Surface and coatings Technology,* Vol.98, No.1, (January1998), pp. 1102-1106, ISSN 0257-8972

Fay, F. et al., (2010). Gene delivery using dimethyldidodecylammonium bromide-coated PLGA nanoparticles. *Biomaterials,* Vol.31, No.14, (May 2010), pp. 4214-4222, ISSN 0142-9612

Freitas, M. N. & Marchetti, J. M. (2005). Nimesulide PLA microspheres as a potential sustained release system for the treatment of inflammatory diseases. *International Journal of Pharmaceutics,* Vol.295, No.1-2, (May 2005), pp. 201-211, ISSN 0378-5173

Fu, B. et al., (2008). Preparation of lactic acid based polyurethanes modified by castor oil. *Advanced Materials Research,* Vols. 47-50, (June 2008), pp. 1458-1461, ISSN 1022-6680

Ganta, S. et al., (2008). A review of stimuli-responsive nanocarriers for drug and gene delivery. *Journal of Controlled Release,* Vol.126, No.3, (March 2008), pp. 187-204, ISSN 0168-3659

Gao, Q. W. et al., (2002). Direct synthesis with melt polycondensation and microstructure analysis of poly(L-lactic acid-co-glycolic acid). *Polymer Journal,* Vol.34, No.11, (November 2002), pp. 786-793, ISSN 0032-3896

Garlotta, D. (2001). A literature review of poly(lactic acid). *Journal of Polymers and the Environment,* Vol.9, No.2, (April 2001), pp. 63-84, ISSN 1566-2543

Ge, Z. et al., (2009). Proliferation and differentiation of human osteoblasts within 3D printed poly-lactic-co-glycolic acid scaffolds. *Journal of Biomaterials Applications*, Vol.23, (May 2009), pp.533-547, ISSN 1530-8022

Ge, Z. et al., (2009). Histological evaluation of osteogenesis of 3D-printed poly-lactic-co-glycolic acid (PLGA) scaffolds in a rabbit model. *Biomedical Materials,* Vol. 4, No. 2, (April 2009), 021001, ISSN 1748-6041

Ghosh, D. et al., (2007). Interactive wound cover. US Patent 0,258,958, filed February 12, 2007, issued November 8, 2007

Graupner, N.; Herrmann, A. S. & Müssig, J. (2009). Natural and man-made cellulose fibre-reinforced poly(lactic acid) (PLA) composites: An overview about mechanical characteristics and application areas. *Composites Part A: Applied Science and Manufacturing,* Vol.40, No.6-7, (July 2009), pp. 810-821, ISSN 1359-835X

Gref, R. et al., (1994). Biodegradable long-circulating polymeric nanospheres. *Science,* Vol.263, No.5153, (March1994), pp. 1600-1603, ISSN 0036-8075

Gu, M. et al., (2007). Surface biofunctionalization of PLA nanoparticles through amphiphilic polysaccharide coating and ligand coupling: Evaluation of biofunctionalization and drug releasing behavior. *Carbohydrate Polymers,* Vol.67, No.3, (February 2007), pp. 417-426, ISSN 0144-8617

Gupta, B.; Revagade, N. & Hilborn, J. Poly(lactic acid) fiber: An overview. *Progress in polymer science,* Vol.32, No.4, (April 2007), pp. 455-482, ISSN 0079-6700

Haag, R. (2004). Supramolecular drug-delivery systems based on polymeric core–shell architectures. *Angewandte Chemie International Edition,* Vol.43, No.3, (January 2004), pp. 278-282, ISSN 1433-7851

Hirotsu, T. et al., (2002). Plasma surface treatments of melt-extruded sheets of poly(L-lactic acid). *Polymer Engineering and Science,* Vol.42, No.2, (February 2002), pp. 299-306, ISSN 0032-3888

Hollister, S. J. (2005). Porous scaffold design for tissue engineering. *Nature Materials,* Vol.4, (July 2005), pp. 518-524, ISSN 1476-1122

Hong, Y. et al., (2005). Collagen-coated polylactide microspheres as chondrocyte microcarriers. *Biomaterials,* Vol.26, No.32, (November 2005), pp. 6305-6313, ISSN 0142-9612

Hu, K. et al., (2009). Lactoferrin-conjugated PEG-PLA nanoparticles with improved brain delivery: In vitro and in vivo evaluations. *Journal of Controlled Release,* Vol.134, No.1, (February 2009), pp. 55-61, ISSN 0168-3659

Huang, J. W. et al., (2009), Polylactide/nano and microscale silica composite films. I. Preparation and characterization. *Journal of Applied Polymer Science,* Vol.112, No.3, (May 2009), pp. 1688-1694, ISSN 0021-8995

Huang, L. et al., (2007). Synthesis and characterization of electroactive and biodegradable ABA block copolymer of polylactide and aniline pentamer. *Biomaterials,* Vol.28, No.10, (April 2007), pp. 1741-1751, ISSN 0142-9612

Huda, M. S. et al., (2008). Effect of fiber surface-treatments on the properties of laminated biocomposites from poly(lactic acid)(PLA) and kenaf fibers. *Composites Science and Technology,* Vol.68, No.2, (February 2008), pp. 424-432, ISSN 0266-3538

Ichikawa, F. et al., (1995). Process for preparing polyhydroxycarboxylic acid. US Patent 5,440,008, filed May 19, 1994, issued August 8, 1995

Ignatova, M. et al., (2009). Electrospun non-woven nanofibrous hybrid mats based on chitosan and PLA for wound-dressing applications. *Macromolecular Bioscience,* Vol.9, No.1, (January 2009), pp. 102-111, ISSN 1616-5187

Ilan, D. I. & Ladd, A. L. (2002). Bone graft substitutes. *Operative Techniques in Plastic and Reconstructive Surgery,* Vol.9, No.4, (November 2002), pp. 151-160, ISSN 1071-0949

Iwasa, J. et al., (2009). Clinical application of scaffolds for cartilage tissue engineering. *Knee Surgery, Sports Traumatology, Arthroscopy,* Vol.17, No.6, (June 2009), pp. 561-577, ISSN 0942-2056

Iwata, T. & Doi, Y. (1998). Morphology and enzymatic degradation of poly(L-lactic acid) single crystals. *Macromolecules,* Vol.31, No.8, (April 1998), pp. 2461-2467, ISSN 0024-9297

Jamshidian, M. et al., (2010). Poly-lactic acid: Production, applications, nanocomposites, and release studies. *Comprehensive Reviews in Food Science and Food Safety,* Vol.9, No.5, (September 2010), pp. 552-571, ISSN 1541-4337

Jiang, T. et al., (2006). In vitro evaluation of chitosan/poly(lactic acid-glycolic acid) sintered microsphere scaffolds for bone tissue engineering. *Biomaterials,* Vol.27, No.28, (November 2006), pp. 4894-4903, ISSN 0142-9612

Jiang, T. et al., (2010). Functionalization of chitosan/poly(lactic acid-glycolic acid) sintered microsphere scaffolds via surface heparinization for bone tissue engineering. *Journal of Biomedical Materials Research Part A,* Vol.93A, No.3, (June 2010), pp. 1193-1208, ISSN 1549-3296

Jiang, T. et al., (2010). Chitosan-poly(lactide-co-glycolide) microsphere-based scaffolds for bone tissue engineering: In vitro degradation and in vivo bone regeneration studies. *Acta Biomaterialia,* Vol.6, No.9, (September 2010), pp. 3457-3470, ISSN 1742-7061

Jie, P. et al.,(2005). Micelle-like nanoparticles of star-branched PEO-PLA copolymers as chemotherapeutic carrier. *Journal of Controlled Release,* Vol.110, No.1, (December 2005), pp. 20-33, ISSN 0168-3659

John, A. et al., (2007) Ni (II) and Cu (II) complexes of phenoxy-ketimine ligands: Synthesis, structures and their utility in bulk ring-opening polymerization (ROP) of L-lactide. *Polyhedron,* Vol.26, No.15, (April 2007), pp. 4033-4044, ISSN 0277-5387

Jung, Y. et al., (2005). A poly(lactic acid)/calcium metaphosphate composite for bone tissue engineering. *Biomaterials*, Vol.26, No.32, (November 2005), pp. 6314-6322, ISSN 0142-9612

Jérôme, C. & Lecomte, P. (2008). Recent advances in the synthesis of aliphatic polyesters by ring-opening polymerization. *Advanced drug delivery reviews,* Vol.60, No.9, (June 2008), pp. 1056-1076, ISSN 0169-409X

Kang, Y. et al., (2008). Preparation, characterization and in vitro cytotoxicity of indomethacin-loaded PLLA/PLGA microparticles using supercritical CO_2 technique. *European Journal of Pharmaceutics and Biopharmaceutics,* Vol.70, No.1, (September 2008), pp. 85-97, ISSN 0939-6411

Kasuga, T. et al., (2001). Preparation and mechanical properties of polylactic acid composites containing hydroxyapatite fibers. *Biomaterials,* Vol.22, No.1, (January 2001), pp. 19-23, ISSN 0142-9612

Kasuga, T. et al., (2003). Preparation of poly(lactic acid) composites containing calcium carbonate (vaterite). *Biomaterials,* Vol.24, No.19, (August 2003), pp. 3247-3253, ISSN 0142-9612

Khondker, O. A. et al., (2006). A novel processing technique for thermoplastic manufacturing of unidirectional composites reinforced with jute yarns. *Composites Part A: Applied Science and Manufacturing,* Vol.37, No.12, (December 2006), pp. 2274-2284, ISSN 1359-835X

Kim, I. H. & Jeong, Y. G. (2010). Polylactide/exfoliated graphite nanocomposites with enhanced thermal stability, mechanical modulus, and electrical conductivity.

Journal of Polymer Science Part B: Polymer Physics, Vol.48, No.8, (April 2010), pp. 850-858, ISSN 0887-6266

Kocbek, P. et al., (2007). Targeting cancer cells using PLGA nanoparticles surface modified with monoclonal antibody. *Journal of Controlled Release,* Vol.120, No.1-2, (July 2007), pp. 18-26, ISSN 0168-3659

Koo, G. H. & Jang, J. (2008). Surface modification of poly(lactic acid) by UV/Ozone irradiation. *Fibers and Polymers,* Vol.9, No.6, (December 2008), pp. 674-678, ISSN 1875-0052

Köhn, R. D. et al., (2003). Ring-opening polymerization of,-lactide with bis (trimethyl triazacyclohexane) praseodymium triflate. *Catalysis Communications,* Vol.4, No.1, (January 2003), pp. 33-37, ISSN 1566-7367

Lassalle, V. L. & Ferreira M. L. (2008). Lipase-catalyzed synthesis of polylactic acid: An overview of the experimental aspects. *Journal of Chemical Technology and Biotechnology,* Vol.83, No.11 , (October 2008), pp. 1493-1502, ISSN 0268-2575

Li, S. & Vert, M. (2003). Synthesis, characterization, and stereocomplex-induced gelation of block copolymers prepared by ring-opening polymerization of L (D)-lactide in the presence of poly(ethylene glycol). *Macromolecules,* Vol.36, No.21, (September 2003), pp. 8008-8014, ISSN 0024-9297

Lieb, E. et al., (2005). Mediating specific cell adhesion to low-adhesive diblock copolymers by instant modification with cyclic RGD peptides. *Biomaterials,* Vol.26, No.15, (May 2005), pp. 2333-2341, ISSN 0142-9612

Liggins, R. T. et al., (2004). Intra-articular treatment of arthritis with microsphere formulations of paclitaxel: Biocompatibility and efficacy determinations in rabbits. *Inflammation Research,* Vol.53, No.8, (August 2004), pp. 363-372, ISSN 1023-3830

Lin, Y. M. et al., (2010). Tissue engineering of lung: The effect of extracellular matrix on the differentiation of embryonic stem cells to pneumocytes. *Tissue Engineering Part A,* Vol.16, No.5, (April 2010), pp. 1515-1526, ISSN 1937-3341

Liu, C. et al., (2004). Effects of DBD plasma operating parameters on the polymer surface modification. *Surface and Coatings Technology,* Vol.185, No.2-3, (July 2004), pp. 311-320, ISSN 0257-8972

Liu, C. X. et al., (2011). Enhanced gene transfection efficiency in CD13-positive vascular endothelial cells with targeted poly(lactic acid)-poly(ethylene glycol) nanoparticles through caveolae-mediated endocytosis. *Journal of Controlled Release,* (in press), ISSN 0168-3659

Liu, P. et al., (2010). The research of RGD-conjugated PLA-PLL nanoparticles carriers on targeted delivery to tumor, *Proceedings of Nanoelectronics Conference (INEC), 2010 3rd International,* 978-1-4244-3543-2, Hong Kong, January 2010

Liu, X, & Ma, P. X. (2004). Polymeric scaffolds for bone tissue engineering. *Annals of Biomedical Engineering,* Vol.32, No.3, (March 2004), pp. 477-486, ISSN 0090-6964

Liu, X.; Won, Y. & Ma, P. X. (2005). Surface modification of interconnected porous scaffolds. *Journal of Biomedical Materials Research Part A,* Vol.74A, No.1, (July 2005), pp. 84-91, ISSN 1549-3296

Ljungberg, N. & Wesslén, B. (2002). The effects of plasticizers on the dynamic mechanical and thermal properties of poly(lactic acid). *Journal of Applied Polymer Science,* Vol.86, No.5, (October 2002), pp. 1227-1234, ISSN 0021-8995

Ljungberg, N.; Colombini, D. & Wesslén, B. (2005). Plasticization of poly(lactic acid) with oligomeric malonate esteramides: Dynamic mechanical and thermal film properties. *Journal of Applied Polymer Science,* Vol.96, No.4, (March 2005), pp. 992-1002, ISSN 0021-8995

Lowe, N. & Beer, K. (2005). A new material in subdermal sculpting, In: *Procedures in Cosmetic Dermatology Series; Soft Tissue Augmentation,* Carruthers, J. & Carruthers, A., pp. 143-146, Elsevier Saunders, 1-4160-4214-8, Oxford

Lu, J. W. et al., (2009). Surface engineering of poly(D, L-lactic acid) by entrapment of soluble eggshell membrane protein. *Journal of Biomedical Materials Research Part A,* Vol.91A, No.3, (December 2009), pp. 701-707, ISSN 1549-3296

Lunt, J. (1998). Large-scale production, properties and commercial applications of polylactic acid polymers. *Polymer Degradation and Stability,* Vol.59, No.1, (January 1998), pp. 145-152, ISSN 0141-3910

Lv, Q. et al., (2007). Preparation and characterization of PLA/fibroin composite and culture of HepG2 (human hepatocellular liver carcinoma cell line) cells. *Composites Science and Technology,* Vol.67, No.14, (June 2007), pp. 3023-3030, ISSN 0266-3538

Ma, H.; Davis, R. H. & Bowman, C. N. (1999). A novel sequential photoinduced living graft polymerization. *Macromolecules,* Vol.33, No.2, (January 2000), pp. 331-335, ISSN 0024-9297

Ma, P. X. (2004). Scaffolds for tissue fabrication. *Materials Today,* Vol.7, No.5, (May 2004), pp. 30-40, ISSN 1369-7021

Ma, W. et al., (2008). Evaluation of blood circulation of polysaccharide surface-decorated PLA nanoparticles. *Carbohydrate Polymers,* Vol.72, No.1, (April 2008), pp. 75-81, ISSN 0144-8617

Maharana, T.; Mohanty, B. & Negi, Y. S. (2009). Melt-solid polycondensation of lactic acid and its biodegradability. *Progress in Polymer Science,* (January 2009), Vol.34, No.1, pp. 99-124, ISSN 0079-6700

Martin, O. & Averous, L. (2001). Poly(lactic acid): Plasticization and properties of biodegradable multiphase systems. *Polymer,* Vol.42, No.14, (June 2001), pp. 6209-6219, ISSN 0032-3861

Mathew, A. P.; Oksman, K. & Sain, M. (2005). Mechanical properties of biodegradable composites from poly lactic acid (PLA) and microcrystalline cellulose (MCC). *Journal of applied polymer science,* Vol.97, No.5, (September 2005), pp. 2014-2025, ISSN 0021-8995

Mathieu, L. M. et al., (2006). Architecture and properties of anisotropic polymer composite scaffolds for bone tissue engineering. *Biomaterials,* Vol.27, No.6, (February 2006), pp. 905-916, ISSN 0142-9612

Matsumoto, K. & Taguchi, S. (2010). Enzymatic and whole-cell synthesis of lactate-containing polyesters: Toward the complete biological production of polylactate. *Applied microbiology and biotechnology,* Vol.85, No.4, (December 2009), pp. 921-932, ISSN 0175-7598

Matsumura, S.; Mabuchi, K. & Toshima, K. (1997). Lipase-catalyzed ring-opening polymerization of lactide. *Macromolecular Rapid Communications,* Vol.18, No.6, (April 2003), pp. 477-482, ISSN 1022-1336

Mauney, J. R. et al., (2007). Engineering adipose-like tissue in vitro and in vivo utilizing human bone marrow and adipose-derived mesenchymal stem cells with silk fibroin

3D scaffolds. *Biomaterials,* Vol.28, No.35, (December 2007), pp. 5280-5290, ISSN 0142-9612

Mavrogenis, A. F. et al., (2009). Early experience with biodegradable implants in pediatric patients. *Clinical Orthopaedics and Related Research,* Vol.467, No.6, (June 2009), pp. 1591-1598, ISSN 0009-921X

Mehta, R. et al., (2005). Synthesis of poly(lactic acid): A review. *Journal of Macromolecular Science, Part C: Polymer Reviews,* Vol.45, No.4, (October 2005), pp. 325-349, ISSN 1532-1797

Metters, A. T.; Anseth, K. S. & Bowman, C. N. (2000). Fundamental studies of a novel, biodegradable PEG-b-PLA hydrogel. *Polymer,* Vol.41, No.11, (May 2000), pp. 3993-4004, ISSN 0032-3861

Moon, S. I. et al., (2000). Melt polycondensation of L-lactic acid with Sn(II) catalysts activated by various proton acids: A direct manufacturing route to high molecular weight Poly(L-lactic acid). *Journal of Polymer Science Part A: Polymer Chemistry,* Vol.3, No.89, (May 2000), pp. 1673-1679, ISSN 0887-624X

Moon, S. I. et al., (2001). Synthesis and properties of high-molecular-weight poly(L-lactic acid) by melt/solid polycondensation under different reaction conditions. *High Performance Polymers,* Vol.13, No.2, (June 2001), pp. 189-196, ISSN 0954-0083

Moran, M. J.; Pazzano, D. & Bonassar, L. J. (2003). Characterization of polylactic acid-polyglycolic acid composites for cartilage tissue engineering. *Tissue Engineering,* Vol.9, No.1, (February 2003), pp. 63-70, ISSN 1076-3279

Na, K. et al., (2006). Biodegradable thermo-sensitive nanoparticles from poly(L-lactic acid)/poly(ethylene glycol) alternating multi-block copolymer for potential anti-cancer drug carrier. *European Journal of Pharmaceutical Sciences,* Vol.27, No.2–3, (December 2006), pp. 115–122, ISSN 0928-0987

Nafee, N. et al., (2007). Chitosan-coated PLGA nanoparticles for DNA/RNA delivery: Effect of the formulation parameters on complexation and transfection of antisense oligonucleotides. *Nanomedicine,* Vol.3, No.3, (September 2007), pp. 173-183, ISSN 1743-5889

Nalawade, S. P.; Picchioni F. & Janssen L. P. B. N. (2006). Supercritical carbon dioxide as a green solvent for processing polymer melts: Processing aspects and applications. *Progress in Polymer Science,* Vol.31, No.1, (January 2006), pp. 19-43, ISSN 0079-6700

Nomura, N. et al., (2002). Stereoselective ring-opening polymerization of racemic lactide using aluminum-achiral ligand complexes: Exploration of a chain-end control mechanism. *Journal of the American Chemical Society,* Vol.124, No.21, (May 2002), pp. 5938-5939, ISSN 0002-7863

Ohta, M.; Obuchi, S. &Yoshida, Y. (1995). Preparation process of polyhydroxycarboxylic acid. US Patent, 5,444,143, filed December 21, 1993, issued August 22, 1995

Oksman, K.; Skrifvars, M. & Selin J. F.. (2003). Natural fibres as reinforcement in polylactic acid (PLA) composites. *Composites Science and Technology,* Vol.63, No.9, (July 2003), pp. 1317-1324, ISSN 0266-3538

Pan, P. et al., (2007). Crystallization behavior and mechanical properties of bio-based green composites based on poly(L-lactide) and kenaf fiber. *Journal of Applied Polymer Science,* Vol.105, No.3, (Auguest 2007), pp. 1511-1520, ISSN 0021-8995

Patil, Y. & Panyam, J. (2009). Polymeric nanoparticles for siRNA delivery and gene silencing. *International Journal of Pharmaceutics.* Vol.367, No.1-2, (February 2009), pp. 195-203, ISSN 0378-5173

Peesan, M.; Supaphol, P. & Rujiravanit R. (2005). Preparation and characterization of hexanoyl chitosan/polylactide blend films. *Carbohydrate Polymers,* Vol.60, No.3, (March 2005), pp. 343-350, ISSN 0144-8617

Penczek, S. et al., (2000). What we have learned in general from cyclic esters polymerization. *Macromolecular Symposia,* Vol.153, No.1, (March 2000), pp. 1-15, ISSN 1022-1360

Phong, L. et al. (2010). Properties and hydrolysis of PLGA and PLLA cross-linked with electron beam radiation. *Polymer Degradation and Stability,* Vol.95, No.5, (May 2010), pp. 771-777, ISSN 0141-3910

Pluta, M. et al., (2002) Polylactide/montmorillonite nanocomposites and microcomposites prepared by melt blending: Structure and some physical properties. *Journal of Applied Polymer Science,* Vol.86, No.6, (November 2002), pp. 1497-1506, ISSN 0021-8995

Qiu, Y. & Park, K. (2001). Environment-sensitive hydrogels for drug delivery. *Advanced Drug Delivery Reviews*, Vol.53, No.3, (December 2001), pp. 321-339

Quirk, R. A. et al., (2000). Surface engineering of poly(lactic acid) by entrapment of modifying species. *Macromolecules,* Vol.33, No.2, (January2000), pp. 258-260, ISSN 0024-9297

Quynh, T. M. et al., (2007). Properties of crosslinked polylactides (PLLA & PDLA) by radiation and its biodegradability. *European Polymer Journal,* Vol.43, No.5, (May 2007) pp. 1779-1785, ISSN 0014-3057

Quynh T. M. et al., (2008). The radiation crosslinked films based on PLLA/PDLA stereocomplex after TAIC absorption in supercritical carbon dioxide. *Carbohydrate Polymers,* Vol.72, No.4, (June 2008), pp. 673-681, ISSN 0144-8617

Quynh, T. M. et al., (2009). Properties of radiation-induced crosslinking stereocomplexes derived from poly(L-lactide) and different poly(D-lactide). *Polymer Engineering and Science,* Vol.49, No.5, (May 2009), pp. 970-976, ISSN 0032-3888

Raghoebar, G. M. et al., Resorbable screws for fixation of autologous bone grafts. *Clinical Oral Implants Research,* Vol.17, No.3, (June 2006), pp. 288-293, ISSN 0905-7161

Rasal, R. M. & Hirt, D. E. (2009). Toughness decrease of PLA-PHBHHx blend films upon surface-confined photopolymerization. *Journal of Biomedical Materials Research Part A,* Vol.88A, No.4, (March 2009), pp. 1079-1086, ISSN 1549-3296

Rasal, R. M.; Janorkar, A. V. & Hirt, D. E. (2010). Poly(lactic acid) modifications. *Progress in Polymer Science,* Vol.35, No.3, (March 2010), pp. 338-356, ISSN 0079-6700

Reddy, N.; Nama, D. & Yang, Y. (2008). Polylactic acid/polypropylene polyblend fibers for better resistance to degradation. *Polymer Degradation and Stability,* Vol.93, No.1, (January 2008), pp. 233-241, ISSN 0141-3910

Rhim, J. W. et al., (2006). Effect of the processing methods on the performance of polylactide films: Thermocompression versus solvent casting. *Journal of Applied Polymer Science,* Vol.101, No.6, (September 2006), pp. 3736-3742, ISSN 0021-8995

Riley, T. et al., (2001). Physicochemical evaluation of nanoparticles assembled from poly(lactic acid)-poly(ethylene glycol) (PLA-PEG) block copolymers as drug delivery vehicles. *Langmuir,* Vol.17, No.11, (May 2001), pp. 3168-3174, ISSN 0743-7463

Robinson, T. M.; Kieswetter, K. & McNulty, A. (2009). System and method for healing a wound at a tissue site. US Patent 0,216,170, filed February 27, 2009, issued August 27, 2009

Roney, C. et al., (2005). Targeted nanoparticles for drug delivery through the blood-brain barrier for Alzheimer's disease. *Journal of Controlled Release,* Vol.108, No.1-2, (November 2005), pp. 193-214, ISSN 0168-3659

Rujitanaroj, P.; Pimpha, N. & Supaphol, P. (2008). Wound-dressing materials with antibacterial activity from electrospun gelatin fiber mats containing silver nanoparticles. *Polymer,* Vol.49, No.21, (October 2008), pp. 4723-4732, ISSN 0032-3861

Sarazin, Y.; Schormann, M. & Bochmann, M. (2004). Novel zinc and magnesium alkyl and amido cations for ring-opening polymerization reactions. *Organometallics,* Vol.23, No.13, (May 2004), pp. 3296-3302, ISSN 0276-7333

Sawahata, K. et al., (1990). Electrically controlled drug delivery system using polyelectrolyte gels. *Journal of Controlled Release,* Vol.14, No.3, (December 1990), pp. 253-262, ISSN 0168-3659

Sethuraman, V. A. & Bae, Y. H. (2007). TAT peptide-based micelle system for potential active targeting of anti-cancer agents to acidic solid tumors. *Journal of Controlled Release,* Vol.118, No.2, (April 2007), pp. 216-224, ISSN 0168-3659

Sheth, M. et al., (1997) Biodegradable polymer blends of poly(lactic acid) and poly(ethylene glycol). *Journal of Applied Polymer Science,* Vol.66, No.8, (November 1997), pp. 1495-1505, ISSN 0021-8995

Shin, I. K. et al., (2010). Novel three-dimensional scaffolds of poly(L-lactic acid) microfibers using electrospinning and mechanical expansion: Fabrication and bone regeneration. *Journal of Biomedical Materials Research B: Applied Biomaterials*, Vol.95B, No.1, (October 2010), pp.150-160, ISSN 1552-4981

Singh, R. P. et al., (2003). Biodegradation of poly(ε-caprolactone)/starch blends and composites in composting and culture environments: The effect of compatibilization on the inherent biodegradability of the host polymer. *Carbohydrate Research,* Vol.338, No.17, (August 2003), pp. 1759-1769, ISSN 0008-6215

Stupack, D. G. et al. (October 2001). Apoptosis of adherent cells by recruitment of caspase-8 to unligated integrins. *The Journal of Cell Biology,* Vol.155, No.3, (2001), pp. 459-470, ISSN 0021-9525

Taguchi, S. et al., (2008). A microbial factory for lactate-based polyesters using a lactate-polymerizing enzyme. *Proceedings of the National Academy of Sciences,* Vol.105, No.45, (November 2008), pp. 17323-17327, ISSN 0027-8424

Taguchi, S. (2010). Current advances in microbial cell factories for lactate-based polyesters driven by lactate-polymerizing enzymes: Towards the further creation of new LA-based polyesters. *Polymer Degradation and Stability,* Vol.95, No.8, (January 2010), pp. 1421-1428, ISSN 0141-3910

Tajima, K. et al., (2009). Chemo-enzymatic synthesis of poly(lactate-co-(3- hydroxybutyrate)) by a lactate-polymerizing enzyme. *Macromolecules,* Vol.42, No.6, (February 2009), pp. 1985-1989, ISSN 0024-9297

Tokoro, R. et al., (2008). How to improve mechanical properties of polylactic acid with bamboo fibers. *Journal of Materials Science,* Vol.43, No.2, (January 2008), pp. 775-787, ISSN 0022-2461

Tomé, C. L. et al.,(2011). Transparent bionanocomposites with improved properties prepared from

acetylated bacterial cellulose and poly(lactic acid) through a simple approach. *Green Chemistry,* Vol.13, No.2, (February 2011), pp. 419-427, ISSN 1463-9262

Torchilin, V. P. (2006). Multifunctional nanocarriers. *Advanced Drug Delivery Reviews,* Vol.58, No.14, (December 2006), pp. 1532–1555, ISSN 0169-409X

Torchilin, V. P. (2008). Cell penetrating peptide-modified pharmaceutical nanocarriers for intracellular drug and gene delivery. *Biopolymers,* Vol.90, No.5, (September 2008), pp. 604-610, ISSN 0006-3525

Tyagi, P. et al., (2004). Sustained intravesical drug delivery using thermosensitive hydrogel. *Pharmaceutical Research,* Vol.21, No.5, (May 2004), pp. 832-837, ISSN 0724-8741

Vaupel, P.; Kallinowski, F. & Okunieff, P. (1989). Blood flow, oxygen and nutrient supply, and metabolic microenvironment of human tumors: A review. *Cancer Research,* Vol.49, No.23, (December 1989), pp. 6449–6465, ISSN 0008-5472

Venkatraman, S. S. et al., (2005). Micelle-like nanoparticles of PLA-PEG-PLA triblock copolymer as chemotherapeutic carrier. *International Journal of Pharmaceutics,* Vol.298, No.1, (July 2005), pp. 219-232, ISSN 0378-5173

Vila, A. et al., (2005). PLA-PEG particles as nasal protein carriers: The influence of the particle size. *International Journal of Pharmaceutics,* Vol.292, No.1-2, (March 2005), pp. 43-52, ISSN 0378-5173

Vink, E. T. H. et al., (2003). Applications of life cycle assessment to NatureWorks™ polylactide (PLA) production. *Polymer Degradation and Stability,* Vol.80, No.3, (2003), pp. 403-419, ISSN 0141-3910

Vleggaar, D. & Bauer, U. (2004). Facial enhancement and the European experience with Sculptra. *Journal of Drugs in Dermatology,* Vol.3, No.5, (September-October 2004), pp. 542-547, ISSN 1545-9616

Wan, Y. Z. et al., (2001). Influence of surface treatment of carbon fibers on interfacial adhesion strength and mechanical properties of PLA-based composites. *Journal of Applied Polymer Science,* Vol.80, No., (April 2001), pp. 367-376, ISSN 0021-8995

Wang, M. (2007). Surface modification of biomaterials and tissue engineering scaffolds for enhanced osteoconductivity, *Proceedings of IFMBE,* 978-3-540-68016-1, Kuala Lumpur, Malaysia, December 2006

Wang, R. Y. et al., (2009). Morphology, mechanical properties, and durability of poly(lactic acid) plasticized with di(isononyl) cyclohexane-1, 2-dicarboxylate. *Polymer Engineering and Science,* Vol.49, No.12, (December 2009), pp. 2414-2420, ISSN 0032-3888

Wang, S.; Cui, W. & Bei, J. (2005). Bulk and surface modifications of polylactide. *Analytical and Bioanalytical Chemistry,* Vol.381, No.3, (February 2005), pp. 547-556, ISSN 1618-2642

Woo, S. I. et al., (1995) Polymerization of aqueous lactic acid to prepare high molecular weight poly(lactic acid) by chain-extending with hexamethylene diisocyanate. *Polymer Bulletin,* Vol.35, No.4, (May1995), pp. 415-421, ISSN 0170-0839

Wu, C. S. & Liao, H.T. (2007). Study on the preparation and characterization of biodegradable polylactide/multi-walled carbon nanotubes nanocomposites. *Polymer,* Vol.48, No.15, (July 2007), pp. 4449-4458, ISSN 0032-3861

Wu, J. C. et al., (2005). Ring-opening polymerization of lactide initiated by magnesium and zinc alkoxides. *Polymer,* Vol.46, No.23, (November 2005), pp. 9784-9792, ISSN 0032-3861

Wu, X. L. et al., (2010). Tumor-targeting peptide conjugated pH-responsive micelles as a potential drug carrier for cancer therapy. *Bioconjugate Chemistry,* Vol.21, No.2, (February 2010), pp. 208-213, ISSN 1043-1802

Xiong, X. Y. et al., (2007). Vesicles from Pluronic/poly(lactic acid) block copolymers as new carriers for oral insulin delivery. *Journal of Controlled Release,* Vol.120, No.1-2, (July 2007), pp. 11-17, ISSN 0168-3659

Xu, C. et al., (2009). Preparation and characterization of polylactide/thermoplastic konjac glucomannan blends. *Polymer,* Vol.50, No.15, (July 2009), pp. 3698-3705, ISSN 0032-3861

Yang, G. et al., (2011). Composite from modified biomass fiber and polylactide (PLA) and the method of preparation. CN Patent 201010510498, filed October 15, 2010, and issued January 26, 2011

Yang, S. L. et al., (2008) Thermal and mechanical properties of chemical crosslinked polylactide (PLA). *Polymer Testing,* Vol.27, No.8, (December 2008), pp. 957-963, ISSN 0142-9418

Yoda, S.; Bratton, D. & Howdle, S. M. (2004). Direct synthesis of poly(L-lactic acid) in supercritical carbon dioxide with dicyclohexyldimethylcarbodiimide and 4-dimethylaminopyridine. *Polymer,* Vol.45, No.23, (October 2004), p. 7839-7843, ISSN 0032-3861

Yu, D. G.; Lin, W. C. & Yang, M. C. (2007). Surface modification of poly(L-lactic acid) membrane via layer-by-layer assembly of silver nanoparticle-embedded polyelectrolyte multilayer. *Bioconjugate Chemistry,* Vol.18, No.5, (September-October2007), pp. 1521-1529, ISSN 1043-1802

Zhang, L. et al., (2004). Camptothecin derivative-loaded poly(caprolactone-co-lactide)-b-PEG-b-poly(caprolactone-co-lactide) nanoparticles and their biodistribution in mice. *Journal of controlled release,* Vol.96, No.1, (April 2004), pp. 135-148, ISSN 0168-3659

Zhang, Z. & Feng, S. S. (2006). Nanoparticles of poly(lactide)/vitamin E TPGS copolymer for cancer chemotherapy: Synthesis, formulation, characterization and in vitro drug release. *Biomaterials,* Vol.27, No.2, (January 2006), pp. 262-270, ISSN 0142-9612

Zhong, Z. et al., (2003). Single-site calcium initiators for the controlled ring-opening polymerization of lactides and lactones. *Polymer Bulletin,* Vol.51, No.3, (November 2003), pp. 175-182, ISSN 0170-0839

Zhu, A. et al., (2002). Covalent immobilization of chitosan/heparin complex with a photosensitive hetero-bifunctional crosslinking reagent on PLA surface. *Biomaterials,* Vol.23, No.23, (December 2002), pp. 4657-4665, ISSN 0142-9612

Zhu, H.; Ji, J. & Shen, J. (2002). Surface engineering of poly(DL-lactic acid) by entrapment of biomacromolecules. *Macromolecular Rapid Communications,* Vol.23, No.14, (October 2002), pp. 819-823, ISSN 1022-1336

Zhu, Y. et al., Immobilization of biomacromolecules onto aminolyzed poly(L-lactic acid) toward acceleration of endothelium regeneration. *Tissue Engineering,* Vol.10, No.1-2, (January-February 2004), pp. 53-61, ISSN 1076-3279

12

Multifunctional Magnetic Hybrid Nanoparticles as a Nanomedical Platform for Cancer-Targeted Imaging and Therapy

Husheng Yan[1], Miao Guo[1] and Keliang Liu[2]
[1]*Key Laboratory of Functional Polymer Materials, Ministry of Education,*
Institute of Polymer Chemistry, Nankai University, Tianjin,
[2]*Beijing Institute of Pharmacology and Toxicology, Beijing,*
P. R. China

1. Introduction

Nanotechnology offers tremendous potential for use in biomedical applications, including imaging, disease diagnosis, and drug delivery. The development of nanosystems has improved the molecular understanding of many diseases and permitted the controlled nanoscale manipulation of materials (Couvreur & Vauthier, 2006). Nanomedical platforms offer many advantages as delivery, sensing, and image-enhancing agents. In recent years, many studies have focused on multifunctional nanomedical platforms that incorporate therapeutic and diagnostic agents with molecular targeting capabilities. Gregoriadis et al. first proposed liposomes as drug carriers in cancer chemotherapy in 1974 (Gregoria et al., 1974). Today, drug delivery systems made of lipids or polymers frequently are exploited for the controlled delivery of therapeutic drugs in the body (Jain, 2005; Vasir et al., 2005).

Nanosized particles for biomedical platforms can be made from a variety of materials, including lipids (liposomes, nanoemulsions, and solid-lipid nanoparticles), self-assembling amphiphilic molecules, nondegradable and degradable polymers, dendrimers, metals, and inorganic semiconductor nanocrystals. The selection of the platform material is determined by the desired diagnostic or therapeutic goal, payload type, material safety profile, and administration route. Among the various types of functional nanostructures, nanomedical platforms based on magnetic nanoparticles (MNPs) are of particular interest in biomedical applications. Most frequently, MNPs are constructed of superparamagnetic iron oxides (SPIOs) (e.g., Fe_3O_4 or γ-Fe_2O_3), although metals such as cobalt and nickel are also employed. The characteristics of MNPs, including their composition, size, morphology, and surface chemistry, are tailored by various processes for their wide application in the detection, diagnosis, and treatment of illnesses. The most popular MNPs for biomedical applications are comprised of a magnetic inorganic nanoparticle core and a biocompatible surface coating that provides stabilization under physiological conditions. The additional application of a suitable surface chemistry allows the integration of functional ligands, such that MNPs can perform multiple functions. The modification and functionalization of MNPs improve their magnetic properties and affect their behavior in vivo (Tartaj et al., 2003; A.K. Gupta & M. Gupta, 2005).

Multifunctional MNPs (MFMNPs) are a major class of nanoscale materials with the potential to revolutionize clinical diagnostic and therapeutic techniques. Due to their unique magnetic properties and ability to function at the cellular and molecular levels of biological interactions, MFMNPs have been investigated as an attractive nanomedical platform. MFMNPs in the form of SPIOs have been actively investigated as contrast enhancement agents for magnetic resonance imaging (MRI) and hyperthermia in response to an external alternating magnetic field, due to their ability to enhance the proton relaxation of specific tissues. MFMNPs have been evaluated extensively as a nanomedical platform for the targeted delivery of pharmaceuticals through magnetic drug targeting (Neuberger et al., 2005) and through the attachment of high-affinity ligands (Zhang et al., 2002; Torchilin, 2006).

2. Surface coatings and functionalization of MNPs

2.1 Core–shell structure

Iron oxides with a core/shell structure are widely used as sources of MFMNP platforms. Iron oxides have several crystalline polymorphs, but only γ-Fe_2O_3 (maghemite) or Fe_3O_4 (magnetite) can be used for biomedical applications. These particles, which range in diameter from about 5–20 nm, have unique advantages, including (1) superparamagnetic behavior, with no magnetism after removal of the magnetic field; (2) high saturation magnetization values and high magnetic susceptibility, for effective magnetic enrichment; (3) biocompatibility and rapid removal through extravasation and renal clearance; and (4) easily tailored surface chemistry and functionalization.

Iron oxide nanoparticles have a significant tendency to agglomerate as a result of their high surface energy. Massart (1981) first prepared stable aqueous dispersions of Fe_3O_4 nanoparticles (ferrofluids) that were stabilized by electrical double layers. However, the colloidal electrostatic stabilization arising from surface charge repulsion on the nanoparticles typically is inadequate to prevent aggregation in biological solutions, due to the presence of salts or other electrolytes that can neutralize the charges. Furthermore, the iron oxide surfaces may be subjected to plasma protein adsorption or opsonization, leading to their rapid clearance by the reticuloendothelial system (RES) (Berry & Curtis, 2003).

To solve the above problems, proper surface coatings have been exploited as an integral component of the MFMNP platform for biomedical applications. The iron oxide core can be coated by organic materials [e.g., polymers such as dextrant (Thorek et al., 2006) and polyethylene glycol (PEG) (Gref et al., 1994)], inorganic metallic materials [e.g., gold (Ji et al., 2007)], or oxides [e.g., silica or alumina (Bumb et al., 2008)]. Polymer coatings will be introduced in detail in the next section. Silica shells are attractive as protective coatings on the iron oxide core, due to their stability under aqueous conditions and ease of synthesis. Recently, Ma et al. (2006) described one such core–shell MFMNP, composed of an iron oxide core (approximately 10 nm diameter) surrounded by a SiO_2 shell (10–15 nm thick). They doped an organic dye, tris(2,2′-bipyridine) ruthenium, inside a second silica shell to provide luminescence and prevent quenching by interaction with the magnetic core. As a core–shell structure exhibiting superparamagnetic and luminescent properties, this MFMNP platform can be used as a multifunctional imaging agent for biomedical applications.

Gold offers several advantages as a coating material for iron oxide cores, due to its low chemical reactivity and unique ability to form self-assembled monolayers on the core surface using alkanethiols (Prime & Whitesides, 1991). A variety of methods (reversed microemulsion, combined wet chemical, and laser irradiation) can be used to synthesize gold-coated iron oxides (A. H. Lu et al., 2007).

The core/shell structure of MFMNPs offers several advantages, including good dispersibility and high stability against oxidation. In addition, an appreciable amount of therapeutic agent can be loaded on the MFMNP shell. Functionalization chemistries generally are better established when a coating material is used.

2.2 Polymer coatings

Polymers comprise some of the most important materials used as shells. Polymer coatings not only provide a steric barrier to prevent nanoparticle agglomeration, but also allow MNPs to evade uptake by the RES and thereby to maintain a long plasma half-life. Polymer coatings provide a means to tailor the surface properties of MNPs, such as the surface charge and chemical functionality. An ideal polymer coating will have a high affinity for the iron oxide core, as well as nonimmunogenic and nonantigenic properties. It also will prevent opsonization by plasma proteins. Polymer materials comprised of lipids, proteins, dendrimers, gelatin, dextran, chitosan, pullulan, PEG, poly(ethylene-co-vinyl acetate), poly(vinylpyrrolidone), poly(vinyl alcohol) (PVA), or poly(glycerol monoacrylate) (PGA) are often chosen as the surface coatings for MNPs.

PEG is the most widely used polymer for nanoparticle coating in biomedical applications. PEG provides a very attractive combination of properties: excellent solubility in aqueous solutions; high flexibility of its polymer chain; very low toxicity, immunogenicity, and antigenicity; lack of accumulation in the RES cells; and minimal influence on the specific biological properties of modified pharmaceuticals (Yamaoka et al., 1994). As a so-called "stealth" surface, PEG prevents the nanomedical platform from being recognized by RES, and thereby extends its blood circulation time in vivo. On the biological level, coating nanoparticles with PEG sterically hinders the interaction of blood components with the nanoparticle surface and reduces the binding of plasma proteins. Mechanisms of preventing opsonization by PEG include the shielding of the surface charges, increased surface hydrophilicity (Gabizon & Papahadjopoulos, 1992), and enhanced repulsive interaction between polymer-coated nanoparticles and blood components (Needham et al., 1992). Various methods have been utilized to attach PEG to the MNP surface, including silane grafting to the oxide surface (Butterworth et al., 2001), alkaline coprecipitation of ferric and ferrous ions in the presence of PEG-containing block copolymers (Wan et al., 2005), direct attachment of PEG-containing block copolymers (Guo et al., 2010), polymerization at the MNP surface (Flesch et al., 2005), and modification through sol-gel approaches (Y. Lu et al., 2002).

Polysaccharide dextran is another polymer coating that has been used widely and successfully in vivo. Dextran-coated iron oxide nanoparticles have become an important part of clinical cancer imaging, and have been shown to increase the accuracy of cancer nodal staging (Harisinghani & Weissleder, 2004; Ferrari, 2005). Because the dextran coating is not strongly associated with the iron oxide core, the polymer is susceptible to detachment. Accordingly, cross-linked iron oxide nanoparticles have been developed by cross-linking the dextran shell with epichlorohydrin (Josephson et al., 1999). The resulting particle offers superb stability under harsh conditions, without causing any change in size or blood half-life or loss of the dextran coat. Chemical functionality can be established by treating cross-linked iron oxide nanoparticles with ammonia to provide primary amino groups for the attachment of biomolecules such as proteins or peptides (Wunderbaldinger et al., 2002; Schellenberger et al., 2002). These formulations of dextran-coated iron oxide nanoparticles have been evaluated extensively for a variety of MRI applications (Josephson et al., 1999).

In addition to the traditional polymer coatings, a new kind of biocompatible polymer material has been reported by Wan et al. (2005): namely, homopolymers of glycerol monoacrylate or glycerol monomethacrylate, or their block copolymers. Highly stable aqueous magnetic fluids were prepared by coating Fe_3O_4 nanoparticles with poly(glycerol monoacrylate) (PGA), poly(glycerol monomethacrylate) (PGMA), or diblock copolymers with PGA or PGMA segments. As shown in Fig. 1, the proposed mechanism of stabilization was the multidentate interactions of 1,2-diols on the polymer chain with iron atoms at the surface of the iron oxide nanoparticles (Wan et al., 2005). This process was a good choice for the preparation of stable magnetic fluids with tailored surfaces; PGA or PGMA binds very tightly to the iron oxide surface, is highly hydrophilic, and does not introduce charges on the surface. Moreover, various block copolymers containing PGA or PGMA can be used to modify the iron oxides and to introduce tailored functional groups for further functionalization.

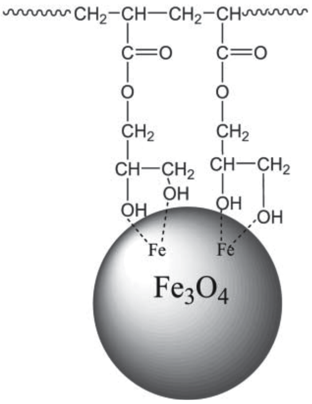


Fig. 1. Proposed structure in the interaction between the iron oxide surface and PGA (Wan et al., 2005).

2.3 Functional ligands

As discussed in the above sections, the core–shell structure of MFMNPs provides a means to tailor the nanoparticle surface properties, such as surface charge and chemical functionality. Various functional ligands, including targeting agents, permeation enhancers, optical dyes, and therapeutic agents, can be conjugated on the surface or incorporated within the nanostructure. The modification of the nanoparticle surface with targeting ligands was described recently as a promising biotargeting strategy. To generate target-specific nanoparticles, various biological molecules, such as antibodies, proteins, small molecular targeting agents, etc., can be bound to the coating surfaces of the MFMNPs by chemical coupling. Tumor cells are rapidly proliferating and overexpress certain receptors that lead to the enhanced uptake of nutrients, including folic acid, vitamins, sugars, and proteins. MFMNPs conjugated with these molecules can be targeted to tumor cells that overexpress the corresponding receptors. Table 1 summarizes a number of different ligands and their corresponding functions that have been investigated for the in vivo targeting of MFMNPs.

Targeting ligand	Functional activity	References
Folic acid	Preferentially targets cancer cells that overexpress folate receptors and facilitates internalization	Zhang et al., 2002
CREKA peptide or F3 peptide	Targets an antigen associated with colorectal carcinoma cells	Reddy et al., 2006; Simberg et al., 2007
Pullulan	Increases receptor-mediated hepatic uptake	Kaneo et al., 2001
Elastin	Cross-linked protein; Provides elasticity for many tissues	Debelle & Tamburro, 1999
RGD peptide	Enhances cell spreading, differentiation, and DNA synthesis	Bhadriraju & Hansen, 2000
Tat-peptide	Membrane-permeating peptide; Enhances intracellular delivery of nanoparticles	Josephson et al., 1999; Lewin et al., 2000
Transferrin	Targets primary proliferating cells by transferrin receptors	Weissleder et al., 2000; Moore et al., 2001
Insulin	Hormone; Regulates blood glucose levels	Gupta et al., 2003
Monoclonal antibody A7	Targets an antigen associated with colorectal carcinoma cells	Toma et al., 2005

Table 1. Selected functional ligands used for MFMNPs in biomedical applications

Organic dyes or fluorophores have been loaded on MNPs as optical imaging agents to allow detection by multiple imaging modalities. In addition to their use as contrast enhancement agents, FITC- (Zhang et al., 2002), rhodamine- (Bertorelle et al., 2006), or other fluorophore-labeled MNPs can be used for the in vitro fluorescent imaging of cells. Since both MRI and optical signals come from the same nanoparticles, the MR image can serve as a roadmap to the fluorescently labeled tumor cells. The conjugation of near-infrared fluorescent (NIRF) dyes to MNPs has received recent attention due to the deep penetration of NIRF light in the tissues (Weissleder & Ntziachristos, 2003). The integration of NIRF detectability allows for these nanoparticles to be used for presurgical planning by MRI and intraoperative resection of malignant tissues by optical imaging.

3. Biomedical applications of MFMNPs

3.1 Targeted drug delivery

One promising biomedical application of MFMNPs is as carriers for site-specific drug delivery. Many therapeutic agents, while pharmacologically effective, also exhibit side-effects because of their toxicities. For example, cytotoxic compounds used in cancer therapy kill not only target cells but also normal cells in the body, resulting in undesired side-effects. Meanwhile, many barriers to the delivery of therapeutic agents are presented, including renal clearance of small molecular therapeutic agents and overexpressed membrane-associated multi-drug resistance developed by tumor cells. Therefore, many therapeutic agents are limited in their clinical application. As widely used nanocarriers, MFMNPs have been considered as alternatives for the target-specific delivery of drugs to different sites in the body. These engineered nanoparticulate carriers offer some advantages, including passive targeting due to the enhanced permeability and retention (EPR) effect and functionalized surface features for target-specific localization. It also may be possible to develop nanocarriers that respond to physiological stimuli, or to combine drugs with energy (heat, light, and sound) delivery for synergistic therapeutic effects.

3.1.1 Passive targeting

Passive targeting relies on the properties of the delivery system and the disease pathology to accumulate the drug preferentially at the site of interest and avoid nonspecific distribution. Long-circulating nanoparticles of 20-200 nm in diameter containing surface PEG or poly(ethylene oxide) (PEO) blocks can accumulate at sites of disease such as tumors, infection, or inflammation through passive targeting via the EPR effect. Maeda and colleagues (2001) first described the EPR effect in their study of vascular abnormalities of solid tumors. Blood vessels in most solid tumors possess unique characteristics that are not usually observed in normal blood vessels, including: active angiogenesis and high vascular density; extensive production of vascular mediators that facilitate extravasation; defective vascular architecture (lack of smooth muscle layer cells, lack of receptors for angiotensin, large gap in endothelial cell-cell junctions, and anomalous conformations); and impaired lymphatic clearance of macromolecules and lipids from interstitial tissue.

Due to the EPR effect, nanopharmaceuticals (macromolecular drugs and drug-loaded nanoparticles) accumulate in tumor tissues with remarkable selectivity as schematically illustrated in Fig. 2. For example, the administration of polymer-drug conjugates results in 10-100 fold higher drug concentrations in the tumor compared to the administration of free drug (van Vlerken et al., 2007). This selective drug targeting to solid tumors results in substantial therapeutic benefits due to the higher drug accumulation in the tumor tissue, as well as fewer side effects. The EPR effect also has been observed in inflammatory and infectious tissues. Thus, the application of nanocarriers is expected to have therapeutic benefits for treating these diseases as well (Allen & Cullis, 2004).

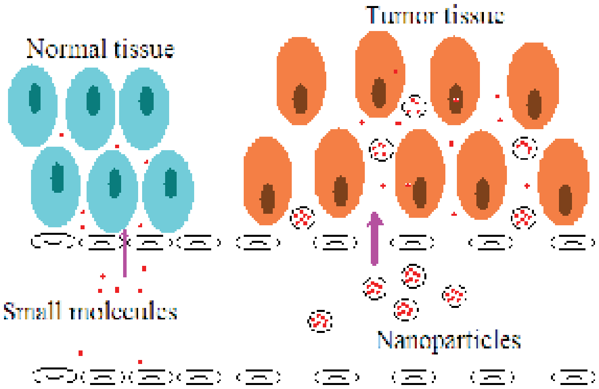


Fig. 2. Schematic illustration for passive targeting using the EPR effect.

Another approach for passive targeting involves the tendency of nanoparticles to localize in the RES. This phenomenon provides an opportunity for nanoparticles to accumulate at high concentrations in the liver or spleen, where many macrophages are present. Overall, the passive targeting strategy provides a means of delivering MFMNPs (as contrast agents or drug carriers) or other nanoparticles to the targeted organs or tissues.

3.1.2 Active targeting

Another promising approach towards increasing the local accumulation of nanoparticles in diseased tissue is known as active (or specific) targeting. Active targeting involves the conjugation of targeting molecules that possess high affinity toward unique molecular signatures found on malignant cells. Targeting ligands, such as proteins, peptides, aptamers and small molecules, have been investigated to increase the site-specific accumulation of MFMNPs. For example, there are certain receptors that are overexpressed on the surface of solid tumor cells, such as antigens, integrin receptors, and folate receptors (Table 1). By bonding with these targeting molecules, MFMNPs can be targeted to the corresponding tumor cells and internalized by receptor- or antigen-mediated endocytosis.

Monoclonal antibodies (mAbs) were the first targeting agents to exploit molecular recognition to deliver MNPs; mAbs continue to be used widely, due to their high specificity. For instance, Herceptin®, an FDA-approved mAb to the HER2/neu (erbB2) receptor, has been used to modify DMSA-coated magnetite nanoparticles. When these MFMNPs were used as contrast enhancement agents, the MR imaging of mice bearing xenograft tumors showed a T2 decrease of ~20% due to the specific accumulation of the nanoprobe in the tumor (Huh et al., 2005). Nanoparticles modified with an HER2-specific antibody (Trastuzumab® or Herceptin®) also are able to localize and deliver the therapeutic payload specifically in HER2-expressing tumor cells (Kirpotin et al., 2006). Certain tumor cells express specific integrin receptors, such as $\alpha_v\beta_5$ or $\alpha_v\beta_3$ that can bind to the arginine-glycine-aspartic acid (RGD) peptide sequence. The RGD peptide has been utilized for the delivery of MNPs to a variety of neoplastic tissues, including breast tumors, malignant melanomas, and squamous cell carcinomas (Montet et al., 2006).

Among the small targeting molecules, folate has been used to modify nanoparticles for targeted delivery to tumor cells that overexpress folate receptors. Recently, our group reported multilayer MFMNPs with a folate-modified surface and doxorubicin (an anticancer chemotherapeutic agent) loaded in the inner shell (Fig. 3) (Guo et al., 2011). The folate-conjugated MFMNPs displayed a much greater cellular uptake than nonfolate-conjugated MFMNPs by a folate receptor–mediated endocytosis process (Fig. 4). Folate conjugation significantly increased nanoparticle cytotoxicity against human cervical carcinoma HeLa cells (Guo et al., 2011).

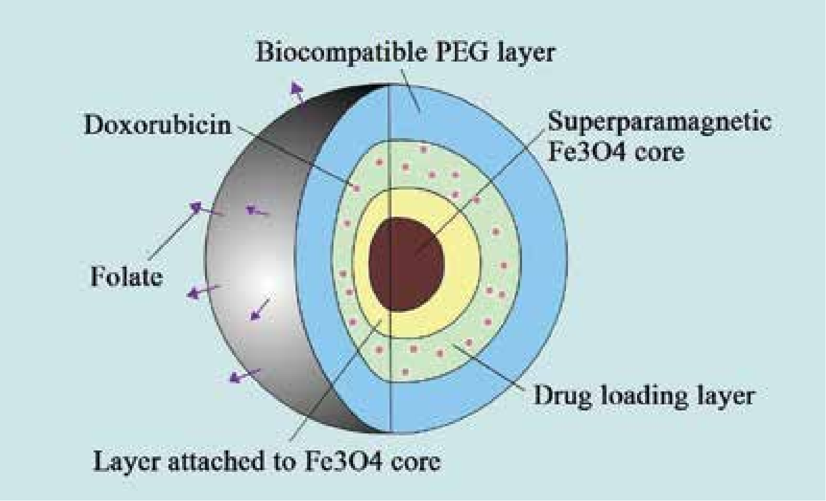


Fig. 3. Schematic illustration of multilayer MFMNPs with folate as the targeting ligand and loaded doxorubicin as the anticancer chemotherapeutic agent in the inner shell.

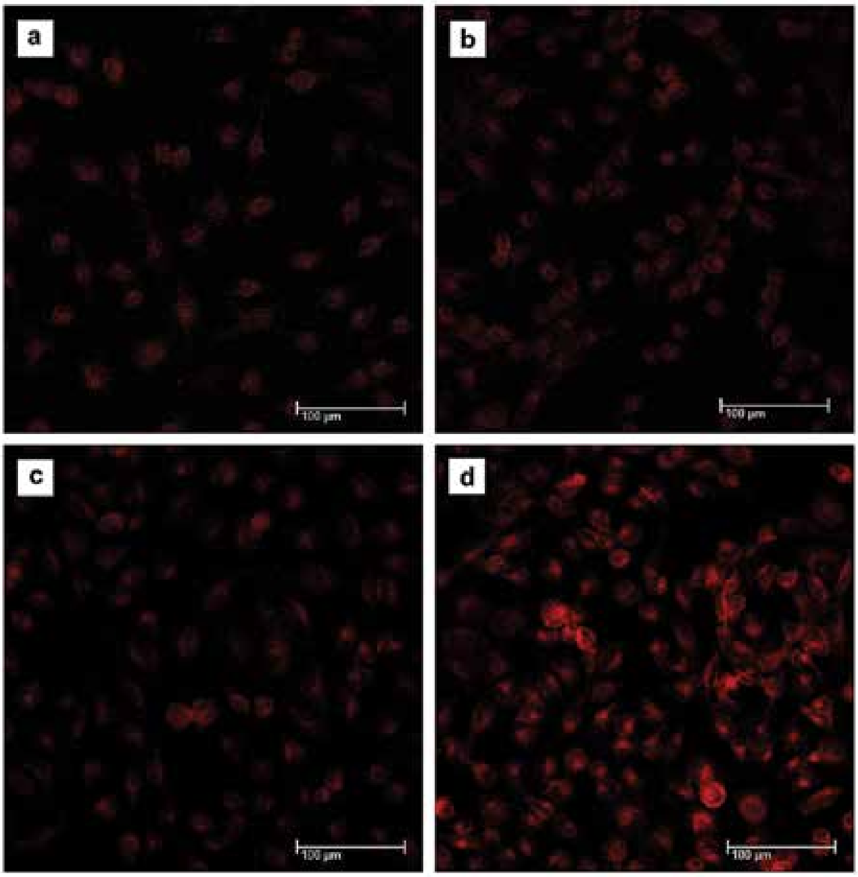


Fig. 4. Confocal microscopic images of HeLa cells incubated with (a, b) doxorubicin-loaded MFMNPs without folate conjugation or (c, d) folate–conjugated and doxorubicin-loaded MFMNPs in (a, c) folate-containing or (b, d) folate-free media (Guo et al., 2011).

3.1.3 Drug loading and controlled release

An essential step in the use of MFMNPs for drug delivery is the controlled release of the therapeutic payload in the desired tumor cells or tissues. The drug-loading capacity and release rate are correlated with the binding affinity of the drug. Strong carrier–drug interactions may enhance the loading capacity and decrease the release rate of the drug from the carrier; therefore, the choice of proper carrier–drug interactions is critical in the design and preparation of nanocarriers for drug delivery.

Successful MNP delivery devices with a prolonged circulation time can carry a chemotherapeutic payload and can be engineered to release its drugs after cell internalization. To successfully integrate a drug into a NP system, several design strategies can be explored, including physical complexation with hydrophobic drugs, or covalent bonding with cleavable linkages for intracellular release. Currently, several chemical drug formulations have been combined with MNPs, including paclitaxel, doxorubicin, and methotrexate, all specifically developed for cancer therapy. For example, methotrexate, an anticancer drug, has an affinity to the target cells, and after grafting the drug to the surface, MNPs can be internalized more rapidly. Kohler et al. (2006) first demonstrated this utility in a study where methotrexate was covalently attached to the surface of PEG-coated MNPs via a cleavable amide linkage. Recently, Sun et al. (2008b) further modified the same MFMNP system with chlorotoxin to enhance the NP's targeting abilities against brain tumor cells.

Ideal drug delivery systems should be stable with a long circulation time, and should keep the loaded drugs unreleased during circulation in the bloodstream or in normal tissues. Upon reaching the tumor tissues and being taken up by cancer cells, the systems should release the drugs rapidly to kill cancer cells.

To achieve this purpose, stimuli-triggered drug delivery systems have been used that respond to characteristics of the local microenvironment, such as pH, temperature, redox potential, etc. [reviewed by Danhier et al. (2010) and Muthu et al. (2009)]. In particular, pH gradients have been used widely to design responsive nanoparticle delivery systems. Various nanocarriers with pH-responsive delivery behaviors have been developed on the basis of the differential pH values of blood plasma (pH 7.4), extracellular tumor matrix (pH 5.8–7.2), and endocytic compartments such as endosomes (pH 5–6) and lysosomes (pH 4–5). Drugs have been loaded into polymeric nanocarriers by acidic pH-induced cleavable covalent bonds, creating smart drug delivery systems that respond to the endosomal/lysosomal pH (Yoo et al., 2002; Bae et al., 2003; Gillies et al., 2004; Hruby et al., 2005). The pH-induced cleavage of such bonds can accelerate drug release from the nanocarriers.

Drugs are otherwise loaded into the core of polymeric micelles by noncovalent (e.g., hydrophobic) interactions. Compared to chemical attachment, noncovalent entrapment is convenient and easy to achieve. Nasongkla et al. (2006) described the preparation of micelles of PEG-b-poly(D, L-lactide) that encapsulated doxorubicin and a cluster of SPIO nanoparticles by noncovalent hydrophobic interactions. The protonation of doxorubicin under acidic conditions increased its water-solubility and induced its release.

Recently, our group reported a MFMNP platform that can load drugs with ionizable groups and hydrophobic moieties by the combined action of ionic bonding and hydrophobic interactions (Guo et al., 2008, 2010) (Fig. 5). The use of double noncovalent interactions resulted in a high loading affinity at a neutral pH (7.4), preventing premature release into the bloodstream. At an endosomal/lysosomal pH (<5.5), protonation of polycarboxylate anions in the polymer chains led to ionic bond breakage and drug release. The release process was controlled, responded well to pH, and displayed good kinetics.

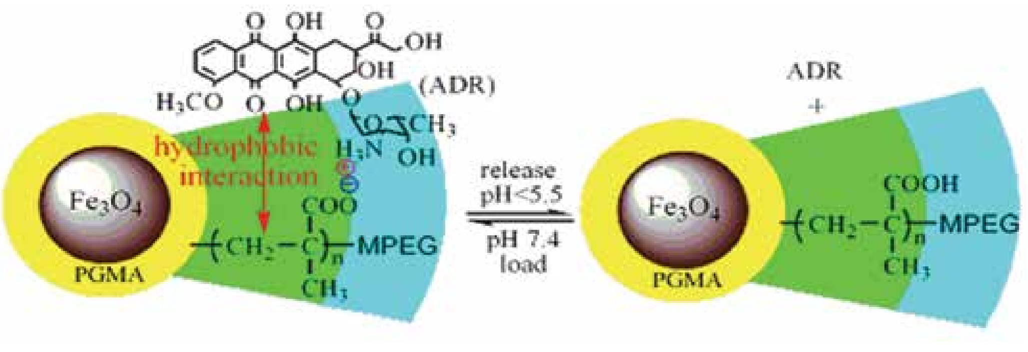


Fig. 5. Schematic illustration of the MFMNP structure, and the load and release of model drug adriamycin (ADR) (Guo et al., 2008).

3.2 MRI

MRI is a powerful noninvasive imaging modality that is utilized widely in clinical medicine. MRI is based on the property that hydrogen protons will align and process around an external alternating magnetic field. The subsequent process through which these protons return to their original state is referred to as the relaxation phenomenon. Two independent processes, longitudinal relaxation (T1-recovery) and transverse relaxation (T2-decay), are monitored to generate the MR image. Local variations in relaxation, corresponding to image contrast, arise from the proton density and the chemical and physical natures of the different tissues. Due to their ability to enhance proton relaxation and accumulation in specific tissues, MFMNPs have been actively investigated as contrast enhancement agents for MRI.

3.2.1 Magnetic properties of iron oxide nanoparticles

Particles whose unpaired electron spins align themselves spontaneously so that the material can exhibit magnetization without being in a magnetic field are called ferromagnetic particles. Materials such as iron oxide nanoparticles that exhibit ferromagnetism can be permanently magnetized. The magnetic properties of iron oxide nanoparticles can be described by the dependence of the magnetic induction B on the magnetic field H. For most materials, the relationship between B and H is linear: $B = \mu H$, where μ is the magnetic permeability of the particles. Iron oxide particles exhibit paramagnetism if $\mu > 1$, and diamagnetism if $\mu < 1$.

Usually, ferromagnetic properties arise only when a certain number of atoms are bound together in solid form; single atoms cannot exhibit ferromagnetism. When the size of particles is smaller than the ferromagnetic domain, they are no longer ferromagnetic but exhibit superparamagnetism (Bansil et al., 1998). Magnetic nanoparticles smaller than ~ 30 nm show superparamagnetic behavior without any magnetic remanence (i.e., restoration of the induced magnetization to zero upon removal of the external magnetic field), but the particles still exhibit very strong paramagnetic properties with a very large susceptibility. This is one important advantage for magnetic nanoparticles: it enables their stability and dispersion upon removal of the magnetic field, as no residual magnetic force exists between the particles.

In MRI, superparamagnetic nanoparticles made of iron oxide act as contrast enhancement agents by shortening both the T1 and T2 relaxations of the surrounding protons. The influence on the T1 relaxation depends strongly on the local MNP concentration, and the shortening processes can be hindered by the coating thickness. The effect of MNPs on T2 shortening is caused by the large susceptibility difference between the particles and surrounding medium, which results in microscopic magnetic field gradients. At low concentrations, a T1-positive contrast can be observed; at high concentrations, the susceptibility effects cause irreversible destruction of the MR signal around the particles.

NMP agglomeration tends to slightly decrease the T1 relaxation times but markedly decrease the T2 times. Therefore, superparamagnetic nanoparticles typically are used to provide negative contrast enhancement using T2-weighted pulse sequences. The effectiveness of a contrast agent can be described by its relaxivity, which is the proportionality constant of the measured rate of relaxation, or R1 (1/T1) and R2 (1/T2). The relaxivity depends on not only the composition, size, and magnetic properties of the MNP, but also depends on experimental variables such as the field strength, temperature, and medium in which the measurements are made.

3.2.2 Molecular imaging in cancer

The generation of new molecular targets that are closely related to pathophysiology will open the way for the development of new treatment paradigms for currently untreatable diseases. In recent years, many new diagnostic technologies have been developed, including molecular diagnostic compounds and new imaging technologies such as MR molecular imaging. Molecular imaging is the noninvasive imaging of targeted macromolecules, cells, and biological or cellular processes in living organisms. Due to their ability to act as molecularly targeted imaging agents, MNPs play an integral role in the applications of early disease detection, individualized treatment, and drug development. In the clinical imaging of tumors, MNPs can be used as contrast enhancement agents to improve the detection, diagnosis, and therapeutic management of solid tumors by exploiting the unique molecular signatures of the diseases. MNPs also have been investigated to improve the delineation of the tumor position, boundaries, and volume.

The first clinical indication for iron oxide nanoparticles was the imaging of liver tumors and metastases. After intravenous injection, MNPs are taken up rapidly by hepatic specialized macrophages. This process causes a drop in MR signal intensity and generates hypointense images, mostly because of a susceptibility effect. However, tumors lack a permanent decrease in signal intensity after MNP administration; tumors are almost devoid of macrophages, which are located exclusively in the healthy hepatic parenchyma. Therefore, MNPs can markedly increase the contrast between healthy and diseased tissue. The clinical imaging of liver tumors and metastases through RES-mediated uptake of MNPs has allowed the detection of lesions as small as 2–3 mm (Semelka & Helmberger, 2001). In combination with MRI, MNPs also have been shown to be effective in the identification of lymph node metastases of 5–10 mm in diameter (Harisinghani et al., 2003). The use of MFMNPs as contrast enhancement agents provides increased lesion conspicuousness and lesion detection compared to nonenhanced imaging.

MFMNPs are currently under evaluation for use in improving the delineation of brain tumor boundaries and quantifying tumor volumes (Enochs et al., 1999; Neuwelt et al., 2004). Some recent approaches have explored utilizing iron oxide nanoparticles as drug delivery vehicles for the MRI-monitored magnetic targeting of brain tumors (Chertok et al., 2008). The accumulation of iron oxide nanoparticles in gliosarcomas is enhanced by magnetic targeting and successfully quantified by MRI (Chertok et al., 2008) (Fig. 6). Such noninvasive approaches for cancer diagnosis and therapy also have been adopted in the treatment of prostate, breast, and colon cancers.

As both drug delivery devices and MRI contrast enhancement agents, MNPs retain the ability to track the movement of drug through the body. This is significant because it allow clinicians to monitor the effectivity of injected therapeutics to reach their target sites. There remains significant flexibility in the contrast agents implemented in these constructs and the manner in which drugs are delivered. Medarova et al. (2007) recently developed cross-linked iron oxide nanoparticles modified with a NIR fluorophore, therapeutic siRNA sequences, and a cell penetrating peptide. The MNPs used passive targeting by the EPR effect to direct tumor localization. In vivo, these MNPs demonstrated therapeutic efficacy against target tissue, as determined by real time PCR and histological evaluation, while simultaneously demonstrating image contrast in both MR and optical imaging. In a study by Sun et al. (2008a) active cell targeting was shown by PEG-coated MNPs to which the chemotherapeutic, methotrexate, and targeting molecule, chlorotoxin, were attached. The selective contrast enhancement of the 9L brain tumor by these MNPs indicates preferential accumulation compared with the same MNP construct without the chlorotoxin peptide in a 3-day study.

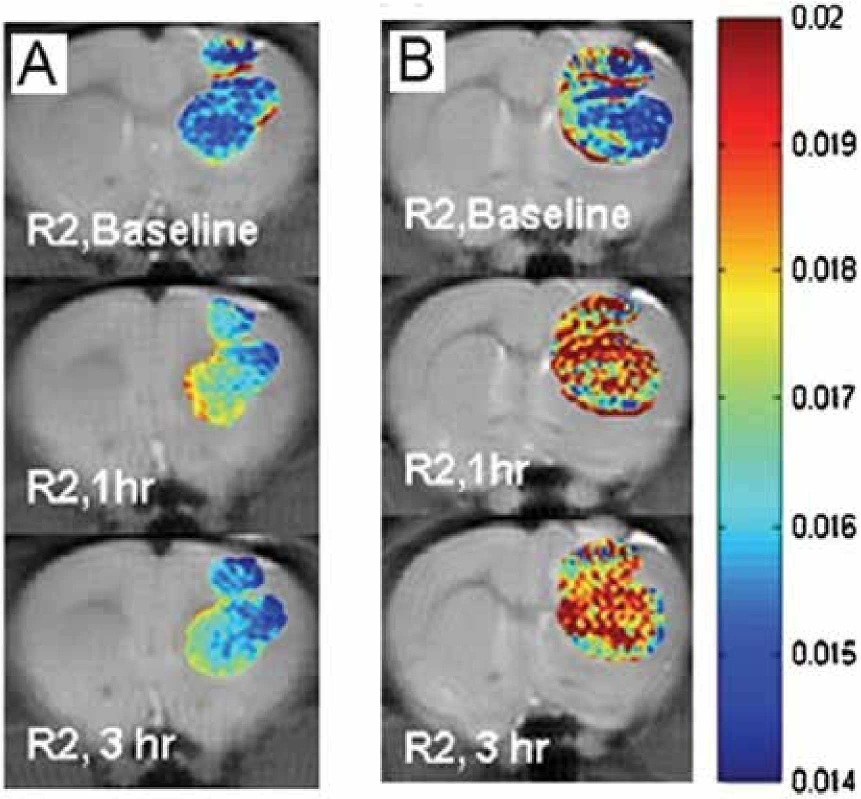


Fig. 6. MR images of brain tumor. Change in the R2 relaxation for the tumor regions before (baseline) and 1–3 h after MNPs administration in (A) control and (B) targeted rats (Chertok et al., 2008).

In another recent study by Yang et al. (2007) simultaneous targeted drug delivery and MR imaging of breast cancer tumors were demonstrated through the use multifunctional magneto-polymeric nanohybrids composed of magnetic nanocrystals and doxorubicin which were simultaneously encapsulated within an amphiphilic block copolymer shell. The surfaces of these micelles were additionally functionalized with the breast cancer targeting/therapeutic ligand, anti-Herceptin antibody. In vivo evaluations of this nanoparticle system were performed in nude mice bearing NIH3T6.7 breast cancer tumors. The quantitative evaluation of MR images revealed preferential accumulation of the targeted MNPs compared to the control MNPs. The therapeutic functionality of the MNPs developed in this study were additionally evaluated and it was determined that the HER-MMPNs which were decorated with targeting ligands and loaded with doxorubicin were most effective in inhibiting tumor growth. Combined, these findings illustrate the functionality and efficacy of targeted multifunctional MNPs for simultaneous MR imaging and drug delivery.

3.3 Hyperthermia

Hyperthermia is the method of using heat as a treatment for malignant tumors. It is based on the observation that tumor cells are more susceptible to heat than normal cells, due to the higher rates of metabolism of cancer cells. Cancer cells typically show signs of apoptosis and necrosis when heated to 41-47 °C, whereas normal cells can survive at higher temperatures (Milleron & Bratton, 2007). Hyperthermia with targeted nanoscale heaters is recognized as a useful therapeutic modality to kill cancer by essentially "cooking" malignant cells from the inside out.

Magnetic nanoparticle hyperthermia is actualized by the exposure of cancer tissues to an alternating magnetic field. The magnetic field cannot be absorbed by the living tissues and can be applied to deep regions in the living body. When MNPs are injected into an organ with a tumor, they tend to accumulate in the tumor due to passive and active targeting strategies (as described above). Subsequent exposure to an alternating magnetic field causes heat to be generated in the tumor tissue due to magnetic hysteresis loss. This process effectively destroys the tumor but not the surrounding healthy tissue. The amount of heat generated depends on the nature of the MNPs and magnetic field parameters used.

The use of MFMNPs for targeted hyperthermia has shown a therapeutic effect in several types of tumors. Using dextran-coated MNPs conjugated to breast cancer–targeting chimeric L6 mAb, DeNardo et al. (2005) demonstrated the feasibility of this method for treating breast cancer cells. Kobayashi et al. constructed a novel therapeutic tool of magnetite nanoparticle-loaded anti-HER2 immunoliposomes that was applicable to the treatment of HER2-overexpressing cancer (Ito et al., 2004).

A clinical breakthrough in MNP use was made in 2007, when Maier-Hauff et al. (2007) reported the results of using heated implanted MNPs for therapeutic hyperthermia in humans. In that study, 14 patients with recurrent glioblastoma multiforme, a type of severe brain cancer, received an intratumoral injection of aminosilane-coated MNPs. The tumor sites were located by several comprehensive MRI scans, and the patients were exposed to an alternating magnetic field to induce particle heating. The nanoparticle deposits were stable for several weeks, and all patients tolerated the nanoparticles without any complications. These findings indicate that MNP hyperthermia may be an effective therapeutic method to cure human brain cancer.

As a potential approach for the treatment of malignant tumors, MNP hyperthermia has the following advantages: it provides a noninvasive way to raise cell temperatures to a therapeutic level; MNPs can be visualized using MRI, thus combining diagnostic and therapeutic approaches in one type of particle; and the particles can be functionalized and combined with other types of treatment, such as chemotherapy or radiotherapy.

4. Conclusions

Multifunctional nanomaterials have been widely used as nanoplatforms for multimodal imaging or simultaneous imaging and therapy. As a multifunctional nanoplatform with many biomedical applications, MFMNPs of various formulations have been developed to diagnose and treat diseases for which conventional therapy has shown limited efficacy, such as cancer. The use of MFMNPs as MRI contrast enhancement agents and anticancer drug carriers has drawn enormous attention, and may provide new opportunities for early cancer detection and targeted therapies. This technology not only will minimize the need for invasive procedures, but also will reduce the side effects to healthy tissues, which are primary concerns in conventional cancer therapies.

5. Acknowledgment

This work was supported by the National Natural Science Foundation of China (20974052), the National Key Technologies R & D Program for New Drugs of China (2009ZX09301-002) and the Natural Science Foundation of Tianjin Municipality (09JCZDJC22900).

6. References

Allen, T. M. & Cullis, P. R. (2004). Drug Delivery Systems: Entering the Mainstream. *Science,* Vol. 303, No. 5665, pp. 1818-1822.

Bae, Y.; Fukushima, S.; Harada, A. & Kataoka, K. (2003). Design of Environment-Sensitive Supramolecular Assemblies for Intracellular Drug Delivery: Polymeric Micelles that Are Responsive to Intracellular pH Change. *Angewandte Chemie International Edition,* Vol. 42, No. 38, pp. 4640-4643.

Bansil, A.; Sawatzky, G.; Fujii, Y.; Shelton, R. & Prosser, D. (1998). Journal of Physics and Chemistry of Solids 40th Anniversary – Preface. *Journal of Physics and Chemistry of Solids,* Vol. 59, No. 4, pp. III-IIV.

Berry, C. C. & Curtis, A. S. G. (2003). Functionalisation of Magnetic Nanoparticles for Applications in Biomedicine. *Journal of Physics D-Applied Physics,* Vol. 36, No. 13, pp. R198-R206.

Bertorelle, F.; Wilhelm, C.; Roger, J.; Gazeau, F.; Menager, C. & Cabuil, V. (2006). Fluorescence-Modified Superparamagnetic Nanoparticles: Intracellular Uptake and Use in Cellular Imaging. *Langmuir,* Vol. 22, No. 12, pp. 5385-5391.

Bhadriraju, K. & Hansen, L. K. (2000). Hepatocyte Adhesion, Growth and Differentiated Function on RGD-Containing Proteins. *Biomaterials,* Vol. 21, No. 3, pp. 267-272.

Bumb, A.; Brechbiel, M. W.; Choyke, P. L.; Fugger, L.; Eggeman, A.; Prabhakaran, D.; Hutchinson, J. & Dobson, P. J. (2008). Synthesis and Characterization of Ultra-Small Superparamagnetic Iron Oxide Nanoparticles Thinly Coated with Silica. *Nanotechnology,* Vol. 19, No. 33, pp. 335601-335606.

Butterworth, M. D.; Illum, L. & Davis, S. S. (2001). Preparation of Ultrafine Silica- and PEG-Coated Magnetite Particles. *Colloids and Surfaces A-Physicochemical and Engineering Aspects,* Vol. 179, No. 1, pp. 93-102.

Chertok, B.; Moffat, B. A.; David, A. E.; Yu, F. Q.; Bergemann, C.; Ross, B. D. & Yang, V. C. (2008). Iron Oxide Nanoparticles as a Drug Delivery Vehicle for MRI Monitored Magnetic Targeting of Brain Tumors. *Biomaterials,* Vol. 29, No. 4, pp. 487-496.

Couvreur, P. & Vauthier, C. (2006). Nanotechnology: Intelligent Design to Treat Complex Disease. *Pharmaceutical Research,* Vol. 23, No. 7, pp. 1417-1450.

Danhier, F.; Feron, O. & Preat V. (2010). To Exploit the Tumor Microenvironment: Passive and Active Tumor Targeting of Nanocarriers for Anti-Cancer Drug Delivery. *Journal of Controlled Release,* Vol. 148, No. 2, pp. 135–146.

Debelle, L. & Tamburro, A. M. (1999). Elastin: Molecular Description and Function. *International Journal of Biochemistry & Cell Biology,* Vol. 31, No. 2, pp. 261-272.

DeNardo, S. J.; DeNardo, G. L.; Miers, L. A.; Natarajan, A.; Foreman, A. R.; Gruettner, C.; Adamson, G. N. & Ivkov, R. (2005). Development of Tumor Targeting Bioprobes (In-111-Chimeric L6 Monoclonal Antibody Nanoparticles) for Alternating Magnetic Field Cancer Therapy. *Clinical Cancer Research,* Vol. 11, No. 19, pp. 7087S-7092S.

Enochs, W. S.; Harsh, G.; Hochberg, F. & Weissleder, R. (1999). Improved Delineation of Human Brain Tumors on MR Images Using a Long-Circulating, Superparamagnetic Iron Oxide Agent. *Journal of Magnetic Resonance Imaging,* Vol. 9, No. 2, pp. 228-232.

Ferrari, M. (2005). Cancer Nanotechnology: Opportunities and Challenges. *Nature Reviews Cancer,* Vol. 5, No. 3, pp. 161-171.

Flesch, C.; Unterfinger, Y.; Bourgeat-Lami, E.; Duguet, E.; Delaite, C. & Dumas, P. (2005). Poly(ethylene glycol) Surface Coated Magnetic Particles. *Macromolecular Rapid Communications,* Vol. 26, No. 18, pp. 1494-1498.

Gabizon, A. & Papahadjopoulos, D. (1992). The Role of Surface-Charge and Hydrophilic Groups on Liposome Clearance In Vivo. *Biochimica et Biophysica Acta,* Vol. 1103, No. 1, pp. 94-100.

Gillies, E. R.; Goodwin, A. P. & Frechet, J. M. J. (2004). Acetals as pH-Sensitive Linkages for Drug Delivery. *Bioconjugate Chemistry,* Vol. 15, No. 6, pp. 1254-1263.

Gref, R.; Minamitake, Y.; Peracchia, M. T.; Trubetskoy, V.; Torchilin, V. & Langer, R. (1994). Biodegradable Long-Circulating Polymeric Nanospheres. *Science,* Vol. 263, No. 5153, pp. 1600-1603.

Gregoria, G.; Wills, E. J.; Swain, C. P. & Tavill, A. S. (1974). Drug-Carrier Potential of Liposomes in Cancer Chemotherapy. *Lancet,* Vol. 1, No. 7870, pp. 1313-1316.

Guo, M.; Que, C. L.; Wang, C. H.; Liu, X. Z.; Yan, H. S. & Liu, K. L. (2011). Multifunctional Superparamagnetic Nanocarriers with Folate-Mediated and pH-Responsive Targeting Properties for Anticancer Drug Delivery. *Biomaterials,* Vol. 32, No. 1, pp. 185-194.

Guo, M.; Yan, Y.; Liu, X. Z.; Yan, H. S.; Liu, K. L.; Zhang, H. K.; & Cao, Y. J. (2010). Multilayer Nanoparticles with a Magnetite Core and a Polycation Inner Shell as pH-Responsive Carriers for Drug Delivery. *Nanoscale,* Vol. 2, No. 3, pp. 434-441.

Guo, M.; Yan, Y.; Zhang, H. K.; Yan, H. S.; Cao, Y. J.; Liu, K. L.; Wan, S. R.; Huang, J. S. & Yue, W. (2008). Magnetic and pH-responsive Nanocarriers with Multilayer Core-Shell Architecture for Anticancer Drug Delivery. *Journal of Materials Chemistry,* Vol. 18, No. 42, pp. 5104-5112.

Gupta, A. K.; Berry, C.; Gupta, M. & Curtis, A. S. G. (2003). Receptor-Mediated Targeting of Magnetic Nanoparticles Using Insulin as a Surface Ligand to Prevent Endocytosis. *IEEE Transactions on Nanobioscience,* Vol. 2, No. 4, pp. 255-261.

Gupta, A. K. & Gupta, M. (2005). Synthesis and Surface Engineering of Iron Oxide Nanoparticles for Biomedical Applications. *Biomaterials,* Vol. 26, No. 18, pp. 3995-4021.

Harisinghani, M. G. & Weissleder, R. (2004). Sensitive, Noninvasive Detection of Lymph Node Metastases. *Plos Medicine,* Vol. 1, No. 3, pp. 202-209.

Hruby, M.; Konak, C. & Ulbrich, K. (2005). Polymeric Micellar pH-Sensitive Drug Delivery System for Doxorubicin. *Journal of Controlled Release,* Vol. 103, No. 1, pp. 137-148.

Huh, Y. M.; Jun, Y. W.; Song, H. T.; Kim, S.; Choi, J. S.; Lee, J. H.; Yoon, S.; Kim, K. S.; Shin, J. S.; Suh, J. S. & Cheon, J. (2005). In Vivo Magnetic Resonance Detection of Cancer by Using Multifunctional Magnetic Nanocrystals, *Journal of the American Chemical Society,* Vol. 127, No. 35, pp. 12387-12391.

Ito, A.; Kuga, Y.; Honda, H.; Kikkawa, H.; Horiuchi, A.; Watanabe, Y. & Kobayashi, T. (2004). Magnetite Nanoparticle-Loaded Anti-Her2 Immunoliposomes for Combination of Antibody Therapy with Hyperthermia. *Cancer Letters,* Vol. 212, No. 2, pp. 167-175.

Jain, K. K. (2005). The Role of Nanobiotechnology in Drug Discovery. *Drug Discovery Today,* Vol. 10, No. 21-24, pp. 1435-1442.

Ji, X. J.; Shao, R. P.; Elliott, A. M.; Stafford, R. J.; Esparza-Coss, E.; Bankson, J. A.; Liang, G.; Luo, Z. P.; Park, K.; Markert, J. T. & Li, C. (2007). Bifunctional Gold Nanoshells with a Superparamagnetic Iron Oxide-Silica Core Suitable for both MR Imaging and Photothermal Therapy. *Journal of Physical Chemistry*, Vol. 111, No. 17, pp. 6245-6251.

Josephson, L.; Tung, C. H.; Moore, A. & Weissleder, R. (1999). High-Efficiency Intracellular Magnetic Labeling with Novel Superparamagnetic-Tat Peptide Conjugates. *Bioconjugate Chemistry,* Vol. 10, No. 2, pp. 186-191.

Kaneo, Y.; Tanaka, T.; Nakano, T. & Yamaguchi, Y. (2001). Evidence for Receptor-Mediated Hepatic Uptake of Pullulan in Rats. *Journal of Controlled Release,* Vol. 70, No. 3, pp. 365-373.

Kirpotin, D. B.; Drummond, D. C.; Shao, Y.; Shalaby, M. R.; Hong, K. L.; Nielsen, U. B.; Marks, J. D.; Benz, C. C. & Park, J. W. (2006). Antibody Targeting of Long-Circulating Lipidic Nanoparticles does not Increase Tumor Localization but does Increase Internalization in Animal Models. *Cancer Research*, Vol. 66, No. 13, pp. 6732-6740.

Kohler, N.; Sun, C.; Fichtenholtz, A.; Gunn, J.; Fang, C. & Zhang, M. Q. (2006). Methotrexateimmobilized poly(ethylene glycol) magnetic nanoparticles for MR imaging and drug delivery. *Small,* Vol. 2, No. 6, pp. 785-792.

Lewin, M.; Carlesso, N.; Tung, C. H.; Tang, X. W.; Cory, D.; Scadden, D. T. & Weissleder, R. (2000). Tat Peptide-Derivatized Magnetic Nanoparticles Allow In Vivo Tracking and Recovery of Progenitor Cells. *Nature Biotechnology,* Vol. 18, No. 4, pp. 410-414.

Lu, A. H.; Salabas, E. L. & Schuth, F. (2007). Magnetic Nanoparticles: Synthesis, Protection, Functionalization, and Application. *Angewandte Chemie International Edition,* Vol. 46, No. 8, pp. 1222-1244.

Lu, Y.; Yin, Y. D.; Mayers, B. T. & Xia, Y. N. (2002). Modifying the Surface Properties of Superparamagnetic Iron Oxide Nanoparticles through a Sol-Gel Approach. *Nano Letters,* Vol. 2, No. 3, pp. 183-186.

Ma, D. L.; Guan, J. W.; Normandin, F.; Denommee, S.; Enright, G.; Veres, T. & Simard, B. (2006). Multifunctional Nano-Architecture for Biomedical Applications. *Chemistry of Materials,* Vol. 18, No. 7, pp. 1920-1927.

Maeda, H. (2001). The Enhanced Permeability and Retention (EPR) Effect in Tumor Vasculature: The Key Role of Tumor-Selective Macromolecular Drug Targeting. *Advances in Enzyme Regulation,* Vol 41, v. 41, pp. 189-207.

Maier-Hauff, K.; Rothe, R.; Scholz, R.; Gneveckow, U.; Wust, P.; Thiesen, B.; Feussner, A.; von Deimling, A.; Waldoefner, N.; Felix, R. & Jordan, A. (2007). Intracranial Thermotherapy Using Magnetic Nanoparticles Combined with External Beam Radiotherapy: Results of a Feasibility Study on Patients with Glioblastoma Multiforme. *Journal of Neuro-Oncology,* Vol. 81, No. 1, pp. 53-60.

Massart, R. (1981). Preparation of Aqueous Magnetic Liquids in Alkaline and Acidic Media. *IEEE Transactions on Magnetics*, Vol. 17, No. 2, pp. 1247-1248.

Medarova, Z.; Pham, W.; Farrar, C.; Petkova, V. & Moore, A. (2007). In Vivo Imaging of siRNA Delivery and Silencing in Tumors. *Nature Medicine,* Vol. 13, No. 3, pp. 372-377.

Milleron, R. S. & Bratton, S. B. (2007). 'Heated' Debates in Apoptosis. *Cellular and Molecular Life Sciences,* Vol. 64, pp. 2329-2333.

Montet, X.; Montet-Abou, K.; Reynolds, F.; Weissleder, R. & Josephson, L. (2006). Nanoparticle Imaging of Integrins on Tumor Cells. *Neoplasia,* Vol. 8, No. 3, pp. 214-222.

Moore, A.; Josephson, L.; Bhorade, R. M.; Basilion, J. P. & Weissleder, R. (2001). Human Transferrin Receptor Gene as a Marker Gene for MR Imaging. *Radiology,* Vol. 221, No. 1, pp. 244-250.

Muthu, M. S.; Rajesh, C. V.; Mishra A. & Singh S. (2009). Stimulus-Responsive Targeted Nanomicelles for Effective Cancer Therapy. Nanomedicine, Vol. 4, No. 6, 657-667.

Nasongkla, N.; Bey, E.; Ren, J. M.; Ai, H.; Khemtong, C.; Guthi, J. S.; Chin, S. F.; Sherry, A. D.; Boothman, D. A. & Gao, J. M. (2006). Multifunctional Polymeric Micelles as Cancer-Targeted, MRI-Ultrasensitive Drug Delivery Systems. *Nano Letters,* Vol. 6, No. 11, pp. 2427-2430.

Needham, D.; Mcintosh, T. J. & Lasic, D. D. (1992). Repulsive Interactions and Mechanical Stability of Polymer-Grafted Lipid-Membranes. *Biochimica et Biophysica Acta,* Vol. 1108, No. 1, pp. 40-48.

Neuberger, T.; Schopf, B.; Hofmann, H.; Hofmann, M. & von Rechenberg, B. (2005). Superparamagnetic Nanoparticles for Biomedical Applications: Possibilities and Limitations of a New Drug Delivery System. *Journal of Magnetism and Magnetic Materials,* Vol. 293, No. 1, pp. 483-496.

Neuwelt, E. A.; Varallyay, P.; Bago, A. G.; Muldoon, L. L.; Nesbit, G. & Nixon, R. (2004). Imaging of Iron Oxide Nanoparticles by MR and Light Microscopy in Patients with Malignant Brain Tumours. *Neuropathology and Applied Neurobiology,* Vol. 30, No. 5, pp. 456-471.

Prime, K. L. & Whitesides, G. M. (1991). Self-Assembled Organic Monolayers - Model Systems for Studying Adsorption of Proteins at Surfaces. *Science,* Vol. 252, No. 5009, pp. 1164-1167.

Reddy, G. R.; Bhojani, M. S.; McConville, P.; Moody, J.; Moffat, B. A.; Hall, D. E.; Kim, G.; Koo, Y. E. L.; Woolliscroft, M. J.; Sugai, J. V.; Johnson, T. D.; Philbert, M. A.; Kopelman, R.; Rehemtulla, A. & Ross, B. D. (2006). Vascular Targeted Nanoparticles for Imaging and Treatment of Brain Tumors. *Clinical Cancer Research,* Vol. 12, No. 22, pp. 6677-6686.

Schellenberger, E. A.; Bogdanov, A.; Hogemann, D.; Tait, J.; Weissleder, R. & Josephson, L. (2002). Annexin V-CLIO: a Nanoparticle for Detecting Apoptosis by MRI. *Molecular Imaging,* Vol. 1, No. 2, pp. 102-107.

Semelka, R. C. & Helmberger, T. K. G. (2001). Contrast Agents for MR Imaging of the Liver. *Radiology,* Vol. 218, No. 1, pp. 27-38.

Simberg, D.; Duza, T.; Park, J. H.; Essler, M.; Pilch, J.; Zhang, L.; Derfus, A. M.; Yang, M.; Hoffman, R. M.; Bhatia, S.; Sailor, M. J. & Ruoslahti, E. (2007). Biomimetic Amplification of Nanoparticle Homing to Tumors. *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 104, No. 3, pp. 932-936.

Sun, C.; Fang, C.; Stephen, Z.; Veiseh, O.; Hansen, S.; Lee, D.; Ellenbogen, R. G.; Olson, J. & Zhang, M. Q. (2008a). Tumor-Targeted Drug Delivery and MRI Contrast Enhancement by Chlorotoxin-Conjugated Iron Oxide Nanoparticles. *Nanomedicine,* Vol. 3, No. 4, pp. 495-505.

Sun, C.; Veiseh, O.; Gunn, J.; Fang, C.; Hansen, S.; Lee, D.; Sze, R.; Ellenbogen, R. G.; Olson, J. & Zhang, M. (2008b). In Vivo MRI Detection of Gliomas by Chlorotoxin-Conjugated Superparamagnetic Nanoprobes. *Small,* Vol. 4, No. 3, pp. 372-379.

Tartaj, P.; Morales, M. D.; Veintemillas-Verdaguer, S.; Gonzalez-Carreno, T. & Serna, C. J. (2003). The Preparation of Magnetic Nanoparticles for Applications in Biomedicine. *Journal of Physics D-Applied Physics,* Vol. 36, No. 13, pp. R182-R197.

Thorek, D. L. J.; Chen, A.; Czupryna, J. & Tsourkas, A. (2006). Superparamagnetic Iron Oxide Nanoparticle Probes for Molecular Imaging. *Annals of Biomedical Engineering,* Vol. 34, No. 1, p. 23-38.

Toma, A.; Otsuji, E.; Kuriu, Y.; Okamoto, K.; Ichikawa, D.; Hagiwara, A.; Ito, H.; Nishimura, T. & Yamagishi, H. (2005). Monoclonal Antibody A7-Superparamagnetic Iron Oxide as Contrast Agent of MR Imaging of Rectal Carcinoma. *British Journal of Cancer,* Vol. 93, No. 1, pp. 131-136.

Torchilin, V. P. (2006). Multifunctional nanocarriers. *Advanced Drug Delivery Reviews,* Vol. 58, No. 14, pp. 1532-1555.

van Vlerken, L. E.; Duan, Z. F.; Seiden, M. V. & Amiji, M. M. (2007). Modulation of Intracellular Ceramide Using Polymeric Nanoparticles to Overcome Multidrug Resistance in Cancer. *Cancer Research,* Vol. 67, No. 10, pp. 4843-4850.

Vasir, J. K.; Reddy, M. K. & Labhasetwar, V. D. (2005). Nanosystems in Drug Targeting: Opportunities and Challenges. *Current Nanoscience,* Vol. 1, No. 1, pp. 47-64.

Wan, S. R.; Zheng, Y.; Liu, Y. Q.; Yan, H. S. & Liu, K. L. (2005). Fe_3O_4 Nanoparticles Coated with Homopolymers Of Glycerol Mono(meth)acrylate and Their Block Copolymers. *Journal of Materials Chemistry,* Vol. 15, No. 33, pp. 3424-3430.

Weissleder, R.; Moore, A.; Mahmood, U.; Bhorade, R.; Benveniste, H.; Chiocca, E. A. & Basilion, J. P. (2000). In Vivo Magnetic Resonance Imaging of Transgene Expression. *Nature Medicine,* Vol. 6, No. 3, pp. 351-354.

Weissleder, R. & Ntziachristos, V. (2003). Shedding Light onto Live Molecular Targets. *Nature Medicine,* Vol. 9, No. 1, pp. 123-128.

Wunderbaldinger, P.; Josephson, L. & Weissleder, R. (2002). Tat Peptide Directs Enhanced Clearance and Hepatic Permeability of Magnetic Nanoparticles. *Bioconjugate Chemistry,* Vol. 13, No. 2, pp. 264-268.

Yamaoka, T.; Tabata, Y. & Ikada, Y. (1994). Distribution and Tissue Uptake of Poly(ethylene glycol) with Different Molecular-Weights after Intravenous Administration to Mice. *Journal of Pharmaceutical Sciences,* Vol. 83, No. 4, pp. 601-606.

Yang, J.; Lee, C. H.; Ko, H. J.; Suh, J. S.; Yoon, H. G.; Lee, K.; Huh, Y. M. & Haam, S. (2007). Multifunctional Magneto-Polymeric Nanohybrids for Targeted Detection and Synergistic Therapeutic Effects on Breast Cancer. *Angewandte Chemie International Edition,* Vol. 46, No. 46, pp. 8836-8839.

Yoo, H. S.; Lee, E. A. & Park, T. G. (2002). Doxorubicin-Conjugated Biodegradable Polymeric Micelles Having Acid-Cleavable Linkages. *Journal of Controlled Release*, Vol. 82, No. 1, pp. 17-27.

Zhang, Y.; Kohler, N. & Zhang, M. Q. (2002). Surface Modification of Superparamagnetic Magnetite Nanoparticles and Their Intracellular Uptake. *Biomaterials,* Vol. 23, No. 7, pp. 1553-1561.

13

Arterial Mass Transport Behaviour of Drugs from Drug Eluting Stents

Barry M. O'Connell and Michael T. Walsh
Centre for Applied Biomedical Engineering Research (CABER),
Department of Mechanical Aeronautical and Biomedical Engineering,
and Materials and Surface Science Institute,
University of Limerick, Limerick,
Ireland

1. Introduction

Coronary artery disease (CAD) is the foremost cause of morbidity in the worlds industrialised nations. Consequently our ability in understanding the treatments available and the mechanisms of their success/failure is of particular importance as we strive to improve procedural success rates through the evolution of existing and the innovation of new interventional technologies. Today, the gold standard for treating CAD is to deploy a drug eluting stent (DES) in the blocked artery to first restore luminal blood flow and second to resist the body's tendency to block the artery once more through its overly aggressive healing response known as restenosis. Intrinsically, the understanding of mass transport is elemental to all aspects of DES design from the type of drug used and the polymer release characteristics to the shape and thickness of the stent struts. Design optimisation of DES is vital to achieving increases in procedural success rates as the technology evolves, as is the case for its newest embodiment, the biodegradable DES.

This chapter addresses issues fundamental to drug mass transport from DES, starting with an overview of CAD and associated interventional procedures. DES mass transport theory is then presented and followed by accounts of mass transport problem classifications employed by researchers under a number of deployment scenarios. The importance of experimental mass transport validation is highlighted and a computational investigation is developed to demonstrate how differences in DES deployment conditions alter drug concentrations in the artery wall.

2. Coronary artery disease

Atherosclerosis is a degenerative disease that affects coronary, carotid and other peripheral arteries in the body. Disease formation can occur as early as childhood with the development of fatty streaks within the artery wall. As the aging process progresses, these fatty streaks accumulate to become larger lipid deposits within the artery. This gradual propagation of plaque can be detrimental to the smooth operation of the vasculature. Occlusions ensuing from aggressive atherosclerotic plaque progression can often culminate in an ischemic attack, such as a stroke or a heart attack. CAD pertains to a blockage or narrowing of the coronary arteries that provide oxygen and nutrients vital to the smooth operation of the heart muscle. Once identified as such, there are a number of interventional procedures available to the cardiologist but the successful emergence of stents, and more recently DES, has seen them become the preferred choice of treatment for CAD, so much so that by the beginning of 2006 more than 8 out of 10 coronary stents were DES (Head et al., 2007) at a cost of between $4 and $5 billion annually (Kaul et al., 2007).

CAD has been intrinsically linked to atherosclerosis since the early 20th century (Chen et al., 2005) and refers to the localisation of disease within the coronary arteries. It is generally asymptomatic and those afflicted often only realise they have the condition when it manifests itself in the form of a heart attack. For this reason CAD is the foremost cause of mortality in the world's industrialised nations (Khakpour &Vafai, 2008). CAD alone is reportedly responsible for approximately 700,000 deaths in the United States of America annually (Kaazempur-Mofrad et al., 2005). Lifestyle choices made by an individual, such as not smoking, regular exercise and a balanced diet, have been shown to influence CAD development but it is the concentration of lipid rich cholesterol in the blood that is considered the most important factor (Sun et al., 2006).

2.1 Traditional interventional procedures

Numerous ways exist to alleviate a stenosis in a coronary artery once detected. The first such interventional procedure commonly practiced was coronary artery bypass graft (CABG) surgery. CABG surgery was successfully performed first by Robert H. Goetz and his team in 1960 at the Albert Einstein College of Medicine (Haller & Olearchyk, 2002). Prior to surgery an angiogram is conducted in order to locate the occlusion within the artery, after which a median sternotomy is performed, which exposes the heart and enables the blocked coronary arteries to be bypassed. This procedure is traumatic for the patient with extensive recovery times and significant scarring to the chest. Furthermore, the long term patency rates of these grafts were moderate and several efforts have been made in vain to optimise downstream graft artery junctions. Despite this, CABG remained the gold standard in the treatment of CAD until 1977 when Andreas Gruntzig first performed percutaneous transluminal coronary angioplasty (PTCA) (Kukreja et al., 2008). CABG is a procedure still used today as not every patient is eligible for minimally invasive surgeries such as PTCA due to highly tortuous or extensively blocked arteries.

Initially PTCA was welcomed by the clinical community due to its minimally invasive approach to arterial stenosis alleviation. A balloon catheter is introduced through an incision in the femoral artery and is manoeuvred through the vasculature until it reaches the stenosis. Once inflated, the balloon pushes the plaque back against the artery wall and enables blood flow to recommence after it has been deflated. The initial success of PTCA was short lived as investigators soon discovered that a substantial percentage of patients, reported to be anywhere between 30% and 60%, experienced recurrent ischemia due to the re-blocking of the artery (restenosis) within the first 6 months. This was attributed to mechanical injury caused by over dilating a device within the vessel (Head et al., 2007).

The next major advance in the field of minimally invasive interventional cardiology came in the early 1990's with the advent of the coronary artery stent (CAS). Prior to surgery a cylindrical metallic scaffold, or stent, is placed on the end of an existing balloon catheter and deployed in the same way as traditional PTCA. Initially, these stents were mounted on the balloon catheter by the physician, however, in more recent times the manufacturers supply the catheters with a stent already in situ. After deployment the stent remains within the artery in an attempt to retain arterial patency. CAS reduced failure rates tobetween 10% and 40% (Duraiswamy et al., 2007; Mongrain et al., 2007) through the elimination of elastic recoil and negative remodelling of the artery associated with PTCA (Costa and Simon, 2005).

2.2 Restenosis and the advent of the drug eluting stent era

Restenosis can best be described as an overly aggressive inflammatory healing response in the artery wall due to the mechanical injury inflicted by balloon/stent expansion. It can be quantified by the reduction of lumen size after an intravascular interventional procedure. The development of restenosis can be described by three processes after PTCA; 1) elastic recoil, 2) arterial negative remodelling and 3) neointimal hyperplasia (Rajagopal&Rockson, 2003). Elastic recoil can occur within an hour of PTCA and is due to passive recoil of the elastic medial layer of the artery. Arterial remodelling on the other hand can be both positive (vessel enlargement) or negative (vessel shrinking) and is characterised as such by a change in vascular dimension. Investigators report contrasting views on the mechanisms behind negative remodelling but whatever the underlying pathology behind vascular remodelling, it is believed to be virtually eliminated when angioplasty is used in conjunction with a stent.

Over inflation of a balloon catheter can result in the fracture of atherosclerotic plaque and in some cases can cause partial fracture of the artery wall (Schwartz et al., 2004). The same crushing/fracturing effect is witnessed when a stent is used in conjunction with an over inflated balloon. However, a stent can also cause excessive injury by penetrating the media which in turn increases neointimal formation. In some extreme cases stents have been known to penetrate as deep as the adventitial layer of the artery (Costa and Simon, 2005). The introduction of the DES to market has gone someway to alleviating the issue of arterial restenosis and excessive vessel injury via stent expansion. Variations of anti-restenotic drugs have been used to coat the stent in order to prevent post-operative in stent restenosis (ISR) and these modern stents can have strut profiles in the order of 80μm which would minimise the possibility of adverse artery wall penetration.

It is generally accepted that one of the main causes of restenosis following BMS implantation is SMC proliferation from the medial artery layer to the injured site. Attempts at systemic drug delivery to inhibit restenosis after stenting failed because effective dosing levels had a toxic effect and could not be tolerated by the patients (Waksman, 2002). Therefore the concept of local drug delivery was developed to redress the issue through the application of a drug eluting coating to the stent platform. This enables site specific local delivery of drugs that can be applied to the injured vessel at the exact location and time that damage occurs. The anti-restenotic coating on DES inhibits the formation of neointimal hyperplasia via suppression of the inflammatory reaction, platelet activation and SMC proliferation, curbing the overly aggressive healing response. Most of the early drugs explored originally were those used as agents for anti-transplant rejection or immunosuppressive drugs (van der Hoeven et al., 2004).

In April 2003 the first DES to gain commercial approval from the Food and Drug Administration (FDA) in the United States was the Cypher stent, which was developed by Cordis Corporation (Miami, FL. USA). The drug used on the Cypher stent is called sirolimus. Boston Scientific's (Natik, MA. USA) TAXUS family of stents were the second DES platform approved by the FDA in March of the following year. The drug employed on the TAXUS stent is called paclitaxel (Venkatraman and Boey, 2007).The first generation DES had a profound effect on reducing restenosis rates compared to bare metal stent (BMS) models. Clinical trials carried out on the Cypher stent (SIRIUS-1) showed restenosis rates of 8.9% after 8 months compared to 36.6% for BMS in the same study. Likewise the TAXUS IV trials heralded a dramatic reduction in restenosis rates when compared to BMS after 9 months, 7.9% versus 26.6% respectively (Venkatraman and Boey, 2007).

A successful DES procedure is defined by its ability to transport the right amount of an appropriate drug within the correct timeframe that will ultimately deem the operation a success through the prevention of ISR. There are aspects of the performance of a DES that can be controlled by the manufacturer which has led to considerable reductions in the instances of ISR, such as the type of drug used and the characteristics of the coating that the drug is stored in. Even the stent shape, thickness and width of the struts can all influence the manner in which the drug is transported. However, the ability for drug uptake within the artery wall is also governed by its interaction with the patient specific arterial environment, making it near impossible to completely eradicate ISR. The degree of initial stenosis, the presence of luminal and abluminal thrombus on the stent and even the advent of re-endothelialisation will all contribute to the DES ability to transport drugs throughout the artery wall.

3. Mass transport theory of drug eluting stents

Mass transport refers to the movement of mass, i.e. the species of interest which is drugs in the case of a DES, within a defined system. This transport of species may be provoked by concentration gradients between two points, but quite often in systems, especially in the vasculature, overpowering complex flow dynamics will ultimately be responsible for the mass transport outcome. In the absence of a free flowing system the presence of these concentration gradients induces diffusion, e.g. between the DES and the artery wall. Mass transport can be broken up into two types within the human vasculature. Firstly blood side mass transport (BSMT) refers to species transport within the vessel lumen and is subject to the haemodynamics therein. Often evanescent due to haemodynamic washout, BSMT can only be effective in transporting anti-proliferative agents to the wall in regions of high recirculation.

The second, and most important, mode of mass transport is in relation to transport within the wall of the artery, referred to as wall side mass transport (WSMT). Along with the properties of the species being transported within the artery wall, WSMT depends on the structural condition of the wall itself, whereby a damaged intimal layer could facilitate accelerated mass transport through to the medial layer. WSMT can be governed by two transport forces, a pressure driven convective force and a diffusive force. The Peclet number (Pe), see equation 11, is a dimensionless parameter that can be used to determine the relative influences of these two forces. A small Pe (i.e. $Pe<1$) is representative of transport which is dominated by diffusion, while a higher Pe (i.e. $Pe>1$) indicates convection dominated mass transport (Friedman, 2008).

3.1 Governing equations

Computational Fluid Dynamics has emerged as one of the most powerful numerical tools for engineers, scientists and mathematicians alike. Its foundations are based on theoretical analysis drawn from experimental observations over various branches of physics. The starting point for any computational analysis is the appropriate allocation of the governing equations. These equations are then substituted with equivalent numerical descriptions that are then solved using appropriate mathematical techniques. There are a number of numerical techniques available that will return a solution to a specified problem. Two of the more popular methods are the Finite Volume Method and the Finite Element Method. The assumptions generally applied when modelling fluid flow problems of this nature are as follows:

- The flow is incompressible and isothermal
- The fluid is Newtonian and possesses constant physical properties
- Flow is considered to be laminar

3.1.1 Conservation of mass: Continuity equation

The conservation of mass is a form of continuity equation which states the net mass flow into a control volume is equal to the rate at which mass leaves the control volume. That is providing there are no sinks or sources within the control volume. The differential form of the equation can be obtained by simply considering the flow into and out of elementary control volume. For the Cartesian co-ordinate system, having coordinates x, y, z referenced to a stationary frame with the corresponding velocity components u, v, w (m/s), the continuity equation can be written as:

$$\frac{\partial \rho}{\partial t}+\frac{\partial(\rho u)}{\partial x}+\frac{\partial(\rho v)}{\partial y}+\frac{\partial(\rho w)}{\partial z}=0 \tag{1}$$

Where the density (ρ, kg/m^3) is a constant, as is the case of incompressible flow, this reduces further to a volume continuity equation.

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \tag{2}$$

3.1.2 Balance of momentum: Navier-Stokes Equations

The balance of momentum is derived from Newton's second law of motion, which states that the rate of change of momentum of a fluid particle is equal to the sum of the forces on the particle. The Navier-Stokes Equations describe the full three dimensional, viscous nature of fluid motion in a control system:

$$\begin{aligned}
\rho\frac{\partial u}{\partial t}&=-\frac{\partial P}{\partial x}+\mu\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\right)\\
\rho\frac{\partial v}{\partial t}&=-\frac{\partial P}{\partial y}+\mu\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial z^2}\right)\\
\rho\frac{\partial w}{\partial t}&=-\frac{\partial P}{\partial z}+\mu\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}+\frac{\partial^2 w}{\partial z^2}\right)
\end{aligned} \tag{3}$$

Where μ(Pa s) is dynamic viscosity and P(Pa) is pressure. Due to the porous nature of the artery wall, flow within it must consider the influence of the tissues permeability. Therefore the flow within the wall is assumed to follow Darcy's law and is demonstrated (in the x direction) by the inclusion of the permeability term in equation 4, where K(m^2) is the permeability of the arterial tissue.

$$\rho\frac{\partial u}{\partial t}=-\frac{\partial P}{\partial x}+\mu\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\right)-\left(\frac{\mu}{K}\right)u \tag{4}$$

3.1.3 Fick's laws of diffusion

Species transport via diffusion is a process driven by concentration gradients between two locations. Fick's first law can be used to describe the diffusional flux (J_x, mol/m^{2}s) of such species, shown in 1D in equation 5, where D(m^2/s) is diffusivity and c is concentration (mol/m^3):

$$J_x = -D\frac{\partial c}{\partial x} \tag{5}$$

The negative term in equation 5 indicates that the flux is positive in the presence of a negative concentration gradient. Biological mass transport often requires the application of a time-dependent mass transport process that can predict variations in concentration over time. Fick's second law (equation 6) can provide such a relationship and is defined here in one dimension:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} \tag{6}$$

3.1.4 Conservation of species: Convection-diffusion equation

The addition of a convective term, equal to the product of the fluid velocity and the local concentration, to equation 6 demonstrates the 3D transport of species in a flowing solution. This is known as the convection-diffusion equation.

$$\frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} + v\frac{\partial c}{\partial y} + w\frac{\partial c}{\partial z} = D\left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} + \frac{\partial^2 c}{\partial w^2}\right) \tag{7}$$

3.1.5 Ratio of convective to diffusive forces

The Peclet number (Pe) is a dimensionless number that determines the relative contribution of convective and diffusive forces to species transport within a defined system. It can be defined as a product of the Reynolds number (Re) and the Schmidt number (Sc).

$$Pe = Re.Sc \tag{8}$$

The Reynolds number is a non-dimensional parameter concerning fluid forces due to viscosity and inertia and is essentially used to determine whether a flow is laminar, transitional or turbulent in nature. For example a Reynolds number of approximately 90 can be obtained for a mean arterial velocity (u) of 0.1m/s in an artery with a diameter (a) of 3mm. When considering transmural flow through the porous artery wall the value $a$would represent the thickness of the porous wall.

$$Re = \frac{\rho u a}{\mu} \tag{9}$$

The Schmidt number (Sc) is defined as the ratio of kinematic viscosity (v, m^2/s) to diffusivity (D).

$$Sc = \frac{v}{D} = \frac{\mu}{\rho D} \tag{10}$$

Substituting equations 9 and 10 into 8 describes how convective and diffusive forces can influence the outcome of the Peclet number.

$$Pe = \frac{ua}{D} \tag{11}$$

3.1.6 Diffusion in porous materials

When considering diffusion in a fluid saturated porous media, as is the case with the artery wall, diffusion takes place over a tortuous path. Because these pores are not straight, the distance over which diffusion takes place becomes effectively longer than for a homogenous material of the same thickness. The effective diffusivity (D_{eff}) can therefore be deduced by considering the impact of the materials structure on the species free diffusivity (D_{free}). The effective diffusivity of a porous material is a function of its porosity (ε) and tortuosity (τ).

$$D_{eff} = \frac{\varepsilon}{\tau} D_{free} \tag{12}$$

One of the more common ways to determine the free diffusivity of a species in a solvent is to use the Stokes-Einstein equation (13), where k(J/K) is the Boltzmann constant, T(K) is temperature and R(m) is the radius of the solute. For the purpose of diffusion in the artery wall, this solvent is considered to be plasma.

$$D_{free} = \frac{kT}{6\pi\mu R} \tag{13}$$

The radius of the solute can be calculated from equation 14 assuming that the particle is spherical in shape, where M(Kg/mol) is the solute molecular weight and Na(mol^{-1}) is Avogadro's number.

$$R = \left(\frac{3M}{4\pi\rho Na} \right)^{1/3} \tag{14}$$

The structure of the porous medium is defined by the tortuosity(τ) of its porous network (15) and by the porosity(ε) (16) of the material itself.

$$\tau = \frac{L}{X} \tag{15}$$

Where L = pore path length and X = distance between beginning and end of the pore path.

$$\varepsilon = \frac{Pore\,Volume}{Total\,Volume} \tag{16}$$

4. Problem classification

In reality the classification of problems of this nature are inherently patient specific and as such no one representation of the problem is correct. However, there are innate similarities between patients. Blood flow within the vasculature is a highly complex 3D process to model given the pulsatile nature of arterial haemodynamics. Coupled with this pulsatile process, the coronary arteries are situated on the surface of the heart and as such are subject to cyclic motion due to the beating of the organ. Therefore the modelling of drug transport from a DES in these arteries is multifaceted in nature, comprising of both luminal and artery wall mass transport, the latter of which may also be subject to a reaction giving that some drug may bind to the arterial tissue. The introduction of a multi-layered artery wall to the model increases the complexity of the domain even further. So to what extent does one go to when modelling DES mass transport?

Comprehensively modelling the behaviour of a DES computationally over a given time period would require the application of the following *in vivo* conditions experience by the device.

- The DES would have to be placed in multi-layered diseased artery.
- Both BSMT and WSMT would have to be considered.
- The real time occurrence of thrombus formation and re-endothelialisation under pulsating flow conditions would need to be modelled.
- The structural deformation of the artery wall due to DES deployment would need to be taken into account.
- Also the movement of the vessel in space due to its location on the surface of the beating heart and how this may alter depending on the extent of the patient's physical activeness would have to be considered.

In light of the computational requirements to undertake such a model it is possible, and almost necessary, to make certain assumptions in order to simplify both BSMT and WSMT models whilst retaining enough detail of the actual model to draw relevant conclusions from the analysis.

The implementation of an arterial pulse and a beating heart are neglected by most researchers. Often the artery is modelled as rigid in space in order to analyse mass transport post DES deployment. This is an effective assumption but one must consider the deformation of the artery wall due to the dynamic expansion of the stent, as this can have an impact on the mass transport outcome due to the porous nature of the wall and the compression it incurs upon stent expansion. As for the application of laminar blood flow, it can be seen that the majority of drug that enters the artery wall from the DES does so via physical contact with the wall and the drugs emanating from the areas of the stent exposed to flow, be it laminar or pulsatile, are predominantly carried downstream.

4.1 Artery wall classification

Arteries transport oxygen rich blood around the body providing essential nutrients to vital organs. The artery wall consists of a complex multilayer porous substructure with an interstitial phase comprising predominantly of plasma. In a healthy artery this substructure (Figure 1) is comprised of three concentric layers; the tunica intima, the tunica media and the tunica adventitia. The tunica intima is the innermost layer, consisting of a single layer of endothelial cells and a subendothelial layer mainly consisting of delicate connective tissues and collagen fibres. The outer boundary of the tunica intima is surrounded by an elastic tissue with fenestral pores known as the internal elastic lamina (IEL). The medial layer consists primarily of concentric sheets of smooth muscle cells (SMC) within a loose connective tissue framework. This configuration of SMC enables the artery wall to contract and relax. The tunica media and the tunica adventitia are separated by another thin band of elastic fibres known as the external elastic lamina (EEL). The outermost layer of the artery, the tunica adventitia, is comprised of connective tissue fibres and some capillaries. These fibres blend into the surrounding connective tissues and aid in stabilising the arteries within the body (Khakpour and Vafai, 2007).

The target layer for the anti-restenotic drugs is the tunica media, where the SMC reside, and quite often computational studies will consider just this arterial layer not only because of this fact but also due to the possible erosion of the tunica intima upon stent deployment. Regardless of the level of complexity modelled, the artery wall is porous in composition and drug transport is facilitated through the surrounding plasma not only via diffusion but there is also the presence of a transmural velocity due to a pressure gradient observed across the artery wall. However, the presence of arterial plaque will reduce the magnitude of this transmural velocity and can even stem it altogether. As DES are deployed in highly occluded arteries it is reasonable to reduce the complexity of the problem by neglecting convection in the wall. Equation 12 gives us an indication of how arterial properties such as porosity, tortuosity and free diffusivity can influence the transport of drugs within the respective artery wall layers. The compression of these layers will alter these properties which in turn may inhibit the transport of species as governed by the mass transport equations. The compression of a porous structure not only reduces the materials porosity but it results in the creation of a more arduous pore path over which mass transport would normally occur. The combination of a reduced porosity with an increased tortuosity, when the artery wall has been compressed, has a net effect of reducing the effective diffusivity thus hindering mass transport within the vessel.

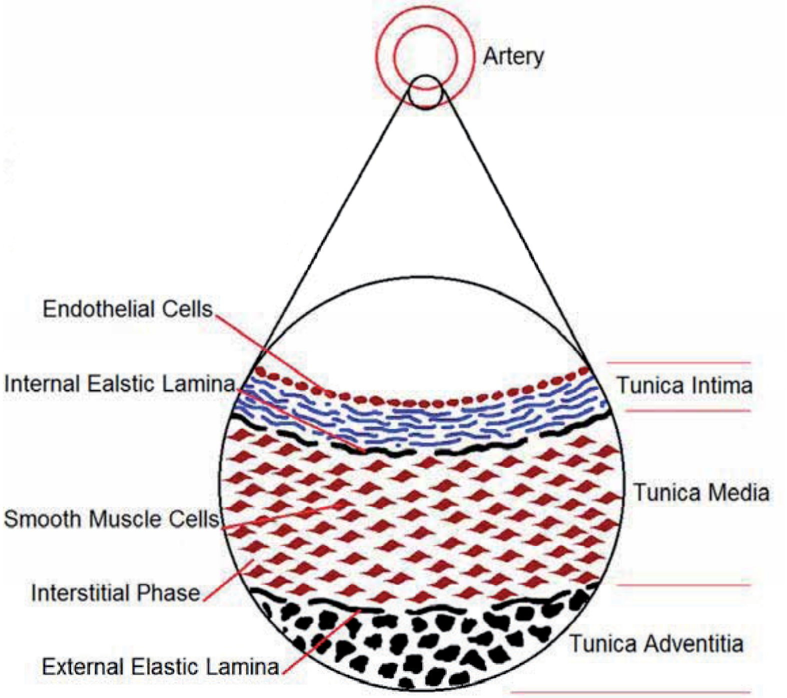


Fig. 1. Illustration of the cross-sectional structure of a healthy artery wall.

4.2 Influence of thrombus

Hwang et al. (2005) were among the first to explore the influence of thrombus height, width and type on the arterial drug uptake. Stents can be deployed at sites of thrombus and after implantation a clot will inevitably develop once the struts become covered with plasma proteins. In most cases this will not be angiographically present or clinically evident but even a fine layer of clotting blood deposited on the surface of the DES can alter drug distribution within the wall. The presence of clot alters the local environment of the stent strut and the physiological transport forces that regulate arterial uptake and retention. Balakrishnan et al.(2008) reported that drug eluting stents clot at a rate of 0.6% each year after implantation for up to 3 years. Strut position within a clot also has a major influence on the arterial uptake. The greater the volume of clot covering the strut, and the closer the strut is to the wall, creates improved conditions for greater drug delivery. Hwang et al. (2005) discovered that in this configuration concentration distribution in the wall can be 30 fold higher than situations where no clot is present. Similarly thrombus or plaque between the strut and the artery wall act as a buffer layer and reduce wall concentrations. Clot diffusivities higher than that of the artery wall will result in drug transfer to the blood at a rate faster than can be absorbed by the wall. Clots with diffusivities equal to or lower than the artery wall can transport drugs to the wall at a rate where the wall can effectively absorb the drugs, thus reducing drug loss to the bloodstream.

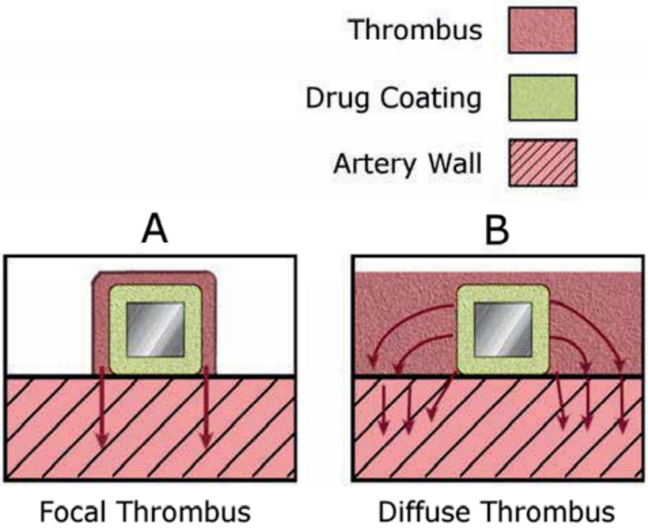


Fig. 2. An illustration of how Focal (A) and Diffuse (B) thrombus, surrounding a DES strut, contribute to drug mass transport within the artery wall.

In 2008, Balakrishnan et al. demonstrated how variations in thrombus size and distribution contribute to fluctuating arterial drug intensities (Figure 2). Their simulations indicate that thrombus cannot influence the slow rate of drug release from the stent because the polymer resistance to drug transport is significantly greater than that of the thrombus, which is consistent with *in vivo* experimental drug release. Focal thrombus, with a thickness of 0.1mm, increased peak average drug concentration by 80%. Greater clot formation between stent struts will also have an effect on arterial concentrations because once the species is transported within the clot it effectively increases the surface area from which the artery wall can absorb drugs. The formation of this interstrut thrombus, described as diffuse thrombus, acts as a shield from drug washout and culminates in an increase in arterial concentration by up to 3.5-fold.

The variability of thrombus can have a major impact on arterial drug concentrations. It can aid in drug uptake and retention within the wall when it covers the stent, but too much thrombus will effectively block the artery, thus creating a problem that the DES aimed to alleviate. Also it can act as a barrier in preventing drugs from reaching the wall when it is located between the strut and the wall. The likely scenario following DES implantation within the vasculature is that at some location along the length of the stent each of these situations will be present. In addition to issues arising from the development of thrombus, the presence of plaque in a freshly stented artery will have a similar effect on mass transport but does so immediately upon implantation, as opposed to thrombus which develops with time. The initial presence of the plaque coupled with the time dependent formation of thrombus, and even the occurrence of re-endothelialisation, would create a realistic stenting scenario if modelled but the volume and formation of these will be different for each person. So to what extent does one model this?

4.3 Polymer and drug characterisation

Stent coatings play a vital role in the regulation of drug release from the stent. Careful consideration must be observed when allocating a polymer as a drug carrier for a DES. If the polymer is not biocompatible an inflammatory response ensues. It has been shown that the development of neointimal hyperplasia can be doubled when certain polymer configurations are used compared with a controlled substance (Granada et al., 2003). Each biologically viable polymer must be able to endure the stresses exerted when stents are deployed, resist cracking, peeling and maintain their physiochemical properties after sterilisation. Drug release rates can be altered with the addition of an extra layer of polymer coating to modulate between slow and fast release formulations. This adds an extra degree of complexity when designing functional DES and indeed analysing them computationally.

Sirolimus, everolimus and paclitaxel were among the first anti-restenotic agents used in DES. Sirolimus, or rapamycin, is a naturally occurring macrocyclic lactone. It was approved by the FDA in 1999 for the prophylaxis of renal transplant rejection because of its potent immunosuppressive properties (Costa & Simon, 2005). Essentially sirolimus inhibits the activation of multiple kinases, associated with cell proliferation, resulting in the cessation of cellular division between the G1/S phases and is said to be cytostatic in nature (Burt & Hunter, 2006). Everolimus is an analogue of the sirolimus immunosuppressant which binds to cytosolic immunophillin FKBP12. It is very similar to sirolimus in that it prevents the cellular division between the G1/S phases of the cell cycle therefore inhibiting SMC proliferation (Stone et al., 2008).

Paclitaxel is a naturally occurring drug that was originally extracted from the pacific yew tree Taxusbrevifolia. It was initially used to treat several types of cancer such as breast and ovarian cancer (Tanabe et al., 2004). Like sirolimus and everolimus, paclitaxel is effective in reducing restenosis but it does so in a manner that results in cell death, suggesting that it works through a cytotoxic mechanism (Parry et al., 2005). Paclitaxel achieves its anti-proliferative effects by binding to aminoterminus β-tubulin thus disrupting microtubule dynamics. This results in the arrest of cells at the M stage and even G2 stages of the cell cycle, leading to cell death (Hara et al., 2006). Due to its lipophilic properties paclitaxel has been known to be loaded directly onto stents without a polymer coating but there could be potential implications with regards to a lack of controlled release (Burt & Hunter, 2006).

4.4 DES problem design and deployment

This problem can be approached in one of two ways, do you a) want to compare the mass transport capabilities between different stent designs or b) want to analyse arterial mass transport for a single stent design under various stenting and artery wall conditions. The latter of which is of more initial interest because a comprehensive understanding of a multifaceted mass transport study with a generic stent design will give a greater understanding of the interactions between drug/polymer characteristics and the arterial condition. Once these interactions are better understood the researcher can revert to comparing stent designs for a predetermined deployment configuration.

The main goal of a DES is to prevent the onslaught of arterial restenosis, which occurs in part due to damage inflicted on the artery during stent deployment. However, researchers to date have generally neglected the artery wall damage induced and its influence on mass transport. An exploratory DES mass transport computational study, even if it is only 2D modelling, should consider both cause and effect. A stent should not just be placed flush with the artery wall, there is going to be some wall indentation and intimal damage, the extent of which is a study onto itself. To this end the resulting artery wall compression will alter the effect that is the transport of the anti-restenotic drugs throughout the artery wall. Stent design and drug/polymer properties, although of significant importance, should be a secondary consideration until these fundamental issues have been addressed.

5. Typical computational boundary conditions for DES models

Once the governing equations have been applied to the model domain the boundary conditions need to be allocated. Often with biological modelling it is necessary to make assumptions when applying boundary conditions. For example treating the artery wall as rigid (Mongrain et al., 2007; Devereux, 2005; Kaazempur-Mofrad and Ethier, 2001) or assuming that mass transport within the wall is modulated solely by diffusion (Balakrishnan et al., 2008, 2007, 2005; Mongrain et al., 2007, 2005) are two examples of ways commonly employed to simplify what is in reality a very complex problem. However, as previously mentioned the fundamental *in vivo* issues should be taken into account as much as possible when applying such simplifications.

5.1 Application of momentum boundary conditions

5.1.1 Inlet: Velocity

The heart is a muscular organ that undergoes repetitive contraction and relaxation of its walls in order to propel blood through the circulatory system. Coupled with the complex geometry of the coronary arteries, the pulsating blood velocity profile is an integral part in the mass transport behaviour of blood borne species. However, when considering the transport of drugs from a DES a common assumption to make is the presence of a steady fully developed flow profile within the lumen (Balakrishnan et al., 2008, 2007, 2005; Mongrain et al., 2007, 2005). When modelling BSMT it may be necessary to incorporate the time-dependent pulsatile nature of blood flow. However, the transient nature of blood side drug transport enables the assumption of a steady flow inlet boundary condition that in most cases will not have a considerable impact on WSMT. Arterial flow can be replicated by applying a pulsatile parabolic velocity profile at the vessels inlet.

5.1.2 Outlet: Pressure

A pressure of zero can be applied to the outlet of DES computational models which reduces the likelihood of encountering backflow through the outlet. The application of such a boundary condition is prevalent in mass transport studies of this nature (Kolachalama et al., 2009; Balakrishnan et al., 2008, 2007, 2005; Rajamohan et al., 2006).

5.1.3 Lumen-wall interface: No-slip

The no-slip boundary condition is standard for a model with a stationary wall and states that the velocity of the fluid is zero relative to the boundary (O'Brien et al., 2005; Walsh et al., 2003).

5.1.4 Axial-symmetry

Due to the idealised nature of some arterial computational models the use of an axial-symmetry boundary condition can greatly reduce the number of degrees of freedom that need to be solved for. This will reduce the computational demand without having to sacrifice the accuracy of the analysis. The Reynolds number for blood flow in DES computational models is normally small (Re<100), indicating laminar flow and as such the application of the axial-symmetry boundary condition should not influence the fluid flow solution.

5.2 Application of mass transport boundary conditions

5.2.1 Luminal inlet: Concentration

Because mass transport within the artery's lumen is convection dominated, it would be virtually impossible for drugs to diffuse in the direction opposing blood flow. Therefore it can be assumed that the luminal inlet concentration has a constant value equal to zero. This boundary condition has been used extensively in mass transport studies on DES (Balakrishnan et al., 2008, 2007, 2005; Mongrain et al., 2007, 2005).

5.2.2 Luminal outlet: Convective flux

A convective flux outflow condition is generally imposed on the outlet of a models lumen, resulting in a zero concentration gradient at the outlet. This boundary condition assumes that all the mass passing through the boundary is convection dominated.

5.2.3 Lumen-stent-wall interface: Continuity

A common assumption for DES mass transport studies is that the intimal layers of the artery are denuded and that the stent is in direct contact with the medial layer of the artery wall (Kolachalama et al., 2009; Balakrishnan et al., 2008, 2007, 2005; Mongrain et al., 2007, 2005). This negates the need to model the endothelial, intima and internal elastic lamina layers. Regardless of the inclusion or exclusion of these layers the continuity equation should be the default setting for all interior boundaries. This condition states that in the absence of sources or sinks, the flux in the normal direction is continuous across the boundary, i.e. the concentration is equal on both sides of the boundary.

5.2.4 External faces: Insulation/symmetry

This condition is specified at the perivascular wall and at the up- and down-stream wall boundaries, which should be a sufficient distance away from the stent. It specifies where the domain is well insulated or it can reduce the size of a model by taking advantage of symmetry. Intuitively this condition states that the gradient across the boundary must be zero, therefore it is impermeable to mass transport.

6. Experimental validation of mass transport behaviour

Historically, experimental validations have been necessary to prove researchers hypotheses across all paradigms of science. In order to validate a theory one must not only be mindful of their goal but also their ability to achieve it. For instance, trying to validate a computational DES model using excised arterial DES would be nice in theory but in practice may prove fruitless because it would be near impossible to obtain the site specific drug concentrations within the artery wall, necessary for the researcher to examine the nuances of stent design that they are interested in analysing. It is often useful to validate a single aspect of the model if possible. With regards to DES, a solitary WSMT model may be acceptable,as a vast body of knowledge pertaining to fluid flow problems already exists, unlike the as of yet mature understanding of WSMT and how the behaviour of the porous artery wall and other pertinent features influence the mass transport therein.

Figure 3 illustrates an example of a validation flowchart for a study of mass transport from DES. The first process in designing a validation experiment is to analyse the problem as a whole and see what you would like to prove. Then certain aspects of the problem that are relatively rudimentary can be neglected from the validation, providing they won't have a fundamental impact on the outcome of the test. The flowchart is divided up into two streams that are developed jointly in order to achieve a desired validation. The experimental mass transport validation is on the left and the development of the corresponding computational mass transport model is on the right. In both streams BSMT is not highlighted for inclusion in this example's validation procedure. This may be due to the readily available examples of validations of this nature in literature or its minimal impact on the outcome of the results. This study was designed to investigate instances of WSMT in relation to DES design and deployment.

How accurate do you want your computational model to be and what assumptions are going to be made? The flowchart (Figure 3) describing this example validation decided to neglect convection-diffusion mass transport in the artery wall and concentrate on validating a diffusion only model. The reason for such a decision could be threefold; 1) an artery requiring a DES would be highly calcified and therefore the plaque can act as a buffer to stop or considerably reduce the flow in the porous wall, 2) an analysis of the Peclet number (equation 8) demonstrates diffusion dominated transport for a give drug or 3) the inability to create and obtain tangible results from an applicable experimental model with both convection and diffusion. The next aspect of the computational model is the characterisation of the artery wall, i.e. should it be modelled as a porous medium. Although flow in the wall has been neglected the application of a porous wall still can have a bearing on the outcome of mass transport due to the characterisation of the effective diffusivity and its propensity to change under varying stenting conditions (equation 12).

In this example a computational model has been identified in which many of the aspects of the *in vivo* stenting conditions remain but more importantly contains the ability to develop and analyse a corresponding experimental validation. Validations of this nature are a powerful tool in a researchers arsenal because once the initial hypotheses has been validated the computational model can be changed to any geometry imaginable to create more realistic stenting scenarios, and when solving the problem using the same physics one can have full confidence in the results.

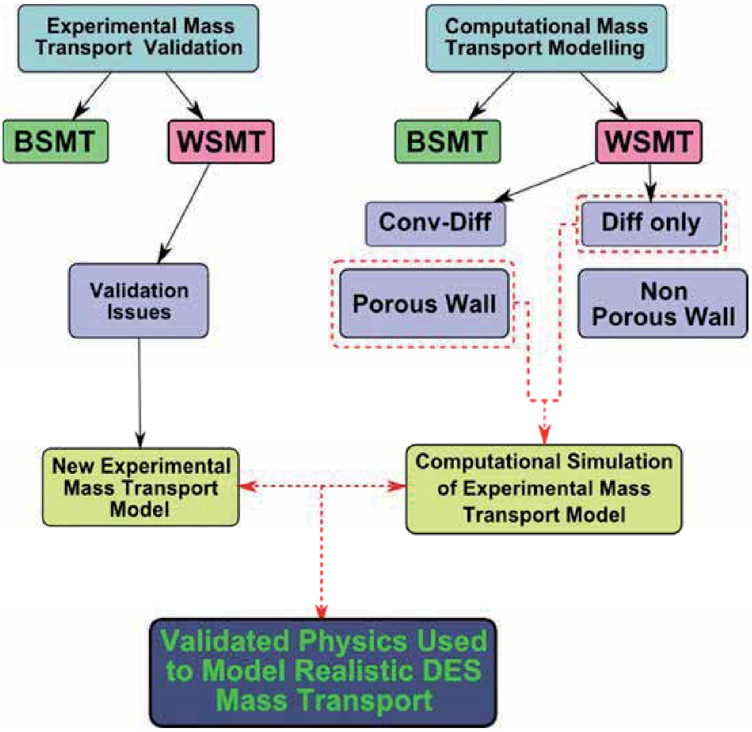

Experimental Mass Transport Validation
Computational Mass Transport Modelling
BSMT
WSMT
BSMT
WSMT
Conv-Diff
Diff only
Validation Issues
Porous Wall
Non Porous Wall
New Experimental Mass Transport Model
Computational Simulation of Experimental Mass Transport Model
Validated Physics Used to Model Realistic DES Mass Transport

Fig. 3. Flowchart demonstrating the process of identifying an appropriate route for the experimental validation of DES mass transport.

6.1 Historical experimental mass transport validations

Experimental techniques to simulate mass transport in the vasculature have been an integral part in the development of DES. Various techniques can be applied depending on the required outcome from the analysis. For example if lumen transport is to be analysed a common approach would be to introduce a dye to a flowing system. Concentrations are then measured at certain time points and locations by withdrawing fluid samples so that the optical dye intensity can be determined.

Markou et al. (1998) employed this approach when analysing the local transport of anti-restenotic agents from a novel drug delivery device. The device consisted of a porous membrane that lined the artery wall and was the location where drugs would be infused into the artery. Their experimental approach consisted of a simulated artery section with a 1cm slit around its circumference. Dye was then infused from this slit in a radial direction into the lumen at uniform speed. It was found that the radial diffusion of the species was minimal in comparison to the axial convection therefore the majority of species remain in close proximity to the wall. This experiment was validated with the commercially available finite volume solver Fluent (Lebanon, NH). The computational models predicted an increase in dye concentration at the wall with an increase in the infusion rate. The same effect was witnessed experimentally and there was good agreement between the results. This study indicates the dominance of convective transport over diffusive transport in the arteries lumen.

Lutostansky et al. (2003) adopted a similar technique when conducting an experimental analysis of mass transport in the recirculation region downstream of a sudden expansion. Dye concentrations were determined at four equidistant locations after the sudden expansion. Experimental concentrations were taken over the course of one hour and showed very good correlation with predicted results from two numerical codes, the finite volume code Fluent (Lebanon, NH) and a finite element code FTSP (developed by Graz University of Technology). Although these experiments provided a validation of their numerical approach, they are limited in that they cannot be used to analyse mass transport within a porous wall.

In 2001, Hwang et al. found that stent based delivery, from a Palmaz-Schatz Crown stent (Cordis), resulted in large variations in concentration gradients. Drug concentrations were found to vary from zero to several times the mean over a few micrometres. The aforementioned stents were spray coated with a fluorescein sodium/ethylene vinyl acetate copolymer and deployed in excised bovine carotid arteries. The arteries were positioned in an *ex vivo* perfusion apparatus and immersed in a perivascular bath where coronary flow was simulated. After 3 hours the arteries were removed and cut into slices. The fluorescein concentration was then measured with a spectrofluorometer. Although the experiment was not used to validate a computational model, the effects witnessed experimentally were compared to variations in simulated drug physicochemical properties. They concluded that the proximity of the device does not necessarily ensure adequate targeting because transport forces can cause local concentrations to deviate from the mean concentration.

In a later paper Hwang et al. (2005) evaluated the paclitaxel uptake in stented abdominal aortas of adult male Sprague-Dawley rats in the presence and absence of controlled mural thrombus. The in vivo clot dimensions were determined and used as boundary conditions and input parameters for the computational model. The computational analysis predicted an arterial drug ratio of 0.56 which correlates with the 50% decrease in arterial uptake ascertained from the animal experiments. Hwang et al. (2005) discovered that by varying clot size and location, large variations in arterial uptake were witnessed.

In 2007, Balakrishnan et al. deployed a Cypher sirolimus eluting stents in porcine arteries. At the desired time points of 1, 8, 14, 30, 60 and 90 days after implantation the stents were harvested and analysed. In each case the stents were carefully removed from the artery and the remaining drug within the polymer was determined. When subtracted from the amount of drug prior to implantation the release fraction can be calculated. At each time point this fractional drug release was compared to numerical predictions using Fluent (Lebanon, NH). A good correlation validated the Fickian diffusion analysis applied with the numerical solver to approximate the drug transport from the coating. However, validating drug release from the polymer coating does not elucidate subsequent drug uptake within the porous artery wall.

In 2010, O'Connell and Walsh developed an analogous model of artery wall mass transport, examining the hypothesis of how compression of a porous media alters mass transport within. Due to the difficulty in measuring site specific concentrations within the artery wall they developed a scaled up experiment. It consisted of a bed of pH paper that was saturated in a neutral pH solution in order to fill the pore space of the material, similar to that of the artery walls interstitial fluid. The wall is then compressed, in increments up to a maximum of 23.75% of its initial thickness, and then the species of interest, an acid of pH 2.0, is introduced to the system and the resulting colour change was used as a marker for concentration. This enabled the site specific measurement of concentration at different depths throughout the porous wall. These experimental results were then validated computationally using the finite element solver COMSOL Multiphysics. The authors concluded that compression of a porous artery wall contributes significantly to the modulation of arterial WSMT and should be considered in future DES computational studies.

7. Computational modelling of mass transport from drug eluting stents

The following computational models were created to illustrate how changes in stenting deployment conditions can influence drug concentrations within the artery wall, analysed after 30 and 60 minutes for each model. Figure 4 describes the five 2-D axis-symmetric computational models that were analysed. Model 1 in figure 4 depicts the locations of drug concentration measurement through the depth of the artery wall and axially down the artery at a depth of 25% and 50% of the wall thickness (WT) beneath the strut. The models are described as follows:

Model 1. In the absence of a lumen and subsequent BSMT, this model only considered WSMT from a single DES stent strut (150μm) deployed flush against a single layer artery wall (200μm). In this instance WSMT is purely diffusive.

Model 2. Similar to Model 1 but with the inclusion of a steady blood flow profile (mean velocity = 0.1m/s) through the arterial lumen. Here WSMT is purely diffusive and BSMT is modelled using the convection-diffusion equation.

Model 3. Similar to Model 2 but with the inclusion of a 20μm thick layer of plaque along the artery wall.

Model 4. Similar to Model 2 only the stent strut becomes embedded in the artery wall as it compresses it by 25% of the wall thickness.

Model 5. Similar to Model 4 except upon compression of the wall the stent strut doesn't become embedded.

A hypothetical drug was used in the analysis with effective diffusivity values in each respective media defined in table 1. It is the combination of both the drug used and the characteristics of the media within which transport takes place that determines the effective diffusivity value. This fact becomes evident as the effective diffusivity of the drug in the compressed wall is determined. The drug remains the same but, as described in equation 12, changes to the tortuosity (τ) and porosity (ε) of the wall alters the effective diffusivity within.

From equation 15 the pore path (L) remains the same length but the distance (X) has reduced due to the 25% compression of the artery wall. Therefore the tortuosity of the compressed wall (τ_{CW}) can be described as a function of the tortuosity of the wall in its original state (τ_W).

$$\tau_{CW} = \frac{L}{0.75X} = 1.333\frac{L}{X} = 1.333\tau_W$$

Similarly, as the wall is compressed the total volume and pore volume of the wall under compression reduces but the fibre volume remains the same. To this end ε_{CW}= 0.787ε_W and the effective diffusivity of the compressed wall (D_{CW}) can be described as follows:

$$D_{CW} = \frac{\varepsilon_{CW}}{\tau_{CW}} D_{free} = \frac{0.787\varepsilon_w}{1.333\tau_W} D_{free} = 0.59 D_W = 0.59 \times 10^{-12} m^2 / s$$

Diffusivity in Lumen	$D_L = 1\times10^{-10}$ m^2/s
Diffusivity in Stent Coating	$D_S = 1\times10^{-14}$ m^2/s
Diffusivity in Plaque	$D_P = 1\times10^{-13}$ m^2/s
Diffusivity in Wall	$D_W = 1\times10^{-12}$ m^2/s
Diffusivity in Compressed Wall	$D_{CW} = 0.59\times10^{-12}$ m^2/s

Table 1. Effective diffusivity values of the different layers of arterial DES models.

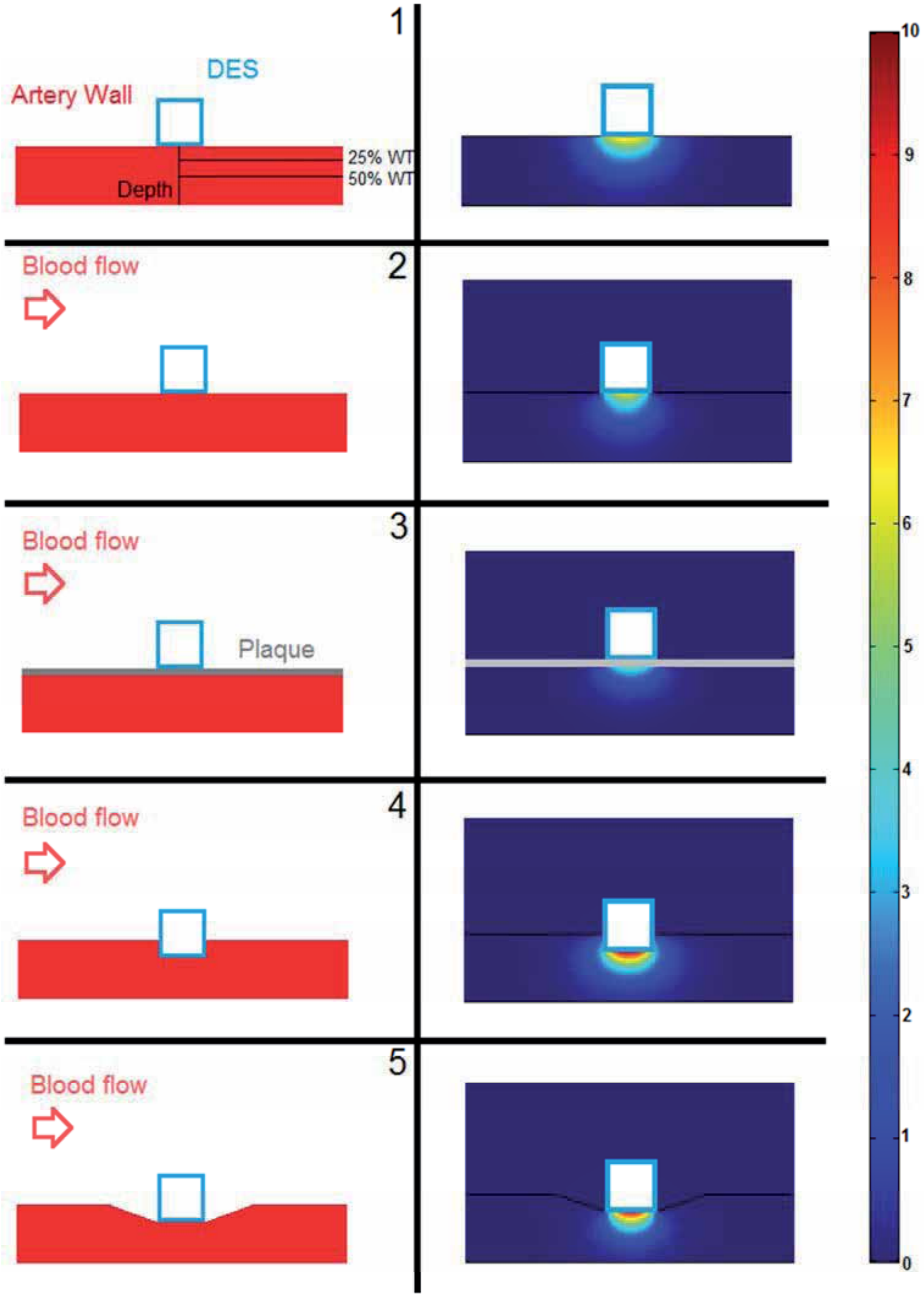

1
DES
Artery Wall
25% WT
50% WT
Depth
2
Blood flow
3
Blood flow
Plaque
4
Blood flow
5
Blood flow
10
9
8
7
6
5
4
3
2
1
0

Fig. 4. Illustration of 2-D axis-symmetric computational models used in the DES mass transport analysis and their resulting drug concentration contours after 30 minutes.

7.1 Computational predictions of mass transport in the artery wall

Even though BSMT doesn't significantly contribute to the distal absorption of drugs into the artery wall, it is a necessary element of the modelling process as it provides a more realistic representation of how the drugs in the DES dissipate in the vasculature. Failure to model BSMT (Model 1) culminates in the entire reservoir of antirestenotic drug, from both the luminal and abluminal side of the stent, having no choice but to eventually transport into the wall. This as we know is not the case as a considerable amount of drug is lost to the blood stream. Figures 5-7 demonstrate that for both time points and at each location the drug concentration for Model 1 is greater than that of Model 2 due to the absence of BSMT. Model 3 examined the influence of arterial plaque, which was given a drug diffusivity of $1\times10^{-13}m^2/s$. In reality plaque size and composition will vary from patient to patient and the implication of this is a study in itself where a range of plaque types would need to be modelled in order to quantitatively predict its influence on WSMT. For the purpose of demonstrating the influence that the presence of any plaque may have on mass transport, a drug diffusivity was chosen that is an order of magnitude between the diffusivities in the DES coating and the uncompressed artery wall respectively. What Model 3 demonstrates is that even a 20μm thick layer of plaque between the stent and the artery wall can significantly reduce uptake within the artery wall, even more so than arterial compression.

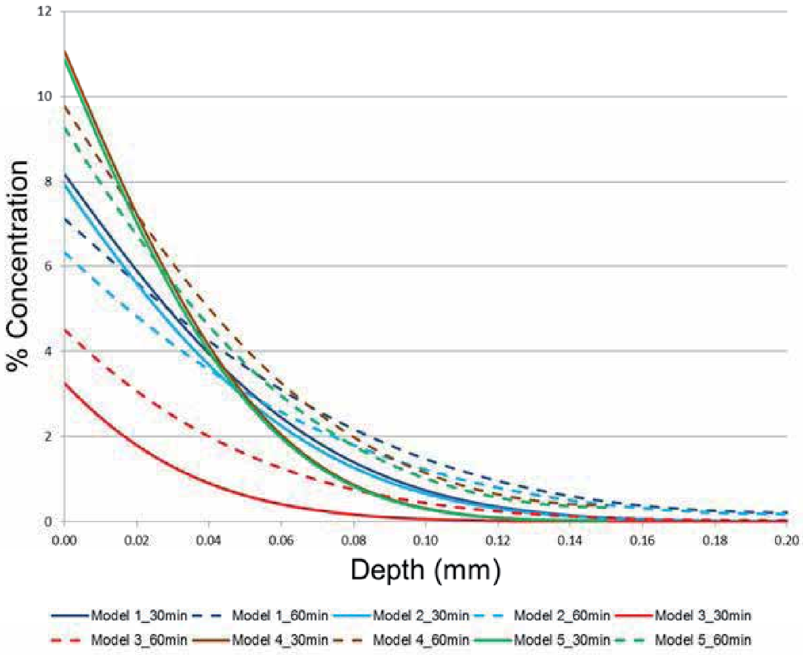


Fig. 5. Drug concentrations through the depth of the artery wall, as illustrated in Fig 4 Model 1. Concentrations are measured after 30 minutes (solid line) and 60 minutes (dashed line) respectively.

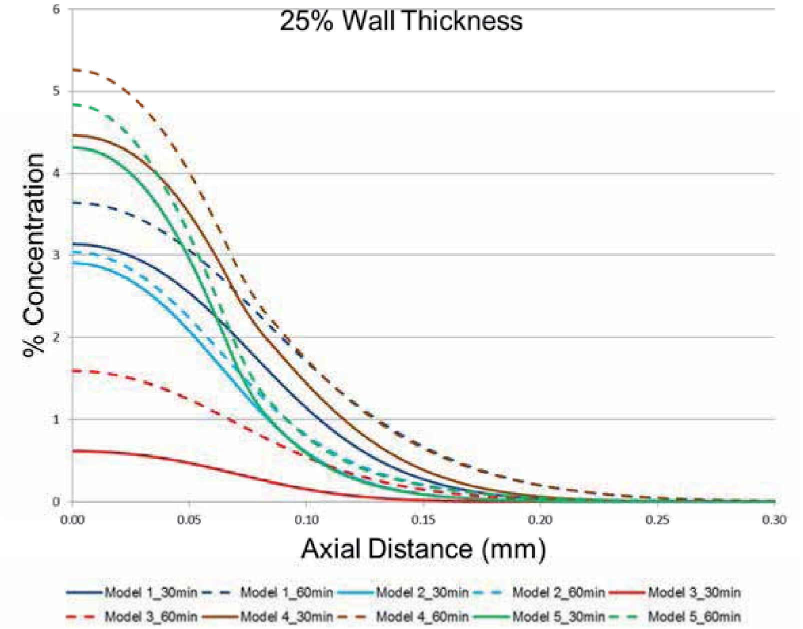


Fig. 6. Axial drug concentrations at 25% of the artery wall depth, illustrated in Fig 4 Model 1 by the 25% WT concentration line. Concentrations are measured after 30 minutes (solid line) and 60 minutes (dashed line) respectively.

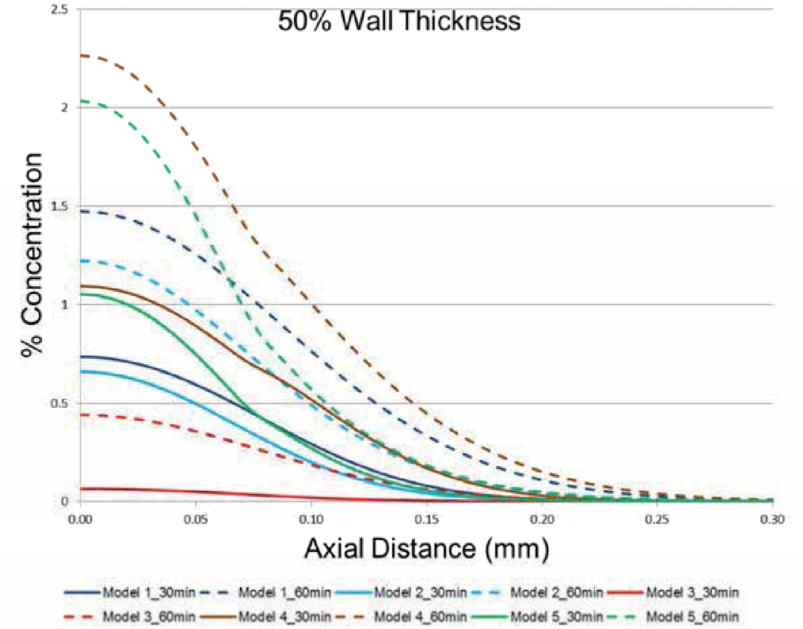


Fig. 7. Axial drug concentrations at 50% of the artery wall depth, illustrated in Fig 4 Model 1 by the 50% WT concentration line. Concentrations are measured after 30 minutes (solid line) and 60 minutes (dashed line) respectively.

The influence of compression on a porous wall in relation to mass transport has been demonstrated experimentally by O'Connell and Walsh (2010) and is evident in the results from Models 4 and 5. Histological evidence demonstrated that upon stent implantation some struts almost cut into the wall (Model 4) while others merely compress it providing gradual recovery either side of the stent strut (Model 5). Contrasting concentration profiles from these models elucidate to the inclusion of compression alone may not necessarily be adequate enough because the surrounding tissue orientation will also play a role in WSMT. Intuitively the greater the DES surface area in contact with arterial tissue the greater its ability to transport drugs into the wall. This holds true when comparing the concentration profiles of Models 4 and 5. The depth concentrations at the early time point for both models are very similar (Figure 5) but there is a notable difference in the axial concentration plots at both 25% and 50% of the wall thickness, whereby the 60% greater wall contact surface area of model 4 results in an increased concentration profile.

8. Conclusion

Understanding the behaviour of DES in the pursuit of improving device functionality is of great importance to clinicians and researchers alike. The modern application of computational techniques have greatly aided in achieving this goal, with researchers continuingly adding to the global understanding of drug mass transport from these devices. However, computational modelling isn't the complete solution because exact recreation of DES deployment and ensuing mass transport is not feasible for a number of reasons. No two patients will have identical stenosis of the coronary artery and therefore a single computational model will not provide the information required to comprehensively assess the viability of a single stent design. Instead a variety of models that will cover the spectrum of DES deployment scenarios is required and to computationally recreate these like for like with *in vivo* stenting conditions would be too computationally demanding for the same rewards that one could yield with a simplified analysis.

However, one must be mindful when simplifying the problem. The computational models developed in this chapter were created as 2-D axis-symmetric models as the study was intent on analysing the influence of relatively simple geometries. However, this may not always be the case and the need may arise where it would be necessary to model the problem in 3-D. The pitfalls of over simplification can be seen in the computational Model 1 where failure to model BSMT results in drug concentrations that are higher than that of the more realistic case, Model 2, that includes the luminal blood flow and mass transport therein.

Computational models where DES struts are flush against a bare artery wall have their merits but a greater degree of complexity needs to be implemented if an improved insight into DES mass transport within the coronary artery environment is to be gained. These arteries are heavily diseased and even a thin layer of plaque between the stent strut and the wall can inhibit WSMT. The compression of the porous artery wall upon stent expansion has an interesting effect on drug concentration within the wall. The reduction in artery wall diffusivity results in higher peak concentrations beneath the stent strut compared to the models where artery wall compression is not present. This pooling of drugs close to the stent strut is undesirable as it retards the early penetration of drugs into the wall, demonstrated in Figure 5 as the uncompressed artery wall of Model 2 recorded concentrations higher than that of Models 4 and 5 from a depth of approximately 50μm through the rest of the artery wall. There is also the danger of potential toxicity due to prolonged exposure of high drug concentrations close to the strut.

Analysis of mass transport from DES requires a multifaceted approach in order to predict behaviour of these devices and subsequently their response to *in vivo* arterial conditions. Future research that is mindful of preoperative DES design and postoperative environmental conditions will increase our knowledge of the second generation DES and enable us as a community to prepare for the advent of biodegradable DES.

9. Acknowledgements

The authors would like to thank the Irish Research Council for Science, Engineering and Technology (IRCSET), Grant no. RS/2005/159, for funding this body of work. The authors would also like to thank the members of CABER and the MSSI.

10. References

Balakrishnan, B., Dooley, J.F., Kopia, G. & Edelman, E.R. (2007).Intravascular Drug Release Kinetics Dictate Arterial Drug Deposition, Retention and Distribution. *Journal of Controlled Release*, Vol. 123, pp. 100-108.

Balakrishnan, B., Dooley, J.F., Kopia, G. & Edelman, E.R. (2008).Thrombus Causes Fluctuations in Arterial Drug Delivery from Intravascular Stents. *Journal of Controlled Release*. Vol. 131, pp. 173-180.

Balakrishnan, B., Tzafriri, A.R., Seifert, P., Groothuis, A., Rogers, C. & Edelman, E.R. (2005).Strut Position, Blood Flow, and Drug Deposition: Implications for Single and Overlapping Drug-Eluting Stents. *Circulation*, Vol. 111, pp. 2958-2965.

Burt, H.M. & Hunter, W.L. (2006).Drug-Eluting Stents: A Multidisciplinary Success Story. *Advanced Drug Delivery Review*, Vol. 58, pp. 350-357.

Chen, M.C., Liang, H.F., Chiu, Y.L., Chang, Y., Wei, H.J. & Sung, H.W. (2005). A Novel Drug-Eluting Stent Spray Coated with Multi-Layers of Collagen and Sirolimus. *Journal of Controlled Release*, Vol.108, pp. 178-189.

Costa, M.A. & Simon, D.I. (2005).Molecular Basis of Restenosis and Drug-Eluting Stents. *Circulation*, Vol. 111, pp. 2257-2273.

Devereux, P.D. (2005). Mass Transport Disturbances in the Downstream Junction of Peripheral Bypass Grafts. *Ph.D. Thesis, University of Limerick, Limerick, Ireland*.

Duraiswamy, N., Schoephoerster, R.T., Moreno, M.R. & Moore Jr, J.E. (2007). Stented Artery Flow Patterns and their Effects on the Artery Wall. *Annual Review of Fluid Mechanics*, Vol. 39, pp. 357–382.

Friedman, M.H. (2008). *Principles and Models of Biological Transport*. Springer Science+Business Media, ISBN 978-0-387-79239-2, New York.

Granada, J.F., Kaluza, G.L. & Raizner, A. (2003).Drug-Eluting Stents for Cardiovascular Disorders.*Current atherosclerosis reports*, Vol. 5, pp. 308-316.

Haller, J.D. & Olearchyk, A.S. (2002).Cardiology's 10 Greatest Discoveries. *Texas Heart Institution Journal*, Vol. 29, pp. 342–344.

Hara, H., Nakamura, M., Palmaz, J.C. & Schwartz, R.S. (2006).Role of Stent Design and Coatings on Restenosis and Thrombosis. *Advanced Drug Delivery Review*. Vol. 58, pp. 377-386.

Head, D.E., Sebranek, J.J., Zahed, C., Coursin, D.B. & Prielipp, R.C. (2007).A tale of Two Stents: Perioperative Management of Patients with Drug-Eluting Coronary Stents. *Journal of Clinical Anesthesia*, Vol. 19, pp. 386-396.

Hwang, C.W., Levin, A.D., Jonas, M., Li, P.H. & Edelman, E.R. (2005).Thrombosis Modulates Arterial Drug Distribution for Drug Eluting Stents. *Circulation*,Vol. 111, pp. 1619-1626.

Hwang, C.W., Wu, D.& Edelman, E. R. (2001). Physiological Transport Forces Govern Drug Distribution for Stent Based Delivery. *Circulation*, Vol. 104, pp. 600-605.

Kaul, S., Shah, P.& Diamond, G.A. (2007).As Time Goes By:Current Status and Future Directions in the Controversy Over Stenting. *Journal of the American College of Cardiology*, Vol. 50, No. 2, pp. 128–137.

Kaazempur-Mofrad, M.R., Wada S., Myers J.G. &Ethier C.R. (2005). Mass Transport and Fluid Flow in StenoticArteries: Axisymmetric and Asymmetric Models. *International Journal of Heat and Mass Transfer*, Vol. 48, pp. 4510-4517.

Kaazempur-Mofrad, M. R. & Ethier C.R. (2001).Mass Transport in an Anatomically Realistic Human Right Coronary Artery. *Annals of Biomedical Engineering*, Vol. 29, pp. 121-127.

Khakpour, M. & Vafai, K. (2008).A Comprehensive Analytical Solution of Macromolecular Transport within an Artery. *International Journal of Heat and Mass Transfer*, Vol. 51, pp. 2905-2913.

Khakpour, M. & Vafai, K. (2007).Critical Assessment of Arterial Transport Models. *International Journal of Heat and Mass Transfer*, Vol. 51, pp. 807-822.

Kolachalama, V.B., Tzafriri A.R., Arifin, D.Y. & Edelman, E.R. (2009). Luminal Flow Patterns Dictate Arterial Drug Deposition in Stent-Based Delivery. *Journal of Controlled Release*, Vol. 133, pp. 24–30.

Kukreja, N., Onuma, Y., Daemen, J. & Serruys, P.W. (2008).The Future of Drug-Eluting Stents. *Pharmacological Research*, Vol. 57, pp. 171–180.

Lutostansky, E.M., Karner, G., Rappitsch, G., Ku, D.N. & Perktold, K. (2003).Analysis of Hemodynamic Fluid Phase Mass Transport in a Separated Flow Region. *Journal of Biomechanical Engineering*, Vol. 125, pp. 189-196.

Markou, C.P., Lutostansky, E.M., Ku, D.N. & Hanson, S.R. (1998).A Novel Method for Efficient Drug Delivery. *Annals of Biomedical Engineering*, Vol. 26, pp. 502–511.

Mongrain, R., Faik, I., Leask, R., Rodes-Cabau, J., Larose, E. & Bertrand, O.F. (2007).Effects of Diffusion Coefficients and Strut Apposition Using Numerical Simulations for Drug Eluting Coronary Stents. *Journal of Biomechanical Engineering*. Vol. 129, pp. 733-742.

Mongrain, R., Leask, R., Brunette, J., Faik, I., Bulman-Feleming, N. & Nguyen, T. (2005).Numerical Modeling of Coronary Drug Eluting Stents. *Studies in Health Technology and Informatics*, Vol. 113, pp. 443-458.

O'Brien, T., Walsh, M. & McGloughlin, T. (2005).On Reducing Abnormal Haemodynamics in the Femoral End-to-Side Anastomosis: The Influence of Mechanical Factors. *Annals of Biomedical Engineering*. Vol. 33, pp. 309-321.

O'Connell, B.M. & Walsh, M.T. (2010).Demonstrating the Influence of Compression on Artery Wall Mass Transport. *Annals of biomedical Engineering*, Vol. 38, pp. 1354-136.

Parry, T.J., Brosius, R., Thyagarajan, R., Carter, D., Argentieri, D., Falotico, R. &Siekierka, J. (2005). Drug–Eluting Stents: Sirolimus and Paclitaxel Differently Affect Cultured Cells and Injured Arteries. *European Journal of Pharmacology*, Vol. 524, pp. 19-29.

Rajagopal, V. &Rockson, S.G. (2003).Coronary Restenosis: A Review of Mechanisms and Management. *American Journal of Medicine*, Vol. 115, pp. 547-533.

Rajamohan, D., Banerjee, R.K., Back, L.H., Ibrahim, A.A., Jog, M.A. (2006). Developing Pulsatile Flow in a Deployed Coronary Stent. *Journal of Biomechanical Engineering*, Vol. 128, pp. 347-359.

Schwartz, R.S., Chronos, N.A. & Vivmani, R. (2004).Preclinical Restenosis Models and Drug-Eluting Stents: Still Important, Still Much to Learn. *Journal of the American College of Cardiology*. Vol. 44, pp. 1373-1385.

Stone, G.W., Midei, M., Newman, W. et al. (2008).Comparison of an Everolimus-Eluting Stent and a Paclitaxel-Eluting Stent in Patients with Coronary Artery Disease: A Randomized Trial. *The Journal of The American Medicine Association*, Vol. 299, pp. 1903-1913.

Sun, N., Wood, N.B., Hughes. A.D., Thom, S. A. M. & Xu, X.Y. (2006). Fluid-Wall Modelling of Mass Transfer in an Axisymmetric Stenosis: Effects of Shear-Dependent Transport Properties. *Annals of Biomedical Engineering*, Vol. 34, pp. 1119-1128.

Tanabe, K., Regar, E., Lee, C.H., Hoye, A., Van der Giessen, W.J.&Serruys, P.W. (2004). Local Drug Delivery Using Coated Stents: New Developments and Future Perspectives. *Current Pharmaceutical Design*, Vol. 10, pp. 357-368.

Van der Hoeven, B.L., Pires, N.M.M., Warda, H.M., Oemrawsingh, P.V., Van Vlijmen, B.J.M., Quax, P.H.A., Schalij, M.J., Van der Wall, E.E. & Jukema, J.W. (2004). Drug Eluting Stents: Results, Problems and Promises. *International Journal of Cardiology*, Vol. 99, pp. 9-17.

Venkatraman, S. &Boey, F. (2007).Release Profiles in Drug-Eluting Stents: Issues and Uncertainties. *Journal of Controlled Release*, Vol. 120, pp. 149-160.

Waksman, R. (2002). Drug-Eluting Stents: From Bench to Bed. *Cardiovascular Radiation Medicine*, Vol. 3, pp. 226– 241.

Walsh, M.T., Kavanagh, E.G., O'Brien, T., Grace, P.A. & McGloughlin, T.(2003). On the Existence of an Optimum End-to-Side Junctional Geometry in Peripheral Bypass Surgery – A Computer Generated Study. *European Journal of Vascular and Endovascular Surgery*, Vol. 26, pp. 649-656.

14

Biosurfactants and Bioemulsifiers Biomedical and Related Applications – Present Status and Future Potentials

Letizia Fracchia[1], Massimo Cavallo[1],
Maria Giovanna Martinotti[1] and Ibrahim M. Banat[2]
[1]*Department of Chemical, Food, Pharmaceutical and Pharmacological Sciences,*
Drug and Food Biotechnology Center,
Università del Piemonte Orientale "Amedeo Avogadro", Novara,
[2]*School of Biomedical Sciences, Faculty of Life and Health Sciences,*
University of Ulster, Coleraine, N. Ireland,
[1]*Italy*
[2]*UK*

1. Introduction

Many microorganisms are able to produce a wide range of amphipathic compounds, with both hydrophilic and hydrophobic moieties present within the same molecule which allow them to exhibit surface activities at interfaces and are generally called biosurfactants or bioemulsifiers. These surface-active compounds (SAC) are mainly classified according to their mode of action, molecular weight and general physico-chemical properties.

In literature, the terms 'biosurfactants' and 'bioemulsifiers' are often used interchangeably, however in general those that reduce surface and interfacial tension at gas-liquid-solid interfaces are called biosurfactants and those that mainly reduce the interfacial tension between immiscible liquids or at the solid-liquid interfaces leading to the formation of more stable emulsions are called bioemulsifiers or bioemulsans. The former group includes low-molecular-weight compounds, such as lipopeptides, glycolipids, proteins, while the latter includes high-molecular-weight polymers of polysaccharides, lipopolysaccharides proteins or lipoproteins (Smyth et al., 2010a, 2010c).

In heterogeneous systems, biosurfactants tend to aggregate at the phase boundaries or interfaces. They form a molecular interfacial film that affects the properties (surface energy and wettability) of the original surface. This molecular layer, in addition to lowering the surface tension in liquids, also lowers the interfacial tension between different liquid phases on the interfacial boundary existing between immiscible phases and therefore can have an impact on the interfacial rheological behaviour and mass transfer.

When at interfaces (solid- liquid, liquid-liquid or vapour-liquid), the hydrophobic moiety of the surface active molecules aggregates at the surface facing the hydrophobic phase (usually the oil phase) while the hydrophilic moiety is oriented towards the solution or hydrophilic phase (mainly water). Their diverse functional properties namely, emulsification, wetting, foaming, cleansing, phase separation, surface activity and reduction in viscosity of heavy liquids such as crude oil, make them suitable for utilization for many industrial and domestic application purposes (Gautam & Tiagi, 2006; Franzetti et al., 2010a; Perfumo et al., 2010a; Satpute et al., 2010b).

During the past two decades biosurfactants have been under continuous investigation as a potential replacement for synthetic surfactants and are expected to have several industrial and environmental applications mainly related to detergency, emulsification, dispersion and solubilisation of hydrophobic compounds (Banat et al., 2000). In addition, biosurfactants present several advantages over surfactants of a chemical origin, particularly in relation to their biodegradability, environmental compatibility, low toxicity, high selectivity and specific activity at extreme temperatures, pH and salinity (Banat 1995a, 1995b). Due to all these properties, they have steadily gained increased significance in industrial and environmental applications such as bioremediation, soil washing, enhanced oil recovery and other general oil processing and related industries (Perfumo et al., 2010b). Furthermore, potential commercial applications in several other industries including paint, cosmetics, textile, detergent, agrochemical, food and pharmaceutical industries begin to emerge (Banat et al., 2000).

Numerous investigations in the field of biosurfactants/bioemulsifiers are leading to the discovery and description of many interesting chemical and biological properties and potential biomedical therapeutic and prophylactic applications. In this chapter we will focus on the most recent and appealing biomedical and therapeutic applications of biosurfactants and bioemulsifiers with special emphasis on the most recent results in the fields of biotechnology, nanotechnology and bioengineering.

2. Classification, properties and functional mechanisms of microbial surface-active compounds

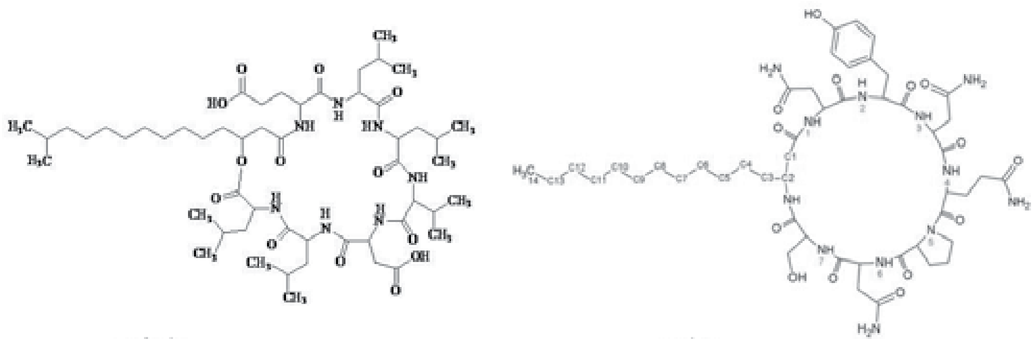
Microbial surface-active compounds are a range of structurally diverse molecules produced by different microorganisms and are mostly therefore classified by their structural features, the producing organism and their molecular mass. Their hydrophilic moiety is mainly comprised of an acid, peptide cations, or anions, mono-, di- or polysaccharides while their hydrophobic moiety can be an unsaturated or saturated hydrocarbon chains or fatty acids. The structural orientation on the surfaces and inter phases confers the range of properties, such as the ability to lower surface and interfacial tension of liquids and the formation of micelles and microemulsions between these different phases (Chen et al., 2010a, 2010b).

2.1 Low molecular weight compounds

2.1.1 Lipopeptides

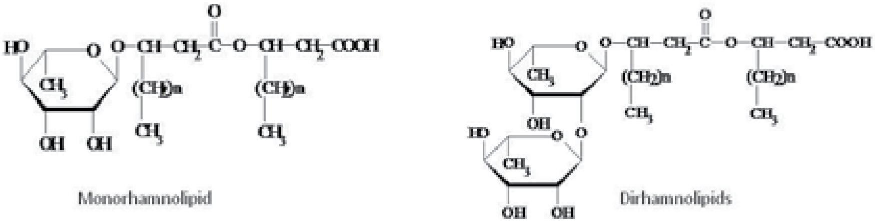
Microbial surface-active compounds can be roughly divided into low molecular weight molecules that efficiently reduce surface and interfacial tension (biosurfactants) (Fig. 1.), and high molecular weight polymers that stabilize emulsions but do not lower the surface tension as much (bioemulsans or bioemulsifiers) (Fig. 2.) (Neu, 1996; Rosenberg, 2006; Rosenberg & Ron, 1997; Smyth et al., 2010a, 2010c).

The most studied low-molecular-weight biosurfactant compounds are lipopeptides and glycolipids. Lipopeptides are mainly produced by members of the *Bacillus* species; they are composed of different families and each family is constituted of several variants, which can differ in their fatty acid chain and their peptide moiety (Dastgheib et al., 2008; Jacques, 2010; Thavasi et al., 2008, 2011).



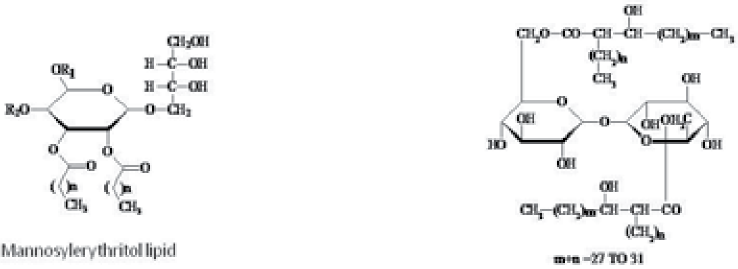
Surfactin

Iturin A



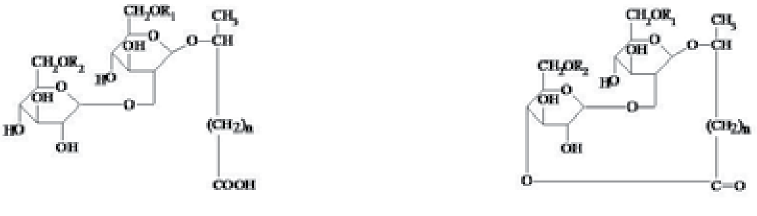
Monorhamnolipid

Dirhamnolipids



Mannosylerythritol lipid

Trehalose dimycolate



Acidic Sophorolipid

Lactonic Sophorolipid

Fig. 1. Chemical structures of the main low molecular weight microbial surface active compounds reported; surfactin, iturin A, mono & di-rhamnolipids, mannosylerythritol lipids, dimycolates trehalose lipids, acidic and lactonic sophorolipids.

Surfactin, a cyclic lipopeptide produced by *Bacillus subtilis* is considered the most active biosurfactant discovered so far (Ron & Rosenberg, 2001). Surfactin was discovered by

Arima et al., (1968) from the culture broth of *Bacillus subtilis* and it was named thus due to its exceptional surfactant activity (Peypoux et al., 1999). Natural surfactins are a mixture of isoforms A, B, C and D which are classified according to the differences in their amino acid sequences and possess various physiological properties (Shaligram & Singhal, 2010). Surfactin is composed of a seven amino-acid ring structure coupled to a fatty-acid chain via a lactone linkage. Surfactin-A has L-leucine, surfactin-B has L-valine and surfactin-C has L-isoleucine at the amino acid position involved in the lactone ring formation with the C14–C15 β-hydroxy fatty acid. The amino-acid residues may vary and the presence of these variants can be related to alterations in the culture conditions such as providing substrate containing some specific amino-acid residues in the culture media (Jacques, 2010).

Another surfactin related compound is lichenysin, a lipopeptide discovered in the supernatant of *Bacillus licheniformis* culture (Horowitz et al., 1990). Its chemical structure and physio-chemical properties are similar to surfactin (McInerney et al., 1990). In particular, lichenysin has Glutamine amino-acid in position 1 while surfactin has Glutamic acid. Other surfactin-like compounds are pumilacidin A, B, C, D, E, F and G, a complex of acylpeptide antibiotics isolated from *Bacillus pumilus* culture supernatants with interesting antiviral properties (Morikawa et al., 1992; Naruse et al., 1990). Among the lipopeptides belonging to the iturin family, iturin A is the most studied compound. It is a heptapeptide interlinked with a β-amino-acid fatty acid with carbon chain length from C14 to C17 (Peypoux, 1978, as cited in Jacques, 2010) produced by *Bacillus subtilis* strains reported to have antifungal activities (Besson et al., 1976).

Other members of the iturin family are iturin C, bacillomycin D, F, and Lc and mycosubtilin (Bonmatin et al., 2003). The family of fengycins includes fengycins A and B, lipodecapeptides which differ by their amino-acid residue in position 6 that can be Alanine or Valine and are known for their interesting fungitoxic and immunomodulating activities (Jacques, 2010). Other interesting lipopeptides are serrawettins, nonionic cyclodepsipeptide biosurfactants produced by *Serratia marcescens* (Matsuyama et al., 2010) and implicated with anti-tumor and anti-nematode activities.

2.1.2 Glycolipids

Are commonly mono or disaccharides compounds acylated with long chain fatty acids or hydroxyl fatty acids. Among them, rhamnolipids, mannosylerythritol lipids (MELs), sophorolipids and trehalolipids are the best-studied structural subclasses.

Rhamnolipids are glycosides, produced mainly by *Pseudomonas aeruginosa* and by the *Burkholderia* genus, that are composed of one (for mono-rhamnolipids) or two (for di-rhamnolipids) rhamnose sugar moieties linked to one or two β-hydroxyfatty acid chains (Perfumo et al., 2006; Raza, 2009). These molecules display high surface activities and many potential applications in the biomedical field due to their antibacterial, antifungal, antiviral, antiadhesive reported properties (Abalos et al., 2001; Cosson et al., 2002; Kim et al., 2000; Remichkova et al., 2008; Sotirova et al., 2008; Yoo et al., 2005). They have also been used in the preparation of nanoparticles (Palanisamy & Raichur, 2009; Xie et al., 2006) and microemulsions (Nguyen & Sabatini, 2009; Xie et al., 2007).

The mannosylerythritol (MELs) glycolipids are produced by the yeasts strains of the genus *Pseudozyma* sp. and *Ustilago* sp. from soybean oil or *n*-alkane (Arutchelvi & Doble, 2010). MELs are a mixture of partially acylated derivative of 4-*O*-*β*-D-mannopyranosyl-D-erythritol, containing $C_{2:0}$, $C_{12:0}$, $C_{14:0}$, $C_{14:1}$, $C_{16:0}$, $C_{16:1}$, $C_{18:0}$ and $C_{18:1}$ fatty acids as the hydrophobic groups (Bhattacharjee et al., 1970, as cited in Arutchelvi & Doble, 2010). Based on the degree of acetylation at C4 and C6 position, and their order of appearance on the thin layer chromatography, the MELs are classified into MEL-A, -B, -C and –D (Arutchelvi & Doble, 2010). MEL-A representing the diacetylated compound while MEL-B and MEL-C are monoacetylated at C6 and C4, respectively. The completely de-acetylated structure is known as MEL-D (Rau et al., 2005, as cited in Arutchelvi & Doble, 2010).

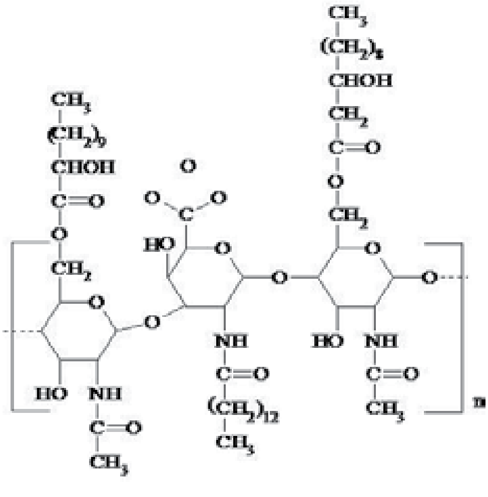
MELs have recently gained attention due to their environmental compatibility, mild production conditions, structural diversity, self-assembling properties and versatile biochemical functions. In particular, interesting applications have been described in the biomedical field as antimicrobial, antitumor and immunomodulating molecules, in the biotechnological field for gene and drug delivery, and in cosmetic applications as skin moisturizers (Arutchelvi & Doble, 2010).

Sophorolipids are another extracellular glycolipids synthesized by some yeast species including *Candida bombicola*, *Candida apicola*, *Rhodotorula bogoriensis, Wickerhaminella domercqiae* and *Candida batistae* (Van Bogaert & Soetaert, 2010). They consist of two glucose units linked β-1,2. The 6- and 6′-hydroxyl groups are generally acetylated. The lipid portion is connected to the reducing end through a glycosidic linkage. The terminal carboxyl group of the fatty acid can be in the lactonic form or hydrolyzed to generate an anionic surfactant (Rosenberg & Ron, 1999). Sophorolipids have been reported suitable for a number of application in the biomedical field including use as antimicrobial, antiviral and anticancer. They also have been used in the synthesis of metal-bound nanoparticles in cosmetic and pharmacodermatological products (Van Bogaert & Soetaert, 2010).

Trehalose lipids are also a glycolipids containing threhalose as the sugar moiety which is a non-reducing disaccharide in which the two glucose units are linked in an α,α-1,1-glycosidic linkage. It is the basic component of the cell wall glycolipids in *Mycobacteria* and *Corynebacteria* (Franzetti et al., 2010b). The most reported trehalose lipid is trehalose 6,6′-dimycolate, which is a α-branched chain mycolic acid esterified to the C6 position of each glucose. Different trehalose containing glycolipids are known to be produced by several other microorganisms belonging to mycolates group, such as *Arthrobacter*, *Nocardia*, *Rhodococcus* and *Gordonia*. *Rhodococcus* genus in particular produced several types of trehalose lipids as reported by Lang & Philp (1998). These glycolipids vary in the number and overall chain length (C20–C90) of the esterified fatty acids. Beside their known industrial applications, trehalose lipids recently attracted attention to their functions in cell membrane interaction and their potential as antitumor therapeutic agents (Aranda et al., 2007, Harland et al., 2009, Imasato et al., 1990, Isoda et al., 1995, as cited in Shao, 2010; Ortiz et al., 2008, 2009; Zaragoza et al., 2009, 2010).

2.2 High molecular weight biosurfactants

These are generally grouped together as polymeric biosurfactants. They are produced by a number of different bacteria and are composed of lipoproteins, proteins, polysaccharides, lipopolysaccharides or complexes containing several of these structural types (Ron & Rosenberg, 2001; Rosenberg & Ron, 1997, 1999). The most commonly studied biopolymer is emulsan (Fig. 2.), a lipopolysaccharide isolated from *Acinetobacter calcoacetius* RAG-1 ATCC 31012 with a molecular weight of around 1,000 kDa (Rosenberg et al., 1979).



Emulsan

Fig. 2. Chemical structure of most known high molecular weight microbial surface active compound; emulsan.

RAG-1 emulsan is a complex of an anionic heteropolysaccharide and protein (Rosenberg & Kaplan, 1987, as cited in Rosenberg & Ron, 1999). Its surface activity is due to the presence of fatty acids, comprising 15% of the emulsan dry weight, which are attached to the polysaccharide backbone via O-ester and N-acyl linkages (Belsky et al., 1979, as cited in Rosenberg & Ron, 1999).

Another high molecular weight biosurfactant is alasan, a complex of an anionic polysaccharide and a protein with a molecular weight of around 1,000 kDa isolated from *Acinetobacter radioresistens* (Navon-Venezia et al., 1995, as cited in Smyth et al., 2010c). These high molecular weight biosurfactants generally possess effective emulsifying activity and are called bioemulsifiers. A large number of other polymeric compounds have been discovered but remain partially or totally uncharacterized (Smyth et al., 2010c). Little is known in general about these bioemulsifiers other than the producing organism and the overall chemical composition of the crude mixture. *Halomonas eurihalina* produces an extracellular sulfated heteropolysaccharide (Calvo et al., 1998, as cited in Rosenberg & Ron, 1999). *Pseudomonas tralucida* produced an extracellular acetylated polysaccharide that was effective in emulsifying several insecticides (Appaiah & Karanth 1991, as cited in Rosenberg & Ron, 1999).

Several bioemulsifiers are effective at high temperature, including the protein complex from *Methanobacterium thermoautotrophium* (De Acevedo et al., 1996, as cited in Rosenberg & Ron, 1999) and the protein-polysaccharide-lipid complex of *Bacillus stearothermophilus* ATCC 12980 (Gunjar et al., 1995, as cited in Rosenberg & Ron, 1999). Yeasts produce a number of emulsifiers, which are particularly interesting because of the food-grade status of several yeasts which allows use in food related industries. Liposan is an extracellular emulsifier produced by *Candida lipolytica* (Cirigliano & Carman, 1985, as cited in Rosenberg & Ron, 1999). It is composed of 83% carbohydrate and 17% protein. Mannanprotein emulsifiers are produced by *Saccharomyces cerevisiae* (Cameron et al., 1988, as cited in Rosenberg & Ron, 1999). Many of these bioemulsifiers have been used in the food, cosmetic, and petroleum industries (Rosenberg & Ron, 1999).

2.3 Properties and functions of biosurfactants

There is a growing interest in the study of the physicochemical and biological properties of biosurfactants because of their potential industrial applications (Cameotra & Makkar, 2004; Desai & Banat, 1997; Lang, 2002; Rodrigues et al., 2006a; Singh & Cameotra, 2004). The interesting biological activities displayed by these compounds constitute an added value to their potential uses (Lang et al., 1989; Lang & Wagner, 1993; Stanghellini & Miller, 1997, as cited in Sánchez et al., 2010). Due to these reasons, an intense research activity is currently directed toward identification of new biosurfactants and characterization of their chemical and biological properties (Biria et al., 2010; Morita et al., 2009a; Satpute et al., 2010a; Singh & Cameotra, 2004; Singh et al., 2007).

The most obvious property of biosurfactants compounds is their ability to effectively lower water surface tension, and a number of approaches that measure directly the surface activity of biosurfactants can be used as screening methods for their detection. Among them, the most frequently used as quick and simple techniques are the drop collapse (Bodour and Miller-Maier, 1998) and the oil spreading tests (Morikawa et al., 2000) (Fig. 3).

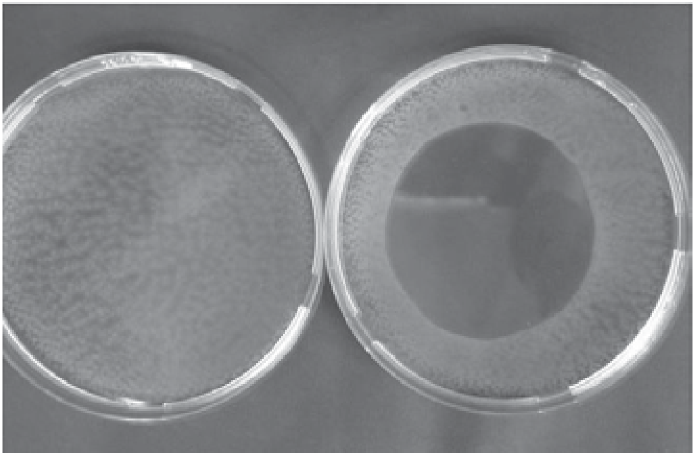


Fig. 3. Oil spreading test. This technique measures the diameter of clear zones caused when a drop of a biosurfactant-containing solution is placed on an oil–water surface (Morikawa et al., 2000). Photo courtesy by Fabrizio Rivardo.

An efficient biosurfactant can reduce the surface tension between pure water and air from 72mN/m to less than 30mN/m. Surfactin, as one of the most powerful biosurfactants, can reduce the surface tension of water from 72mN/m to 27mN/m (which is close to the minimum detectable value) at a concentration as low as 10 µM (Seydlová & Svobodová, 2008). Rhamnolipids can similarly achieve such level of reduction (Hisatsuka et al., 1971, as cited in Muthusamy et al., 2008; Mohammad Abdel-Mawgoud et al., 2010). The sophorolipids from *T. bombicola* have been reported to reduce the surface tension to 33mN/m (Muthusamy et al., 2008) while MELs and trehalose lipids to less than 30mN/m (Arutchelvi & Doble, 2010; Shao et al., 2010).

As surfactant monomers are added into solution, the surface or interfacial tension will decrease until the biosurfactant reaches the critical micelle concentration (CMC). The CMC is defined as the minimum concentration necessary to initiate micelle formation (Becher, 1965). Above the CMC no further reduction in surface or interfacial tension is observed. At the CMC, surfactant monomers begin to spontaneously associate into structured aggregates such as micelles, vesicles or continuous bilayers. These aggregates are produced as a result of numerous weak chemicals interactions such as hydrophobic, van der Waals and hydrogen bonding (Maier, 2003; Raza et al., 2010). Since no chemical bonds are formed, these structures are fluid-like and are easily transformed from one state to another as conditions such as electrolyte concentration and temperature are changed (Lin, 1996).

The aggregate structure depends on the polarity of the solvent in which the surfactant is dissolved. In an aqueous solution, the polar head groups of a micelle will be oriented outward toward the aqueous phase, and the hydrophobic tails will associate in the core of the micelle within an oil-in-water micelle. In contrast, in oil, the polar head groups will associate in the center of the micelle while the hydrophobic tails will be oriented toward the outside within the water–in-oil micelle (Soberón-Chávez & Maier, 2010). Efficient surfactants have a low CMC, i.e. less surfactant is necessary to decrease the surface tension (Seydlová & Svobodová, 2008).

Biosurfactants are most effective and efficient at their CMC which can be 10–40 times lower than that of chemical surfactants, i.e. less surfactant is necessary to get a maximum decrease in surface tension (Desai & Banat, 1997). Another important property for industrial and biotechnological applications is that most biosurfactants surface activities are not affected by environmental conditions such as temperature and pH (Muthusamy et al., 2008) particularly those of glycolipids composition. It has been reported that lichenysin from *B. licheniformis* JF-2 was not affected by temperature (up to 50 °C), pH (4.5–9.0) and NaCl and Ca concentrations up to 50 and 25 g/l respectively. A lipopeptide from *B. subtilis* LB5a was also stable after autoclaving (121°C/20 min) and after 6 months at –18°C; the surface activity did not change from pH 5 to 11 and NaCl concentrations up to 20% (Muthusamy et al., 2008).

Moreover, unlike synthetic surfactants, microbial-produced compounds are easily degraded and are generally considered as low or non-toxic products and therefore, appropriate for pharmaceutical, cosmetic and food uses. Although little is known about the toxicity of microbial surfactants, some data in the literature suggest that they are less toxic than synthetic surfactants (Muthusamy et al., 2008). The synthetic anionic surfactant (Corexit) for example had an LC50 (lethal concentration to 50% of test species) against *Photobacterium phosphoreum* at approximately ten times lower concentrations than that for rhamnolipids, demonstrating the higher toxicity of the chemical-derived surfactant. It was also reported that biosurfactants needed higher effective concentration to decrease 50% of test population values (EC50) and were degraded faster than commercial dispersants. In another report, biosurfactants from *P. aeruginosa* were noted to have much less toxic and mutagenic activities when compared to synthetic surfactant Marlon A-350, which is widely used in the industry.

Understanding the functional mechanisms of biosurfactants and bioemulsifiers is of great help to discover interesting applications. Surfactin, one of the most powerful biosurfactants, is known to destabilize membranes disturbing their integrity and permeability (Bernheimer et al., 1970). This is due to changes in physical membrane structure or through disrupting protein conformations which alter important membrane functions such as transport and energy generation (Ortiz et al., 2008, 2009; Sánchez et al., 2009, 2010; Sotirova et al., 2008; Van Hamme et al., 2006; Zaragoza et al., 2009).

The molecular mechanisms of surfactin interactions with membrane structures have been described by Shaligram & Singhal (2010) and by Seydlová & Svobodová (2008). A key step for membrane destabilization and leakage is the dimerization of surfactin into the bilayer (Carrillo et al., 2003). The hypothetical mechanisms of surfactin interactions with membranes exhibit a complex pattern of effects such as insertion into the lipid bilayers, chelating mono- and divalent cations, modification of membrane permeability by channel formation or membrane solubilization by a detergent-like mechanism. *In vitro,* the incorporation of surfactin into the membrane gives rise to dehydration of the phospholipid polar head groups and the perturbation of lipid packing which strongly compromise the bilayer stability, leading to the disturbance of the membrane barrier properties.

These structural fluctuations may well explain the primary mode of the antibiotic action and the other important biological effects of this lipopeptide (Carrillo et al., 2003). The extent of perturbation of the phospholipid bilayer correlates with the concentration of surfactin. At low concentrations surfactin penetrates readily into the cell membrane, where it is completely miscible with the phospholipids and forms mixed micelles. At moderate concentrations, the lipopeptide forms domains segregated within the phospholipid bilayer that may contribute to the formation of ion-conducting pores in the membrane leading to membraine disruption and permeabilization at high concentrations, showing a stronger activity than that of Triton (Heerklotz et al., 2007, as cited in Seydlová & Svobodová, 2008).

As biological amphiphilic molecules, biosurfactants naturally tend to self-assemble into hierarchically ordered structures using hydrogen bonding, hydrophobic and van der Waals interactions (Kitamoto et al., 2009). Glycolipids, and in particular MELs, are well known for their self-assembling properties, that are influenced by the stereochemistry of the saccharide head groups (Kitamoto et al., 2005). Some of glycolipid type surfactants, which possess relatively large hydrophilic head groups as compared to the hydrophobic parts, generally form micelles in a dilute aqueous solution. Other than spherical micelles, they also form oblate (disk-like) and prolate (rod-like) structures (Söderman, 2000, as cited in Kitamoto et al., 2009). As the surfactant concentration further increases, glycolipid/water systems start to form a range of liquid crystalline phases. In particular, glycolipid biosurfactants spontaneously self-assemble into a variety of molecular assemblies with well-defined and/or unique structures, such as sponge (L3), cubic (V2), hexagonal (H2), or lamellar (L_α) configurations (Imura et al., 2007, as cited in Kitamoto et al., 2009).

Among these molecular assemblies, vesicles are one of the most intensively studied ones. MELs in particular, due to their efficient molecular orientation property and effective balance between hydrophilic and hydrophobic groups, are able to form giant vesicles of diameter larger than 10 µm (Kitamoto et al., 2002). In comparison, rhamnolipids show a pH-sensitive conversion of molecular assemblies due to the presence of a carboxyl group on the side chain (Kitamoto et al., 2002). This leads rhamnolipids to form micelles at pH more than 6.8, lipid particles at pH 6.6-6.2, lamella structures at pH 6.5-6.0, and finally vesicles in the size of 50-100 nm at pH 5.8-4.3. Glycolipid biosurfactant-based vesicles or bilayer membranes appear, thus, to be very promising for exploiting useful nanostructured materials and/or systems.

Another function of microbial surface-active molecules with interesting biotechnological potential is the ability to form stable emulsions (Fig. 4.). High molecular-mass biosurfactants are in general better emulsifiers than low-molecular-mass biosurfactants and are thus called bioemulsifiers. Bioemulsifiers, can form and stabilize oil in water or water in oil emulsions, but are not necessarily efficient detergents that are able to demonstrate remarkable surface tension reduction (Dastgheib et al., 2008). Emulsions can be produced with prolonged lifespan of months and years. Liposan, for example, does not reduce surface tension, but has been used successfully to emulsify edible oils (Cirigliano & Carman, 1985). Emulsan is an effective emulsifier at low concentrations (0.01-0.001%) representing emulsan-to-hydrocarbon ratios of 1:100 to 1:1000, while exhibiting considerable substrate specificity (Ron & Rosemberg, 2001). Polymeric surfactants offer additional advantages because they coat droplets of oil, thereby forming stable emulsions. This property is especially useful for making oil/water emulsions for cosmetics and food industries (Muthusamy et al., 2008).

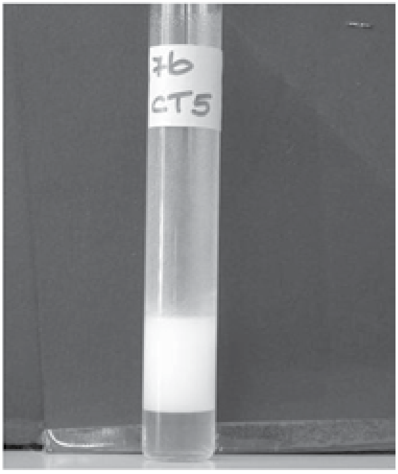


Fig. 4. Example of emulsion produced by the bioemulsifier extracted from the bacterial strain 7bCT5, isolated from a Cambodian soil. This emulsion is stable since 2009.

2.3.1 Natural roles

Although biosurfactants are produced by a large number of microorganisms and are clearly significant in many aspects of growth, it is difficult to generalize on their roles in microbial physiology. Due to their very diverse chemical structures and surface properties, different groups of biosurfactants may have different natural roles in the growth of the producing microorganisms and probably provide advantages in a particular ecological niche. Ron & Rosenberg, (2001) and Van Hamme et al., (2006) recently reviewed the physiological roles of microbial surface-active compounds. Some biosurfactants are essential for the motility of the microorganisms (gliding and swarming). For example, serrawettin plays a fundamental role for surface locomotion and access to water repelling surfaces for *Serratia marcescens* whereas surfactin, together with flagellar biosynthesis, are crucial for swarming motility in *B. subtilis* (Arutchelvi et al., 2008; Van Hamme et al., 2006). Bioemulsifiers also play an important role in regulating the attachment-detachment of microorganisms to and from surfaces (Van Hamme et al., 2006).

In addition, bioemulsifiers are involved in cell-to-cell interactions such as bacterial pathogenesis, quorum sensing and biofilm formation, maintenance and maturation. Rhamnolipids, for example, are essential to maintain the architecture of the biofilms and are considered as one of the virulence factors in *Pseudomonas* sp. (Arutchelvi et al., 2008; Ron & Rosenberg, 2001; Van Hamme et al., 2006). Rhamnolipids, mannosylerythritol lipid and surfactin show antimicrobial and antibiotic properties thus conferring a competitive advantage to the organism during colonization and cell–cell competition. In addition cellular differentiation, substrate accession and resistance to toxic compounds are all roles attributed to microbial surface-active compounds. Their most widespread role however is believed to be the interaction between microbes and insoluble substrates such as hydrocarbons. Some biosurfactants/bioemulsifiers enhance the growth of bacteria on hydrophobic water-insoluble substrates by increasing their bioavailability, presumably by increasing their surface area, desorbing them from surfaces and increasing their apparent solubility (Neu, 1996; Ron & Rosenberg, 2001; Van Hamme et al., 2006).

3. Biomedical applications of microbial surface-active compounds

The use and potential commercial applications of biosurfactants in the medical field have increased during the past decade. Their antibacterial, antifungal and antiviral activities make them relevant molecules for applications in combating many diseases and as therapeutic agents. Furthermore, biosurfactants are generally considered safer than synthetic pharmaceuticals, due to their biological origin. Their pertinence in these fields is related to their biological properties such as the ability to disrupt membranes leading to cell lysis and metabolite leakage through increased membrane permeability and hence antimicrobial activity. Moreover, similarly to organic-conditioning films, their ability to partition at the interfaces can affect the adhesion properties of cells/microorganisms. Biomedical applications of biosurfactants have been thoroughly described (Banat et al., 2010; Cameotra & Makkar, 2004; Rodrigues et al., 2006a; Rodrigues & Teixeira, 2010; Seydlová & Svobodová, 2008; Singh & Cameotra, 2004).

3.1 Antimicrobial activity of biosurfactants

The search for new antimicrobial drugs remains a major concern nowadays because of the newly emerged pathogenic microorganisms and traditional others which have become virtually unresponsive to existing antibiotics. In fact, no novel or effective chemical antibiotics have been discovered during the last few decades (Hancock & Chapelle, 1999). Microbial metabolites have been recognized as a major source of compounds endowed with ingenious structures and potent biological activities (Donadio et al., 2002). Among these, some biosurfactants have been reported to be suitable alternatives to synthetic medicines and antimicrobial agents and may therefore be used as effective and safe therapeutic agents (Banat et al., 2000; Cameotra & Makkar, 2004; Singh & Cameotra, 2004).

3.1.1 Antibacterial activity

Lipopeptides have the most potent antimicrobial activity and have been a subject of several studies on the discovery of new antibiotics. The antibiotic activity is due to the ability of molecules of lipopeptide biosurfactants to self-associate and form a pore-bearing channel or micellular aggregate inside a lipid membrane (Carrillo et al., 2003; Deleu et al., 2008). Surfactin, in particular, has been associated with several physical and biological actions, such as antimicrobial, antiviral, anti-mycoplasma and haemolytic activities. It can penetrate into the membrane through hydrophobic interactions, thus influencing the ordering of the hydrocarbon chains and thus varying the membrane thickness (Bonmatin et al., 2003). Such membrane disruptions are a nonspecific mode of action and are advantageous for action on different cell membranes of both Gram-positive and Gram-negative bacteria (Lu et al., 2007). It has been suggested that such action by surfactin type peptides on membrane integrity rather than other vital cellular processes may perhaps constitute the next generation of antibiotics (Rodrigues & Teixeira, 2010). Similar bioactive fractions from the marine *Bacillus circulans* biosurfactant had antimicrobial action against various Gram-positive and Gram-negative pathogenic and semi-pathogenic bacteria including *Micrococcus flavus, Bacillus* *pumilis, Mycobacterium smegmatis, Escherichia coli, Serratia marcescens, Proteus vulgaris, Citrobacter freundii, Proteus mirabilis, Alcaligenes faecalis, Acetobacter calcoaceticus, Bordetella bronchiseptica, Klebsiella aerogenes* and *Enterobacter cloacae* (Das et al., 2008). The chemical identity of this bioactive biosurfactant fraction showed overlapping patterns with that of surfactin lipopeptides and lichenysin. Mild antimicrobial action was also observed against methicillin-resistant *Staphylococcus aureus* (MRSA) and other MDR pathogenic strains. The biosurfactant was also found to be nonhaemolytic in nature thus indicating possible use as a drug in antimicrobial chemotherapy.

Very recently Huang et al., (2011) evaluated antimicrobial activity of surfactin and polylysine against *Salmonella enteritidis* in milk using a response surface methodology and showed *S. enteritidis* to be very sensitive to both molecules with minimum inhibitory concentrations of 6.25 and 31.25 μg/mL, respectively. The optimization of antimicrobial activity indicated that *S. enteritidis* could be reduced by 6 orders of magnitude at a temperature of 4.45°C, action time of 6.9 h, and concentration of 10.03 μg/mL (surfactin/polylysine weight ratio, 1:1).

In addition to surfactin, *Bacillus subtilis* strains produce a broad spectrum of bioactive peptides with great potential for biomedical applications, such as fengycin (Vanittanakom et al., 1986) and the iturin compounds: iturins (Besson et al., 1978; Peypoux et al., 1978), mycosubtilins (Peypoux et al., 1986), and bacillomycins (Peypoux et al., 1984), all of which are amphiphilic surface and membrane-active compounds with potent antimicrobial activities. Huang et al., (2007) reported that a lipopeptide antimicrobial substance produced by *B. subtilis* fmbj strain, which is mainly composed of surfactin and fengycin, was able to inactivate endospores of *B. cereus* through damaging the surface structure of the spores as seen by Transmission Electron Microscopy.

Lichenysin, pumilacidin and polymyxin B (Grangemard et al., 2001; Landman et al., 2008; Naruse et al., 1990; Yakimov et al., 1995) are other antimicrobial lipopeptides produced by *Bacillus licheniformis*, *Bacillus pumilus* and *Bacillus polymyxa*, respectively. Polymyxin B, in particular, due to its high affinity for the lipid moieties of lipopolysaccharide, has shown antibacterial activities against a wide variety of Gram-negative pathogens. Being a cationic agent, it binds to the anionic bacterial outer membrane, leading to a detergent effect that disrupts membrane integrity. Important nosocomial pathogens such as *Escherichia coli*, *Klebsiella* spp., *Enterobacter* spp., *Pseudomonas aeruginosa*, and *Acinetobacter* spp. are usually susceptible to polymyxins and considerable activity has been reported against *Salmonella* spp., *Shigella* spp., *Pasteurella* spp., and *Haemophilus* spp. (Landman et al., 2008).

Another promising example of an antimicrobial lipopeptide that is under commercial development is daptomycin (Cubicin®). It has been approved for the treatment of skin infections by the FDA in 2003 (Giuliani et al., 2007, as cited in Seydlová & Svobodová, 2008). Daptomycin produced by *Streptomyces roseosporus* has been shown to be highly active against multiresistant bacteria such as MRSA (Tally & De Bruin, 2000, as cited in Seydlová & Svobodová, 2008). Another lipopeptide with antimicrobial activity and other interesting biological properties is viscosin, a cyclic lipopeptide from *Pseudomonas* (Saini et al., 2008).

Glycolipids, both rhamnolipids (Abalos et al., 2001; Benincasa et al., 2004) and sophorolipids (Kim et al., 2002; Van Bogaert et al., 2007) also have shown interesting antimicrobial activities (Fig. 5). Benincasa et al., (2004) reported that a mixture of six rhamnolipides homologues performed very well against *Bacillus subtilis* with a MIC of 8 μg/mL. Mannosylerythritol lipids (MEL-A and MEL-B) produced by *Candida antarctica* strains have also been reported to exhibit antimicrobial action against Gram-positive bacteria (Kitamoto et al., 1993).

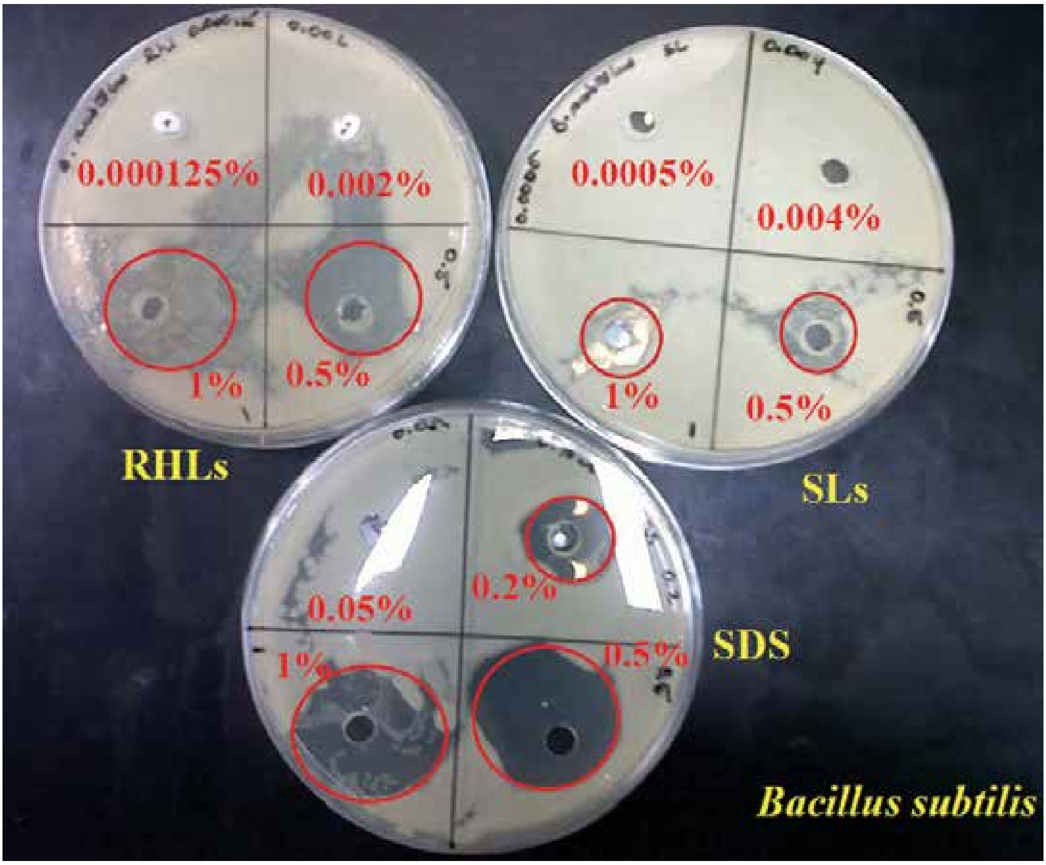


Fig. 5. Measuring antimicrobial activity for rhamnolipids, sophorolipids and SDS at various concentrations above and below the CMC for these surface active molecules against *Bacillus subtilis,* red circles showing clearing/inhibition zones.

Very recently, Nitschke et al., (2010) reported rhamnolipids produced by *P. aeruginosa* LBI with antimicrobial activity against several bacteria and fungi, including *Bacillus cereus*, *Staphylococcus aureus*, *Micrococcus luteus*, *Mucor miehei* and *Neurospora crassa*. Another glycolipid, flocculosin, a cellobiose lipid produced by the yeast-like fungus *Pseudozyma flocculosa*, was particularly effective against *Staphylococcus* species, including MRSA. Its antibacterial activity was not influenced by the presence of common resistance mechanisms against methicillin and vancomycin and it was able to eliminate *C. albicans* cells in a very short period of time (Mimee et al., 2009).

Trehalose lipids produced by *Tsukamurella* sp. strain DSM 44370 together with trisaccharide and tetrasaccharide lipids also showed some activity against Gram-positive bacteria, with the exception of the pathogenic strain *Staphylococcus aureus,* whereas Gram-negatives were either slightly or not inhibited at all (Vollbrecht et al., 1999, as cited in Franzetti et al., 2010b). Studies carried out to elucidate the molecular interactions between this biosurfactant and the lipidic component of the membrane showed that trehalose lipid increased the fluidity of phosphatidylethanolamine and phosphatidylserine membranes and formed domains in the fluid state and did not modify the macroscopic bilayer organization (Ortiz et al., 2008, 2009).

3.1.2 Antiviral activity

Antiviral effects have also been reported for surfactin and its analogues (Naruse et al., 1990). More effective inactivation of enveloped viruses, such as retroviruses and herpes viruses, were noted compared to non-enveloped viruses, suggesting that inhibitory action links may be due to physico-chemical interactions with the virus envelope (Vollenbroich et al., 1997a). Antiviral activity of some lipopeptides therefore may take place as a result of the viral lipid envelope and capsid disintegration due to ion channels formation, with consequent loss of the viral proteins involved in virus adsorption and/or penetration (Jung et al., 2000; Seydlová & Svobodová, 2008).

In vitro experiments showed that both surfactin and fengycin produced by *B. subtilis* fmbj were able to inactivate cell-free virus stocks of porcine parvovirus, pseudorabies virus, newcastle disease virus and bursal disease virus and could effectively inhibit infections and replication of these viruses (Huang et al., 2006).

Sophorolipids are also claimed to have activity against human immunodeficiency virus (Shah et al., 2005) and trehalose lipids (namely trehalose dimycolate, TDM) conferred higher resistance to intranasal infection by influenza virus in mice though inducing proliferation of T-lymphocytes bearing gamma/delta T-cell receptors, associated with the maintenance of acquired resistance to the infection (Hoq et al., 1997, as cited in Franzetti et al., 2010b).

Rhamnolipid alginate complex also showed significant antiviral activity against herpes simplex virus types 1 and 2. In particular, they significantly inhibited the herpesvirus cytopathic effect in the Madin-Darby bovine kidney cell line (Remichkova et al., 2008). The suppressive effect of the compounds on herpes simplex virus replication was dose-dependent and occurred at concentrations lower than the critical micelle concentration.

3.1.3 Anti mycoplasma activity

Some investigations have shown interesting anti-mycoplasma effects for surfactins. Mycoplasma contamination in cell culture is a frequently occurring serious limitation to biomedical research, particularly when it affects the irreplaceable cell lines which ultimately ends up destroyed. Earlier studies showed that surfactin treatment of mammalian cells that had been contaminated with mycoplasmas permitted specific inactivation of mycoplasmas without significantly damaging effects on cell metabolism in the culture (Vollenbroich et al., 1997b). In a more recent study, surfactin was used to eliminate mycoplasma from an extensively infected irreplaceable hybridoma cell line (Kumar et al., 2007). There were apparent indications of limited elimination, suggesting the possible use of surfactin in achieving total decontamination. However, it was observed that surfactin was toxic to the infected hybridoma cells plated at various cell densities and exposure times, therefore it was suggested that preliminary tests should be carried out to determine the cytotoxicity of surfactin before use in decontamination.

Another study confirmed surfactin potential to eliminate mycoplasma cells independently of the target cell, which is a significant advantage over the mode of action of conventional antibiotics (Fassi et al., 2007). This study also reported that surfactin exhibited a synergistic effect in combination with enrofloxacin, and resulted in mycoplasma-killing activity of about two orders of magnitude greater than when the molecules were used separately.

3.1.4 Antifungal activity

The antifungal activities of biosurfactants have long been known, although their action against human pathogenic fungi has been rarely described (Abalos et al., 2001; Chung et al., 2000; Tanaka et al., 1997). The previously mentioned cellobiose lipid flocculosin isolated from *Pseudozyma flocculosa*, was shown to display *in vitro* antifungal activity against several pathogenic yeasts, associated with human mycoses, including *Candida lusitaniae, Cryptococcus neoformans, Trichosporon asahii* and *Candida albicans* (Mimee et al., 2005). This product positively inhibited all pathogenic strains tested under acidic conditions and showed synergistic activity with amphotericin B. Moreover, no significant cytotoxicity was detected when tested against human cell lines. In nature, flocculosin is part of *P. flocculosa* biocontrol arsenal against other fungi. Recent reports however have suggested that flocculosin is also used by *P. flocculosa* as a nutrient source when experiencing food limitation and the molecule is rapidly deacylated under alkaline conditions losing its antimicrobial activity which may explain conflicting results concerning the antimicrobial activity of this class of glycolipids (Mimee et al., 2009).

Other antifungal activity of biosurfactants against phytopathogenic fungi has also been described. It has been recently demonstrated that glycolipids, such as cellobiose lipids (Kulakovskaya et al., 2009, 2010) and rhamnolipids (Debode et al., 2007, Banat et al, 2010) and cyclic lipopeptides (Tran et al., 2007, 2008, as cited in Banat et al, 2010), including surfactin, iturin and fengycin (Kim et al., 2010; Arguelles-Arias et al., 2009, Chen et al., 2009, Grover et al., 2010, Mohammadipour et al., 2009, Snook et al., 2009) can all have varying degrees of antimicrobial activities.

3.2 Antiadhesion activity of biosurfactants

Microbial biofilms formation on medical and technical equipment is an important and mostly hazardous occurrence, especially as the bacteria within such biofilms usually become highly resistant to antibiotics and adverse environmental challenges. Several approaches have been adopted in order to limit pathogen colonization. Strict hygienic practices by healthcare personnel such as hand washing and regular disinfection of equipment and environment become of grave importance. However, it should be noted that routine disinfection is becoming controversial as frequent application becomes less effective (Dettenkofer et al., 2004, 2007, Kramer et al., 2006, as cited in Falagas & Makris, 2009).

3.2.1 Biofilms on medical devices

Device-related infections are often identified as having a biofilm aetiology and biofilm formation can occasionally be facilitated by the host inflammatory response molecules which can make adhesion to the surface of the device easier (Hall-Stoodley et al., 2004). Almost all kind of surfaces are suitable to be colonized by biofilms (Donlan & Costerton, 2002). Biomedical devices are not the exception, biofilms are often found on the surface of urinary catheters (Stickler, 2008), central venous catheters (Petrelli et al., 2006), heart valves (Litzler et al., 2007), voice prostheses (Buijssen et al., 2007), contact lenses (Imamura et al., 2008), hip prostheses (Dempsey et al., 2007) and intrauterine devices (Chassot et al., 2008).

Current biofilm preventive strategies are essentially aimed at coating medical surfaces with antimicrobial agents, a process not always successful (Basak et al., 2009; von Eiff et al., 2005). Bacteria in biofilms become highly resistant to antibiotics, and so they evade host defenses withstanding antimicrobial chemotherapy (Morikawa, 2006). Since nosocomial infections remain an important problem even for hospitals with strict infection control programmes, infection control measures remain highly sought after (Falagas & Makris, 2009). Development of successful technologies based on the biofilm formation and growth control is expected to be a major breakthrough in the clinical practice and preventive medicine.

To eliminate biofilm formation novel compounds capable of specifically targeting biofilm growth while causing no adverse toxicity to the environment of application are needed. Several reports have suggested that, in addition to their direct action against pathogens, biosurfactants are able to interfere with biofilm formation, modulating microbial interaction with interfaces (Federle & Bassler 2003; Merk et al., 2005; Neu, 1996; Rasmussen & Givskov, 2006; Rodrigues et al., 2006b, 2006c; Rodrigues et al., 2007).

Surfactin for example has shown to be an important part of a list of biofilm controlling agents. Surfactin is able to inhibit biofilm formation of *Salmonella typhimurium, S. enterica*, *E. coli* and *Proteus mirabilis* in polyvinyl chloride wells, as well as vinyl urethral catheters (Mireles et al., 2001). Many *Salmonella* species are important opportunistic pathogens of the urinary tract system. Recently, two lipopeptide biosurfactants, produced by *B. subtilis* V9T14 and *B. licheniformis* V19T21, showed the ability to selectively inhibit biofilm formation by pathogenic strains on polystyrene (Rivardo et al., 2009). In particular, *S. aureus* ATCC 29213 and *E. coli* CFT073 biofilm formation were decreased by 97% and 90%, respectively. V9T14 biosurfactant was active on the Gram-negative strain yet ineffective against the Gram-positive while the opposite was observed for V19T21 biosurfactant. These effect were observed either by coating the polystyrene surface with these compounds or by adding the biosurfactant to the inoculum. The chemical characterization of V9T14 lipopeptide biosurfactant carried out by LC/ESI-MS/MS revealed that it was composed of 77% of surfactin and of 23% of fengycin (Pecci et al., 2010).

The activity of $AgNO_3$ combined with the lipopeptide biosurfactant V9T14 has also been studied against a preformed *E. coli* biofilm on the Calgary Biofilm device (Rivardo et al., 2010). Results indicated that the activity of silver can be synergistically enhanced by the presence of V9T14, allowing a reduction in the quantity of silver used to achieve greater antimicrobial impact. The concentrations of silver in the silver–biosurfactant solutions were 129 to 258 fold less than the concentrations needed when silver was used alone. In another study, the V9T14 biosurfactant in association with antibiotics led to a synergistic increase in the efficacy of antibiotics in *E. coli* CFT073 biofilm innhibition and, in some combinations, to total eradication of the uropathogenic strain biofilm (Rivardo et al., 2011).

In another recent work, marine bacterial culture supernatants of *Bacillus pumilus* and *B. indicus* significantly inhibited the initial attachment process and biofilm formation and dispersal of mature biofilms of *Vibrio* spp. strains (Nithya & Pandian, 2010). The bacterial supernatants also reduced the surface hydrophobicity of *Vibrio* spp. which is one of the important requirements for biofilm development. Valle et al., (2006) observed that distinct serotypes of group II capsular polysaccharides, produced by the uropathogenic *E. coli* (UPEC strain CFT073) behaved like surface-active polymers that displayed anti-adhesion properties. The treatment of abiotic surfaces with group II capsular polysaccharides drastically reduced both initial adhesion and biofilm development of important nosocomial pathogens.

More recently, the effect of different temperatures on the anti-adhesive activity of surfactin and rhamnolipid biosurfactants was tested on polystyrene surfaces, regarding the attachment of *Staphylococcus aureus*, *Listeria monocytogenes*, and *Micrococcus luteus* (Zeraik & Nitschke, 2010). Surfactin inhibited bacterial adhesion at all tested conditions, and its activity increased with the decrease in temperature, giving a 63–66% adhesion reduction in the bacterial strains at 4°C. Rhamnolipids promoted a slight decrease in the attachment of *S. aureus* but were not as effective. The ability of rhamnolipid biosurfactant to inhibit adhesion of microorganisms to silicone rubber was also investigated in a parallel-plate flow chamber (Rodrigues et al., 2006c). The results showed an effective reduction in the initial deposition rates and in the number of bacterial cells adhering after 4h, for several microorganisms. Moreover, perfusing the flow chamber with biosurfactant containing solution followed by the passage of a liquid–air interface produced high detachment (96%) of adhered cells for several microorganisms. These capabilites have a lot of implications regarding biofilm formation and microbial contamination and establishments on such biomedical devices made of such compounds.

Antibiofilm activity was also reported for a glycolipid biosurfactant isolated from another marine bacterium *Brevibacterium casei* MSA19 against pathogenic biofilms *in vitro* (Kiran et al., 2010a). The purified glycolipid disrupted the biofilm formation under dynamic conditions and the biofilm-forming capacity of both mixed culture and individual human and fish pathogenic strains was significantly inhibited at 30 mg/mL glycolipid. Raya et al., (2010) analyzed the effects of rhamnolipids and shear on initial attachment of *Pseudomonas aeruginosa* PAO1 in glass flow chambers. The presence of rhamnolipids significantly reduced the initial attachment of PAO1, even at the low concentration of 13 mg/L. Prewashing the cells with a 100 mg/L rhamnolipid solution, however, did not affect the attachment significantly. The initial cell attachment increased with increasing shear at the very low shear range (up to 3.5–5.0 mN/m^2), however the attachment could be minimized with further increase of the shear.

The biosurfactant Lunasan produced by the yeast *Candida sphaerica* UCP0995 completely inhibited the adhesion of *Streptococcus agalactiae*, *Streptococcus sanguis* 12, *Streptococcus mutans*, *Streptococcus mutans* NS, *Staphylococus epidermidis*, *Staphylococus aureus* and *Candida albicans* on plastic tissue culture plates at a concentration of 10 mg/ml and ≈92% inhibition of adhesion occurred for *Pseudomonas aeruginosa* (Luna et al., 2011). Lunasan, tested at the same concentration, also showed antimicrobial activity against the strains *Streptococcus oralis* (68%), *Candida albicans* (57%), and *Staphylococcus epidermidis* (57.6%). The same research group also described antiadhesive and antimicrobial activities of Rufisan, a biosurfactant produced by the yeast *Candida lipolytica* UCP 0988 (Rufino et al., 2011). Crude biosurfactant showed anti-adhesive activity at ≥0.75 mg/L against most of the microorganisms tested (*Staphylococus aureus*, *Streptococcus agalactiae*, *Streptococcus mutans* NS) and the anti-adhesive property was proportional to the concentration of the biosurfactant while antimicrobial activities were also observed at higher biosurfactant concentrations.

In conclusion, the anti-adhesive activity of biosurfactants against several pathogens indicates their potential utility as coating agents for medical insertional materials that may lead to a reduction in a large number of hospital infections without the need for use of synthetic drugs and chemicals.

3.2.2 Biofilms on food processing surfaces

In addition to the treatment of biomaterials used for medical devices, biosurfactants have also been used in the pre-treatment of material surfaces found in food-processing environments. Pathogenic bacteria implicated in food-borne illness outbreaks are able to form biofilms on food contact surfaces that are more resistant to sanitation than free-living cells (Kalmokoff et al., 2001; Kim et al., 2006; Stepanovic et al., 2004). The pre-conditioning of surfaces using microbial surface-active compounds may be an interesting strategy for preventing the adhesion of food-borne pathogens to solid surfaces. Meylheuc et al., (2006b) demonstrated that the preconditioning of stainless steel surfaces with an anionic biosurfactant produced by *Pseudomonas fluorescens* reduced the number of *L. monocytogenes* LO28-adhering cells and thus favoured the bactericidal activities of the disinfectants sodium hypochlorite (NaOCl) and peracetic acid/hydrogen peroxide (PAH).

Similarly, biosurfactants obtained from *Lactobacillus helveticus* and *P. fluorescens* were able to inhibit the adhesion of four *Listeria* strains to stainless steel (Meylheuc et al., 2006a). Whichever strain of *L. monocytogenes* used in combination with biosurfactants, the anti-adhesive biological coating developed both reduced the total adhering flora and the viable and culturable adherent bacteria on stainless steel surfaces. More recently, another group investigated the effect of rhamnolipid and surfactin biosurfactants on the adhesion of the food pathogens *E. sakazakii*, *L. monocytogenes* and *S. enteritidis* to polypropylene and stainless steel surfaces (Nitschke et al., 2009). Preconditioning with surfactin, rather than rhamnolipid, caused a reduction in the number of adhering cells particularly of *L. monocytogenes* and to some extent *E. sakazakii* on stainless steel. Surfactin showed a significant decrease in the adhesion on polypropylene of all strains. The adsorption of surfactin on polystyrene also reduced the adhesion of S. *enteritidis*- and *L. monocytogenes*-growing cells. In addition, surfactin was able to delay bacterial adhesion within short contact periods using non-growing cells or longer contact periods using growing cells.

Other antimicrobial and antiadhesive properties of a biosurfactant produced by *Lactobacillus paracasei* ssp. *paracasei* A20 isolated from a Portuguese dairy plant were also described (Gudiña et al., 2010). The biosurfactant had antimicrobial activity against a broad range of microorganisms including the pathogenic *C. albicans*, *E. coli*, *S. aureus*, *S. epidermidis* and *Streptococcus agalactiae* while exhibiting a considerable antiadhesive activity against a wide range of microorganisms.

The activity demonstrated by biosurfactants suggests that they could be considered as new tools in developing strategies to prevent or delay microbiological colonization of industrial plant surfaces used in foodstuffs preparation.

3.3 Probiotics biosurfactants activity

Probiotics are: “Live microorganisms which when administered in adequate amounts confer a health benefit on the host”. They have been reported to have positive effects on the maintenance of human health (Gupta & Garg 2009). Interest in probiotics has gained great significance due to the increasing antimicrobial resistance of bacteria worldwide. Evidence suggests that probiotic organisms may have a role in lowering the incidence or the duration of antibiotic-related diarrhea, contributing to the prevention or treatment of vaginal candidiasis, bacterial vaginosis and recurrent lower urinary tract infections. Furthermore, they encourage improved immunological defense responses and can decrease the activity of numerous toxic antimetabolites (Falagas et al., 2006a, 2006b, 2007).

Probiotics mechanisms of action vary, however, some are known to produce various antimicrobial agents such as organic acids, hydrogen peroxide, carbon peroxide, diacetyl, low molecular weight antimicrobial substances and bacteriocins (Merk et al., 2005). In addition, probiotics have long been known also for the capacity to interfere with the adhesion and formation of biofilms of pathogens to epithelial cells of urogenital and intestinal tracts (Reid et al., 1998, 2001). The mechanisms of this interference include the release of surface active molecules (Gudiña et al., 2010; Rodrigues et al., 2006d). Hong et al., (2005) reported the production of antimicrobial lipopeptides by *Bacillus* probiotics products as the main mechanisms by which they inhibit the growth of pathogenic microorganisms in the gastrointestinal tract. Similarly, competition with other microorganisms for adherence to epithelial cells as well as biosurfactants production are well known mechanisms used by *Lactobacillus* probiotics to interfere with vaginal pathogens (Barrons & Tassone, 2008; Cribby et al., 2008; Falagas et al., 2007).

Several investigators have pointed to evidence that probiotic type microorganisms and their biosurfactants may antagonize the growth of nosocomial pathogens on inanimate surfaces (Rodrigues et al., 2004a, 2004b, 2006b, 2006c; Walencka et al., 2008). Falagas & Makris, (2009) reviewed studies involving *in vitro* experiments on the potential role of probiotics microorganisms and their products in the inhibition of bacterial or fungal colonisation of artificial surfaces, such as vinyl urethral catheters and silicon rubber voice prostheses (Busscher et al., 1997, 1998; Velraeds et al., 1996, 1997, 2000; Rodrigues et al., 2004a, 2004b, 2006b, Van der Mei et al., 2000). The majority of the investigators examined the preconditioning of the materials surfaces with probiotic biosurfactants, while others added probiotic biosurfactant producing strains to examin adhesion or biofilm development. It was generally demonstrated that both probiotics microorganisms alone (mainly *Streptococcus thermophilus* and *Lactobacillus* spp. strains) or their biosurfactants were able to antagonize growth and development of potentially pathogenic microorganisms including *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus* spp., *Enterococcus faecalis, Candida albicans, Candida tropicalis* (Busscher et al., 1997; Van Hoogmoed et al., 2000).

Rodrigues et al., (2004a) demonstrated that the biosurfactant obtained from the probiotic bacterium *Lactococcus lactis* 53 was able to inhibit the adhesion of bacterial pathogens to silicone rubber with an adsorbed biosurfactant layer. Adhesion of yeasts was also decreased in the presence of biosurfactant, but to a lesser extent. In another work, using an artificial throat model, the same authors showed that biosurfactants obtained from probiotic strains greatly reduced microbial numbers on voice prostheses and induced a decrease in the airflow resistance of voice prostheses after biofilm formation, which may prolong the lifetime of indwelling silicone rubber voice prostheses (Rodrigues et al., 2004b).

In a more recent work, it was demonstrated that the preconditioning of silicon rubber with a biosurfactant produced by the strain *Streptococcus thermophilus* A reduced adhering bacterial pathogens by up to 97% and adhering *Candida* spp. by up to 70% (Rodrigues et al., 2006b). Velraeds et al., (1996) also reported on the inhibition of adhesion of pathogenic enteric bacteria by a biosurfactant produced by a *Lactobacillus* strain and later showed that the biosurfactant caused an important dose-related inhibition of the initial deposition rate of *E. coli* and other bacteria adherent on both hydrophobic and hydrophilic substrata (Velraeds et al., 1997).

Another interesting application area that is gaining increased interest relates to probiotics use in preventing oral infections (Çaglar et al., 2005; Hatakka et al., 2007; Kõll et al., 2008; Meurman, 2005; Meurman & Stamatova, 2007). Van Hoogmoed et al., (2004) demonstrated that *Streptococcus mitis* biosurfactant inhibited adhesion of *Streptococcus sobrinus* HG 1025 and *Streptococcus mutans* ATCC 25175 to bare enamel, while *S. mitis* biosurfactant was able to inhibit the adhesion of *S. sobrinus* HG 1025 to salivary pellicles. The authors later reported that these reductions may be attributed to increased electrostatic repulsion between the bacteria and the biosurfactant-coated pellicles (Van Hoogmoed et al., 2006).

New biosurfactant molecules produced by probiotic bacteria are reported from dairy products and environment. Recent work by Walencka et al., (2008) demonstrated that surfactants obtained from three *Lactobacillus acidophilus* strains inhibited *S. epidermidis* and *S. aureus* biofilm integrity and formation. Moreover, surfactant addition to preformed mature biofilms accelerated their dispersal and altered the characteristics of the biofilm morphology. A novel xylolipid biosurfactant from *Lactococcus lactis*, a probiotic strain isolated from a traditional Indian fermented dairy product, showed a good antibacterial activity against clinical pathogens of *E. coli* and MRSA strains (Saravanakumari & Mani, 2010). Xylolipid was non-pathogenic and safe for oral consumption and dermal applications, suggesting that it could be safely used as a therapeutic agent or as a preservative in food or cosmetic products.

In another recent work, a biosurfactant producing strain, *Lactobacillus* sp. CV8LAC, isolated from fresh cabbage, showed interesting antiadhesive activity against two *C. albicans* pathogenic biofilm-producing strains (CA-2894 and DSMZ 11225) (Fracchia et al., 2010). The CV8LAC biosurfactant significantly inhibited the adhesion of fungal pathogens to polystyrene microtiter plates in pre-coating and co-incubation experiments. In pre-coating assays, biofilm formation of strain CA-2894 was reduced by 82% at concentration of 312.5 µg/mL while that of strain DSMZ 11225 was reduced by 81% at 625 µg/mL. In co-incubation assays, biofilm formation of the two strains was inhibited by 70% at 160.5 µg/well and by 81% at 19.95 µg/well, respectively. It was interesting to note that no inhibition of both *C. albicans* planktonic cells was observed, thus indicating that the biosurfactant displayed specific anti-biofilm formation but not antimicrobial activity.

Considering their importance for human health and their recognized safety, environmental probiotic organisms may, thus, represent a safe and effective intervention for infection control purposes. Probiotics themselves or their products (biosurfactants), could be applied to patient care equipment, such as tubes or catheters, with the aim of decreasing the colonisation of these sites by nosocomial pathogens and potentially impede a central step in the pathogenesis of nosocomial infections (Falagas & Makris, 2009).

3.4 Other promising biological activities

Biosurfactants have been shown to have many other roles in biomedical application. Some of the most powerful molecules (eg. surfactin, mannosylerythritol lipids (MELs), trehalose lipids) are known to have anti-inflammatory, anti-tumour, immunosuppressive and immunomodulating functions, in addition to other properties such as self-assembling, human cells stimulation and differentiation, interaction with stratum corneum lipids, cell-to-cell signaling, hemolytic activity.

3.4.1 Anti-tumor activity

Recently, it has been demonstrated that these interesting microbial products can control a variety of mammalian cell functions. They are considered to participate in various intercellular molecular recognitions such as signal transduction, cell differentiation, cell immune response, etc. (Osada, 1998). Cao et al., (2010) demonstrated that surfactin induces apoptosis in human breast cancer MCF-7 cells through a ROS/JNK-mediated mitochondrial/caspase pathway. In a more recent work, they investigated the reactive oxygen species (ROS) and Ca^{2+} impact on mitochondria permeability transition pore (MPTP) activity, and MCF-7 cell apoptosis induced by surfactin (Cao et al., 2011). The results showed that surfactin initially induced the ROS formation, leading to the MPTP opening accompanied with the collapse of mitochondrial membrane potential which lead to an increase in the cytoplasmic Ca^{2+} concentration. In addition, cytochrome c was released from mitochondria to cytoplasm through the MPTP which activated caspase-9, eventually inducing apoptosis.

In another study, viscosin, an effective surface-active cyclic lipopeptide recovered from *Pseudomonas libanensis* M9-3, inhibited the migration of the metastatic prostate cancer cell line, PC-3M, without visible toxicity effects (Saini et al., 2008). More recently, lipopeptides (namely isoforms of surfactins and fengycins) derived from a marine *Bacillus circulans* DMS-

2 showed interesting cytotoxic activity against cancer cell lines (Sivapathasekaran et al., 2010). The purified lipopeptides at a concentration of 300 µg/mL showed more than 90% inhibition of proliferation on both colon cancer cell lines HCT 15 and HT 29 after 24 h treatment and the antiproliferative activity of lipopeptides was observed in a dose dependent manner.

Significant effects against both tumor cell lines as compared to non-tumor cell line were also observed, thus indicating the selective inhibitory activity of these molecules. Serratamolide AT514, cyclic depsipeptide from *Serratia marcescens,* belonging to the group of serrawettins, has also been reported to be a potent inducer of apoptosis of several cell lines derived from various human tumors and B-chronic lymphocytic leukemia cells, primarily involving the mitochondria-mediated apoptotic pathway and interference with Akt/NF-kB survival signals (Escobar-Díaz et al., 2005, as cited in Matsuyama et al., 2010). Biological studies of AT514 using human B-lymphocytes are now in progress for clinical applications of AT514 in the field of medical oncology.

Interesting anti-tumor activities has also been reported for glycolipids. Mannosylerythritol lipids (MELs) are among the most promising biosurfactants known due to their versatile interfacial and biochemical actions. Interesting studies, thoroughly reviewed by Kitamoto et al. (2002) and by Arutchelvi & Doble, (2010), have shown that MEL-A and MEL-B display excellent growth inhibition and differentiation-inducing activities against human leukemia cells including myelogenous leukemia cell K562, promyelocytic leukemia cell HL60, and the human basophilic leukemia cell line KU812, as well as growth inhibition activity of mouse melanoma B 16 cells. Recently Chen et al., (2006) also demonstrated that a sophorolipid produced from the yeast *Wickerhamiella domercqiae* induced apoptosis in H7402 human liver cancer cells by blocking cell cycle at G1 phase and partly at S phase, activating caspase-3, and increasing Ca^{2+} concentration in cytoplasm.

3.4.2 Anti-inflamatory activity

Byeon et al., (2008) observed that surfactin was able to downregulate LPS-induced nitric oxide production in RAW264.7 cells and primary macrophages by inhibiting NF-κB activation, suggesting a good potential as a bacterium-derived anti-inflammatory agent. Selvam et al., (2009) studied the effect of *B. subtilis* PB6, a natural probiotic, on plasma cytokine levels in inflammatory bowel disease and colon mucosal inflammation. The strain was found to secrete surfactins which are known to inhibit phospholipase A2, involved in the pathophysiology of inflammatory bowel disease. In animal experiments carried out in rat models for trinitrobenzene sulfonic acid-induced colitis, oral administration of PB6 as a probiotic suppressed colitis as measured by mortality rate and changes in colon morphology and weight gain. Plasma levels of pro-inflammatory cytokines were also significantly lowered and the anti-inflammatory cytokine significantly increased after the oral administration of PB6, supporting the concept that PB6 inhibits PLA2 by secreting surfactins.

In another work, surfactin isomers derived from the mangrove bacterium *Bacillus* sp. (No. 061341) showed interesting anti-inflammatory activities (Tang et al., 2010). In particular, this class of cyclic lipopeptide showed strong inhibitory properties on the overproduction of nitric oxide and the release of IL-6 in LPS-induced murine macrophage cell RAW264.7. Moreover, structure-activity relationship studies revealed that the existence of the free carboxyl group in the structure of surfactin isomer was crucial as to the anti-inflammatory activities. An interesting recent study explored the mechanisms responsible for surfactin-induced anti-inflammatory actions in the context of periodontitis caused by *Porphyromonas* *gingivalis,* the major pathogen of periodontal disease (Park et al., 2010). The Authors observed that surfactin significantly reduces pro-inflammatory cytokines, including tumor necrosis factor-α, interleukin (IL)-1β, IL-6, and IL-12, through suppression of nuclear factor-κB activity in *P. gingivalis* LPS-stimulated THP-1 human macrophage cells, in a Heme oxygenase-1 (HO-1)-dependent fashion. Furthermore, surfactin treatment effectively induces HO-1 expression, a major defense in response to oxidative stress.

These observations support the potential of surfactin as a candidate in strategies to prevent caries, periodontitis, or other inflammatory diseases.

3.4.3 Immuno-modulatory action

Park & Kim, (2009) studied the role of surfactin in the inhibition of the immunostimulatory function of macrophages through blocking the NK-κB, MAPK and Akt pathway. This provided a new insight into the immunopharmacological role of surfactin in autoimmune disease and transplantation. Their work indicated that surfactin has potent immunosuppressive capabilities which suggested important therapeutic implications for transplantation and autoimmune diseases, including allergy, arthritis and diabetes.

A biosurfactant glycolipid complex from *Rhodococcus ruber* was also shown to activate the production of IL-1beta and TNF-alpha cytokines without modifying the production of IL-6, thus suggesting good prospects for further studies of immunomodulating and antitumor activities (Kuyukina et al., 2007).

3.4.4 Other biomedical related properties

Han et al., (2008) observed that high surfactin micelle concentration affected the aggregation of amyloid β-peptide (Aβ (1-40)) into fibrils, a key pathological process associated with Alzheimer's disease. Fengycin, another lipopeptide biosurfactant is also able to cause membrane perturbations (Deleu et al., 2008). Recent results by Eeman et al., (2009) emphasized the ability of fengycin to interact with the lipid constituents of the stratum corneum extracellular matrix and with cholesterol. Another interesting property of surfactin and its synthetic analogues is the ability to alter the nanoscale organisation of supported bilayers and to induce nanoripple structures with intriguing perspectives for biomedical and biotechnological applications (Bouffioux et al., 2007; Brasseur et al., 2007; Francius et al., 2008).

Morita et al., (2010) investigated the cell activating property of MELs using cultured fibroblast and papilla cells, and a three dimensional cultured human skin model. The di-acetylated MEL (MEL-A) produced from soybean oil significantly increased the viability of the fibroblast and of the papilla cells over 150% compared with that of control cells, suggesting potential use as new hair growth agent stimulating the papilla cells. Using a three-dimensional cultured human skin model, Morita et al., (2009b) observed that the viability of the SDS damaged cells was markedly improved by the addition of MEL-A in a dose-dependent manner. This demonstrated that MEL-A also had a ceramide-like moisturising activity toward the skin cells. Similarly, (Kitagawa et al., 2007, as cited in Worakitkanchanakul et al., 2008) reported that MEL-B shows excellent moisturizing properties, equivalent to those of natural ceramides, toward human skin.

Trehalose lipids also display various interesting biological activities mainly due to their great tendency to partition into phospholipid membranes (Ortiz et al., 2008, 2009). In particular, the trehalose lipid was suggested to incorporate into the membrane bilayers and produce structural perturbations, which might affect the function of both phosphatidylethanolamine and phosphatidylserine membranes. Zaragoza et al., (2010) observed that a succinoyl trehalose lipid produced by *Rhodococcus* sp. caused the swelling of human erythrocytes followed by hemolysis at concentrations well below its critical micellar concentration. They concluded that trehalose lipid caused the hemolysis of human erythrocytes by a colloid-osmotic mechanism, most likely by formation of enhanced permeability domains, or "pores" enriched with biosurfactant, within the erythrocyte membrane.

Permealization of biological and artificial membranes was also reported to be induced by *Pseudomonas aeruginosa* dirhamnolipid (Sánchez et al., 2010). In particular, it caused the hemolysis of human erythrocytes through a lytic mechanism, as shown by the similar rates of K^+ and hemoglobin leakage, and by the absence of effect of osmotic protectants. Scanning electron microscopy showed that the addition of the biosurfactant changed the usual disc shape of erythrocytes into that of spheroechinocytes.

4. Biotechnological and nanotechnological applications of surface-active compounds

Biosurfactants, have been increasingly attracting attention in the field of nanotechnology (Kitamoto et al., 2005, 2009). During the last decade, unique properties of biosurfactants, like versatile self-assembling and biochemical properties, which are not usually observed in conventional chemical surfactants, have been reported (Kitamoto et al., 2005, 2009). In recent years, the development of new functional structures and/or systems using self-assembly of amphiphilic molecules has evolved into a dynamic and rapidly growing area of nanotechnology (Ariga et al., 2007, Shimizu et al., 2005, as cited in Kitamoto et al., 2009) due to their ability to self-assemble into hierarchically ordered structures using hydrogen bonding, hydrophobic and van der Waals interactions as mentiond earlier.

Mannosylerythritol lipids (MELs) show the most interesting self-assembling properties and numerous related potential applications (Kitamoto et al., 2009). Konishi et al., (2007), Imura et al., (2007, 2008), and Ito et al., (2007), for example, developed and studied the kinetics of interactions in carbohydrate ligand systems composed of self-assembled monolayers of mannosylerythritol lipid-A (MEL-A) serving as a high-affinity, easy to handle and low-cost ligand system for immunoglobulin G and M and lectins..

Table 1 below lists the latest discoveries in the biotechnological and nanotechnological fields applicable to biosurfactants, and in particular the latest successful results of mannosylerythritol lipids (MELs) application in the enhancement of the gene transfection efficiency of cationic liposomes as well as some interesting applications of glycolipids and other biosurfactants in drug delivery and gene therapy. Biosurfactants use as a "green" alternative for high-performance nanomaterials production and, in particular, for the synthesis and stabilization of metal-bound nanoparticles will also be described.

Biosurfactant type	Activity/application	Study
Mannosylerythritol lipids-A	Ligand system for immunoglobulin G and M and lectins	Konishi et al., (2007); Imura et al., (2007, 2008), Ito et al., (2007)
	DNA capsulation and membrane fusion with anionic liposomes	Ueno et al., (2007a)
	In vitro promotion of gene transfection mediated by cationic liposomes	Inoh et al., (2001, 2004, 2010); Igarashi et al., (2006); Ueno et al., (2007b)

Biosurfactant type	Activity/application	Study
	In vivo promotion of liposome-mediated gene transfection	Inoh et al., (2009)
	Herpes simplex virus thymidine kinase gene therapy	Maitani et al., (2006)
	Water-in-oil microemulsions	Worakitkanchanakul et al., (2008)
	Increase membrane fluidity of monolayers composed of L-α-dipalmitoylphosphatidylcholine (DPPC)	Kitamoto et al., (2009)
Mannosylerythritol lipids-B	Self-assembling and vesicle-forming activity	Worakitkanchanakul et al., (2008)
Rhamnolipids and sophorolipids	Deuterated rhamnolipids and sophorolipids	Smyth et al., (2010b)
	Cadmium sulfide nanoparticles	Singh et al., (2011)
	Biocompatible microemulsions of lecithin/rhamnolipid/sophorolipid biosurfactants	Nguyen et al., (2010)
Rhamnolipids	Silver nanoparticles with antibioticmicrobial activity	Kumar et al., (2010)
	Nickel oxide nanoparticles by microemulsion technique	Palanisamy & Raichur, (2009)
	Silver nanoparticles	Xie et al., (2006)
	ZnS nanoparticles	Narayanan et al., (2010)
	Microemulsions	Xie et al., (2005, 2007)
	Alcohol-free microemulsions	Nguyen & Sabatini, (2009)
Sophorolipids	Cobalt nanoparticles	Kasture et al., (2007)
	Silver nanoparticles	Kasture et al., (2008)
	Sophorolipid-coated silver and gold nanoparticles with antibacterial activity	Singh et al., (2009, 2010)
	Biocompatible microemulsions of lecithin/rhamnolipid/sophorolipid biosurfactants	Nguyen et al., (2010)
Glycolipid biosurfactant	Silver nanoparticles	Kiran et al., (2010b)
Fengycin and surfactin	Enhancers for the skin accumulation of aciclovir	Nicoli et al., (2010)
Surfactin	Surfactin-mediated synthesis of gold nanoparticles	Reddy et al., (2009)
	Cadmium sulfide nanoparticles	Singh et al., (2011)

Table 1. Examples of recent biosurfactant applications in the biotechnological and nanotechnological fields.

4.1 Liposomes and gene transfection

Gene transfection into the cells is a fundamental technology not only for molecular and cellular biology processes but also a clinical gene therapy (Ueno et al., 2007b). Although several methods for gene transfection have been investigated (Felgner et al., 1989, Fujiwara, 2000, Gao & Huang, 1991, Hatakeyama et al., 2007, Nishiyama et al., 2005, Ueno et al., 2007b), more efficient and safe systems are still needed (Ueno et al., 2007b). Among the various methods, lipofection using cationic liposomes is considered to be a promising method for introducing foreign gene to the targeted cells due to their high transfection efficiency, low toxicity and immunogenicity, ease of preparation and targeted application (Farhood et al., 1992, Felgner et al., 1989, Kogure et al., 2007, Lasic, 1998, Nakanishi, 2003, Inoh et al., 2010). The physicochemical properties of cationic liposomes, such as lipid packing density, shape, and zeta-potential, have a significant effect on gene transfection efficiency (Lin et al., 2003, Takeuchi et al., 1996, Wittenberg et al., 2008, Xu et al., 1999, as cited by Inoh et al., 2010).

Inoh et al., (2001) reported that MEL-A promoted DNA transfection efficiency mediated by cationic liposomes. Confocal laser scanning microscopic analysis showed the distribution of lipids and oligonucleotide DNA in MEL-A-containing liposome–DNA complex in the plasma membrane and the nucleus of target cells at 1 h after the addition of complex (Inoh et al., 2004). This suggests that MEL-A induces the membrane fusion between the target cells and the cationic liposomes, accelerating the efficiency of gene transfection significantly. Similarly, Igarashi et al., (2006) reported that MEL-A significantly increased the cellular association and the efficiency of gene transfection mediated by cationic liposomes in human cervix carcinoma Hela cells. Analysis of flow cytometric profiles clearly indicated that the amount of DNA associated with the cells was rapidly increased and sustained by addition of MEL-A to the liposome. Confocal microscopic observation also indicated that the MEL-lipoplex distributed widely in the cytoplasm and DNA presence was intensely detected in cytoplasm around the nucleus.

The above results suggested that MEL-A increased gene expression by enhancing the association of the lipoplexes with the cells in serum and, thus, MEL-liposome may prove a significant nonviral vector for gene transfection and gene therapy.

In an attempt to explain how MEL-A-containing liposomes could accelerate gene transfection, Ueno et al., (2007a) examined MEL-containing liposomes properties such as their activity for DNA capsulation and membrane fusion abilities of cationic liposomes with artificial anionic liposomes. They observed that MEL-A-containing liposomes exhibited high activity in DNA incapsulation and membrane fusion with anionic liposomes, which are important properties for gene transfection. On the other hand, MEL-B- and MEL-C-containing liposomes only increased either the incapsulation or the membrane fusion. Ueno et al., (2007b) further examined the mechanism of the transfection mediated by cationic liposomes with NBD-conjugated MEL-A and reported that MEL-A distributed on the intracellular membranes through the plasma membranes of target cells, while the cationic liposomes with MEL-A fused to the plasma membranes within 20–35 min. Thereafter, they noted that the oligonucleotide released from the vesicles was immediately transferred to the nucleus. They therefore suggested that MEL-A was capable of promoting the transfection efficiency of target cells by inducing membrane fusion between liposomes and the plasma membrane of these cells.

Recently Kitamoto et al., (2009) demonstrated that monolayers composed of L-α-dipalmitoylphosphatidylcholine (DPPC) containing MEL-A had greater membrane fluidity than those containing only DPPC. It was also reported that unsaturated fatty acids in MEL-A significantly influenced surface pressure and packing density in the monolayer and thus the physicochemical properties of MEL-A and MEL-A/lipids (Imura et al., 2008). Transfection efficiency of nano vectors with MEL-A was investigated *in vivo* on tumor cells in the mouse abdominal cavity (Inoh et al., 2009). When a complex of the nano vectors with MEL-A and plasmid DNA was injected intraperitoneally into C57BL/6J mice bearing B16/BL6 tumors, the biosurfactant significantly increased liposome-mediated gene transfection to the mouse tumor cells. The transfection efficiency of the plasmids into the solid tumors by the cationic liposomes of cholesteryl-3beta-carboxyamidoethylene-N-hydroxyethylamine (OH-Chol) with MEL-A increased by approximatley 100-fold compared to that by the commercially available DC-Chol cationic liposomes without MEL-A. This suggests that nonviral vectors with MEL-A are very useful for gene transfection *in vivo*. The mechanisms of gene delivery by nano vectors with MEL-A and the numerous biological activities of these biosurfactants have been described by Nakanishi et al., (2009) and Kitamoto et al., (2009).

Inoh et al., (2010) further investigated the effects of unsaturated fatty acid ratio within the MEL-A compound on the physicochemical properties and gene delivery into cells of cationic liposomes using MEL-A with three different unsaturated fatty acid (USF) component ratios. Gene transfer efficiency of cationic liposomes containing MEL-A (containing 21.5% USF) was much higher than that of those containing MEL-A (containing 9.1%USF) and MEL-A (containing 46.3%USF). In particular, MEL-A (21.5% USF)-containing cationic liposomes induced highly efficient membrane fusion after addition of anionic liposomes and led to subsequent DNA release.

Imaging analysis revealed that MEL-A (21.5% USF)-containing liposomes fused with the plasma membrane and delivered DNA into the nucleus of NIH-3T3 cells, MEL-A (46.3% USF)-containing liposomes fused with the plasma membrane did not deliver DNA into the nucleus, and MEL-A (9.1% USF)-containing liposomes neither fused with the plasma membrane nor delivered DNA into the nucleus. These results suggest that the MEL-A unsaturated fatty acid ratio significantly affects transfection efficiency due to changes in membrane fusion activity and the efficiency of DNA release from the liposomes.

Mannosylerythritol lipid-B (MEL-B) with a different configuration of the erythritol moiety was found to self-assemble into a lamellar phase over remarkably wide concentration and temperature ranges; furthermore it showed great potential as a vesicle-forming lipid, suggesting its potential application in drug and gene delivery as well as in transdermal delivery systems (Worakitkanchanakul et al., 2008). In another work, a liposome vector containing betasitosterol beta-D-glucoside biosurfactant-complexed DNA was successfully used for herpes simplex virus thymidine kinase gene therapy (Maitani et al., 2006).

4.2 Biosurfactants potential in drug delivery

Properties such as detergency, emulsification, foaming and dispersion make biosurfactants interesting molecules with potential application in the field of drug delivery (Faivre & Rosilio, 2010). MEL-A for example has much higher emulsifying activity with soybean oil and tetradecane than polysorbate 80 (Kitamoto et al., 2009) and is able to form stable water-in-oil microemulsions without addition of co-surfactant or salt (Worakitkanchanakul et al., 2008).

Rhamnolipids and sophorolipids have also been mixed with lecithins to prepare biocompatible microemulsions in which the phase behavior was unaffected by changes in temperature and electrolyte concentration, making them desirable for cosmetic and drug delivery applications (Nguyen et al., 2010). In 1988, rhamnolipid liposomes were patented as drug delivery systems, useful as microcapsules for drugs, proteins, nucleic acids, dyes and other compounds, as biomimetic models for biological membranes and as sensors for detecting pH variations. These novel liposomes were described as safe and biologically decomposable, with suitable affinity for biological organisms, stable and with long service and shelf life.

The potential of lipopeptides, fengycin and surfactin to act as enhancers for the transdermal penetration and skin accumulation of aciclovir was also recently investigated (Nicoli et al., 2010) to elucidate any possible synergistic effect between surfactin and fengycin associated with anodal iontophoresis. It was demonstrated that these lipopeptides did not enhance aciclovir transport across the skin (not even when associated with iontophoresis) although they increased aciclovir concentration in the epidermis by a factor of 2 (Nicoli et al., 2010).

Microemulsion produced using biosurfactant are thermodynamically stable and their isotropic systems that form spontaneously-consisting of microdomains of oil or water stabilized by an interfacial film - in addition to their long-term stability, easy preparation and high solubilization capacity are considered to be very promising liquid vehicles for future drug delivery systems (Date et al., 2008, as cited in Faivre & Rosilio, 2010).

4.3 Nanoparticles

Another interesting application for natural surfactant is the the synthesis of metal-bound nanoparticles as an alternative environmentally friendly technology (Sharma et al., 2009). Nanomaterials synthesis and use has been an active research area due to interesting properties of the nanomaterials as compared to bulk material use (Palanisamy & Raichur, 2009). Metal nanoparticles have beening explored in various fields such as catalysis, mechano- and electrical applications and biomedical uses (Van Bogaert & Soetaert, 2010). The reduction in size gives rise to size dependent effects such as high surface to volume ratio, lower melting point, changes in electronic structure and changes in lattice structure and interatomic distances which in turn affect the processing parameters (Liveri, 2006, as cited in Palanisamy & Raichur, 2009).

The use of gold nanoparticles, in particular, is currently undergoing a dramatic expansion in the field of drug and gene delivery, targeted therapy and imaging technologies (Boisselier & Astruc, 2009; Pissuwan et al., 2009, 2011). Potential therapeutic applications of gold compound and gold nanoparticles also include anti-HIV activity, anti-angiogenesis, anti-malarial activity, anti-arthritic activity and biohydrogen production (Kalishwaralal et al., 2010). Silver nanoparticles are also been reported to possess anti-fungal activity, anti-inflammatory effect, anti-viral, anti-angiogenesis and anti-platelet activity (Kalishwaralal et al., 2010).

Reddy et al., (2009) successfully synthesized surfactin-mediated gold nanoparticles and investigated the effects of proton concentrations and temperature on the morphology of the obtained nanoparticles. It was demonstrated that the nanoparticles synthesized at pH 7 and 9 remained stable for 2 months, while aggregates were observed at pH 5 within 24 h. Moreover, the nanoparticles formed at pH 7 were uniform in shape and size and were polydispersed and anisotropic at pH 5 and 9. The nanoparticles synthesized produced at room temperature were monodispersed and were more uniform when compared to those formed at 4°C. More recently they also carried out a biological synthesis of gold and silver nanoparticles using the bacteria *Bacillus subtilis.* Gold nanoparticles were synthesized both intra- and extracellularly, while silver nanoparticles were exclusively formed extracellularly (Reddy et al., 2010). According to the Authors the nanoparticles were stabilized by the surface-active molecules i.e., surfactin or other biomolecules released into the solution by *B. subtilis*.

Surfactin produced by *Bacillus amyloliquefaciens* KSU-109 was also used for the synthesis of cadmium sulfide nanoparticles which remained stable up to six months without compromising their functionality (Singh et al., 2011). This kind of nanoparticles works as semiconductors with unique optical properties and tunable photo-luminescence allowing potential applications in solar energy conversion, nonlinear optical, photoelectrochemical cells and heterogeneous photocatalysis (Singh et al., 2011). In addition, surfactin produced by strain KSU-109 was easily extracted and used without further purification for nanoparticles stabilization under ambient conditions (Singh et al., 2011). Such simple, inexpensive and environmental friendly procedure of obtaining surfactin offers a further advantage of use in nanobiotechnology for the large-scale production of highly stable metal nanoparticles.

Both rhamnolipids and sophorolipids have also been successfully used for the synthesis and stabilization of metal-bound nanoparticles. Purified rhamnolipids from *P. aeruginosa* strain BS-161R were used to synthesize silver nanoparticles which exhibited good antibiotic activity against both Gram-positive and Gram-negative pathogens and *Candida albicans*, suggesting their broad spectrum antimicrobial activity (Kumar et al., 2010). In another work, a glycolipid biosurfactant produced from sponge-associated marine bacteria *Brevibacterium casei* MSA19, using agro-industrial and industrial waste as substrate, were used as a "green" stabilizer for the synthesis of stable and uniform silver nanoparticles (Kiran et al., 2010b). The biosurfactant acted as stabilization agent and prevented the formation of aggregates.

Palanisamy & Raichur, (2009) also described a simple and eco-friendly method for synthesizing spherical nickel oxide nanoparticles by microemulsion technique using rhamnolipids as alternative surfactant. The synthesized nanoparticles were found to be fully crystalline and spherical in shape with uniform distribution and increasing the pH of the solution decreased the size of the nanoparticles. Xie et al., (2006) were also able to synthesize silver nanoparticles in rhamnolipid reverse micelles while in another study rhamnolipids were used as capping agents for the synthesis of ZnS nanoparticles in aqueous medium (Narayanan et al., 2010).

Sophorolipids were also tested for use in nanoparticles synthesis and reported to be good reducing and capping agents for cobalt and silver particles (Kasture et al., 2007, 2008, as cited in Van Bogaert & Soetaert, 2010). Singh et al., (2009) demonstrated the antibacterial activity of sophorolipid-coated silver and gold nanoparticles against both Gram-positive and –negative bacteria. They also verified that sophorolipid-coated gold nanoparticles were more cyto and geno-compatible with respect to silver nanoparticles (Singh et al., 2010). They also plan to investigate these nanoparticles suitability for medical and diagnostic applications.

Recently, methodologies for the biological synthesis of metal nanoparticles using microbes have also been described (Narayanan & Sakthivel, 2010; Kalishwaralal et al., 2010; Reddy et al., 2010). In addition Smyth et al., (2010b) reported on the production of selectively deuterated rhamnolipids and sophorolipids using deuterated substrates. The production of such deuterated biosurfactants, in particular, or other bioactive microbial products in general, in which distinct pattern of labeling could be achieved resulting in varing molecular weight products and or stereochemistry unrecognised by existing degradative enzymes is very improtant. Such molecules would have great future implications with regards to efficacy and/or persistence or the development of resistance for some bioactives particularly in biomedical related applications.

4.4 Microemulsions

Microemulsions are thermodynamically stable, isotropic dispersions of oil, water and surfactant (Rosen, 1989, as cited in Nguyen et al., 2010). Microemulsion systems produce high solubilization capacity and ultralow interfacial tensions of oil and water, making them desirable in practical applications such as enhanced oil recovery, drug delivery, food and cosmetic applications (Bourrel & Schechter, 1988, Kogan & Garti, 2006, Komesvarakul et al., 2006, Lawrence & Rees. 2000, Vandamme, 2002, Yuan et al., 2008, as cited in Nguyen et al., 2010). Xie et al., 2005 demonstrated that rhamnolipids could be successfully used to form microemulsions using medium chain alcohols as cosurfactant. Subsequently, the same Authors observed that the phase behavior and microstructure of these microemulsions were rational to the conformational changes of rhamnolipid molecules at the interface of oil/water (Xie et al., 2007). Microemulsion technique using oil–water–surfactant mixture has also emerged as a promising method for nanoparticle synthesis and can be used to synthesize different types of particles (Eastoe et al., 2006, as cited in Palanisamy & Raichur, 2009). Palanisamy & Raichur, (2009), for example, successfully used rhamnolipids as the surfactant to synthesize spherical nickel oxide nanoparticles by microemulsion technique. In another work, Nguyen & Sabatini, (2009) were able to formulate alcohol-free microemulsions using rhamnolipid biosurfactant and rhamnolipid mixtures.

Lecithin-based microemulsions have proven to be desirable in biocompatible formulations due to their tendency to mimic the phospholipid nature of cell membranes (Nguyen et al., 2010). In a recent report Nguyen et al., (2010) formulated and evaluated microemulsions of lecithin/rhamnolipid/sophorolipid biosurfactants with a range of oils. Sophorolipid played an important role as the hydrophobic component in these formulations and the phase behavior of these biocompatible microemulsions did not change significantly with changing temperature and electrolyte concentration, making them desirable for cosmetic and drug delivery applications.

4.5 A survey over biotechnological commercial applications and patents of biosurfactants and bioemulsifiers

Due to their broad-range of functional properties and the diverse synthetic capabilities of microbes, biological surfactants and emulsifiers have been recently used in various industries like detergents and soaps, petroleum, textile, agriculture, cosmetic, medicine and food (Banat et al., 2000, 2010). Due to their environmental acceptability, biodegradability and lower toxicity, they are generally accepted as good candidates to substitute synthetic surfactants. Commercial applications of biosurfactants and bioemulsifiers in the biotechnological field are mainly related to the oil industry, enhanced oil recovery and bioremediation technologies (Desai & Banat, 1997). However, interesting marketable products and patents have been issued in the last few years in the health care and cosmetic industries, reviewed by Shete et al., (2006) and Banat et al., (2010).

Sugar-based biosurfactants, sophorolipids in particular, are very attractive in these fields, because of their good detergency, emulsifying, foaming and dispersing properties (Faivre & Rosilio, 2010). Sophorolipids are better solubilizers than emulsifiers, but their derivatives containing propylene glycol have excellent hygroscopic properties and are applied as moisturizer or softener in cosmetic products (Faivre & Rosilio, 2010). For example, a product containing 1 mol of sophorolipid and 12 mol of propylene glycol has excellent skin compatibility and is used commercially as a skin moisturizer (Yamane, 1987, as cited in Desai & Banat 1997). Sophorolipid is commercially used by Kao Co. Ltd. as a humectant for cosmetic makeup brands such as Sofina. This company has developed a fermentation process for sophorolipid production, and after a two-step esterification process, the product finds application in lipstick and as moisturizer for skin and hair products (Inoue et al., 1979 a, 1979b, as cited in Desai & Banat, 1997). Moreover, sophorolipids are also believed to stimulate the leptin synthesis through adipocytes, in this way reducing the subcutaneous fat overload (Pellecier & André, 2004, as cited in Van Bogaert & Soetaert, 2010).

The French company Soliance (http://www.groupesoliance.com) produces sophorolipid-based cosmetics for the body and skin and the Korean MG Intobio Co. commercializes Sopholine cosmetics (Van Bogaert & Soetaert, 2010). They are also found in cleaning soap mixtures (Ecover™ products). Despite the high number of scientific publications and patents, industrial surfactin applications still remain quite limited (Jacques, 2010). Sold by SIGMA and SHOWA DENKO for analytical or laboratory purposes, the compound is also available in several Japanese cosmetic products.

During the last decades, many patents have been issued worldwide in relation with applications of biosurfactants and bioemulsifiers in the health care field (Shete et al., 2006). Bioemulsifiers produced by *Acinetobacter calcoaceticus*, for examples, have been used in shampoos and soaps against acne and eczema and in personal care products. The skin cleansing cream and lotion containing these bioemulsifiers have, among other properties, the ability to interfere with microbial adhesion on skin or hair (Hayes et al., 1989, 1990, 1991, 1992, as cited in Shete et al., 2006). Viscosin and analogues have been patented as antibacterial, antiviral, antitrypanosomal therapeutic compounds that inhibit the growth of *Mycobacterium tuberculosis*, Herpes simplex virus 2 and/or *Trypanosoma cruzi* (Burke et al., 1999, as cited in Shete et al., 2006). *Lactobacillus* biosurfactants have also been patented as inhibitors of adherence and colonization of bacterial pathogens on medical devices (Reid et al., 2000).

Another interesting patented area is related to antimicrobial biosurfactant peptides produced by probiotic strains able to selectively bind to collagen and inhibit infections around wounds at the site of implants and biofilms associated with infections in mammals (Howard et al., 2002, as cited in Shete et al., 2006). Sophorolipids, in particular have been the object of many patents as moisturizing agents and for the amelioration of skin physiology, skin restructuration and repair (Shete et al., 2006). Sophorolipids are also used for the treatment of skin, as an activator of macrophages, and as agent in fibrinolytic healing, desquamating and depigmenting process (Maingault, 1999 as cited in Shete et al., 2006). A germicidal composition containing fruit acids, a surfactant and a sophorolipid biosurfactant, able to kill in 30 seconds 100% of *E. coli*, *Salmonella* and *Shigella*, has been patented for cleaning fruits, vegetable, skin and hair (Pierce & Heilman, 2001).

Rhamnolipids in comparison have been patented in a process to make some liposomes and emulsions (Ishigami & Suzuki 1997; Ramisse et al., 2000) both important in the cosmetic industry. More recently an activator and anti-aging agent containing MEL as active ingredient has been patented (Suzuki et al., 2010). Another recent invention is directed to polymeric acylated biosurfactants that can self-assemble or auto-aggregate into polymeric micellar structures useful in topically-applied dermatologic products (Owen & Fan, 2010).

Another patent has been deposited about a biosurfactant composition produced by a new *B. licheniformis* strain, with anti-adhesion activity against biofilm producer microbial pathogens (Martinotti et al., 2009).

5. Conclusions and perspectives

As evidenced by the growing number of publications on the topic of biosurfactants, there is an increasing interest in the study of these molecules and their potential applications. The demand for new specialty surfactants in the agriculture, cosmetic, food, pharmaceutical, and environmental industries is steadily increasing and biosurfactants, as effective and environmentally compatible compounds, perfectly meet this demand (Banat et al., 2000, 2010; Mukherjee et al., 2006).

The most important limitation for the commercial use of biosurfactants is the complexity and high cost of production, which has limited the development of their use on a large scale (Soberón-Chávez & Maier, 2010). However, the proven antimicrobial, anti-adhesive, immune-modulating properties of biosurfactants and the recent successful applications in gene therapy, immunotherapy and medical insertion safety suggest that it is worth persisting in this field. Moreover, in pharmaceutical and biomedical sectors, the high cost of production could be compensated for by the small amounts of product required. In fact, it has been elucidated that biosurfactants used as pharmaceutical agents are needed only in very low concentrations (Cameotra & Makkar, 2004). Prerequisites for making biosurfactant production more profitable and economically feasible include optimized growth/production conditions and novel and efficient multi-step downstream processing methods as well as the use of recombinant varieties of microorganisms or selected hyperproducing mutants, which can grow on a wide range of cheap renewable substrates (Muthusamy et al., 2008).

Recent advances in the area of biomedical application are probably going to take the lead due to higher potential economic returns. Moreover, due to their self-assembly properties, new and fascinating applications in nanotechnology are predicted for biosurfactants (Kitamoto et al., 2009; Palanisamy, 2008; Reddy et al., 2009). In-depth studies of their natural roles in microbial competitive interactions, cell-to-cell communication, pathogenesis, motility and biofilm formation and maintenance could suggest improved and interesting future applications.

6. Acknowledgements

This work was partially supported by the Local Research funding of the Italian *Ministero dell'Istruzione, dell'Università e della Ricerca*.

7. References

Abalos A, Pinazo A, Infante MR, Casals M, García F & Manresa A (2001) Physicochemical and antimicrobial properties of new rhamnolipids produced by *Pseudomonas aeruginosa* AT10 from soybean oil refinery wastes. *Langmuir.* 17:1367–1371

Arima K, Kakinuma A & Tamura G (1968) Surfactin, a crystalline peptide-lipid surfactant produced by *Bacillus subtilis*: isolation, characterization and its inhibition of fibrin clot formation. *Biochem Biophys Res Commun.* 31:488–494

Arutchelvi J & Doble M (2010) Mannosylerythritol lipids: microbial production and their applications, In: *Biosurfactants: From Genes to Applications,* Soberón-Chávez G Ed., pp. 145-177, Springer, Münster, Germany

Arutchelvi JI, Bhaduri S, Uppara PV & Doble M (2008) Mannosylerythritol lipids: a review. *J Ind Microbiol Biotechnol.* 35:1559–1570

Banat IM (1995a) Biosurfactants production and use in microbial enhanced oil recovery and pollution remediation: A review. *Bioresource Technol.* 51:1-12

Banat IM (1995b) Biosurfactants characterization and use in pollution removal; state of the art. A review. *ACTA Biotechnologica.* 15:251-26

Banat IM, Franzetti A, Gandolfi I, Bestetti G, Martinotti MG, Fracchia L, Smyth TJ & Marchant R (2010) Microbial biosurfactants production, applications and future potential. *Appl Microbiol Biotechnol.* 87:427-44.

Banat IM, Makkar RS & Cameotra SS (2000) Potential commercial applications of microbial surfactants. *Appl Microbiol Biotechnol*. 53:495–508

Barrons R & Tassone D (2008) Use of *Lactobacillus* probiotics for bacterial genitourinary infections in women: a review. *Clin Ther.* 30:453-468

Basak P, Adhikari B, Banerjee I & Maiti TK (2009) Sustained release of antibiotic from polyurethane coated implant materials. *J Mater Sci Mater Med.* 20:S213–S221

Becher P (Ed.) (1965) *Emulsions, Theory and practice*, Reinhold Publishing, New York, USA.

Benincasa M, Abalos A, Oliveira I & Manresa A (2004) Chemical structure, surface properties and biological activities of the biosurfactant produced by *Pseudomonas aeruginosa* LBI from soapstock. *Anton Leeuw Int J G.* 85:1–8

Bernheimer AW & Avigad LS (1970) Nature and properties of a cytolytic agent produced by *Bacillus subtilis. J Gen Microbiol.* 6:361-366

Besson, F, Peypoux F, Michel G, & Delcambe L (1978). Identification of antibiotics of iturin group in various strains of *Bacillus subtilis*. *J. Antibiot.* (Tokyo) 31:284–288

Besson F et al. (1976) Characterisation of iturin A in antibiotics from various strains of *Bacillus subtilis. J. Antibiot* 29: 1043–1049

Biria D, Maghsoudi E, Roostaazad R, Dadafarin H, Lotfi AS & Amoozegar MA (2010) Purification and characterization of a novel biosurfactant produced by *Bacillus licheniformis* MS3. *World J Microbiol Biotechnol.* 26:871–878

Bodour AA, Miller-Maier RM (1998) Application of a modified drop-collapse technique for surfactant quantitation and screening of biosurfactant-producing microorganisms. *J. Microbiol Methods.* 32:273-280

Boisselier E & Astruc D (2009) Gold nanoparticles in nanomedicine: preparations, imaging, diagnostics, therapies and toxicity. *Chem Soc Rev.* 38:1759–1782

Bonmatin JM, Laprevote O & Peypoux F (2003) Diversity among microbial cyclic lipopeptides: iturins and surfactins. Activity-structure relationships to design new bioactive agents. *Comb Chem High Throughput Screen.* 6:541-556

Bouffioux O, Berquand A, Eeman M, Paquot M, Dufrêne YF, Brasseur R & Deleu M (2007) Molecular organization of surfactin-phospholipid monolayers: effect of phospholipid chain length and polar head. *Biochim Biophys Acta Biomembr.* 1768:1758–1768

Brasseur R, Braun N, El Kirat K, Deleu M, Mingeot-Leclercq MP & Dufrêne YF (2007) The biologically important surfactin lipopeptide induces nanoripples in supported lipid bilayers. *Langmuir* 23:9769–9772

Buijssen KJ, Harmsen HJ, van der Mei HC, Busscher HJ & van der Laan BF (2007) Lactobacilli: important in biofilm formation on voice prostheses. *Otolaryngol Head Neck Surg.* 137:505-507

Busscher HJ, Bruinsma G, van Weissenbruch R, et al. (1998) The effect of buttermilk consumption on biofilm formation on silicone rubber voice prostheses in an artificial throat. *Eur Arch Otorhinolaryngol.* 255:410-413

Busscher HJ, van Hoogmoed CG, Geertsema-Doornbusch GI, van der Kuijl-Booij M & van der Mei HC. (1997) *Streptococcus thermophilus* and its biosurfactants inhibit adhesion by *Candida* spp. on silicone rubber. *Appl Environ Microbiol.* 63:3810-3817.

Byeon SE, Lee YG, Kim BH, Shen T, Lee SY, Park HJ, Park SC, Rhee MH & Cho JY (2008) Surfactin blocks NO production in lipopolysaccharide-activated macrophages by inhibiting NF-κB activation. *J Microbiol Biotechnol.* 18:1984–1989

Çaglar E, Kargul B & Tanboga I (2005) Bacteriotherapy and probiotics role on oral health. *Oral Dis.* 11:131–137

Cameotra SS & Makkar RS (2004) Recent applications of biosurfactants as biological and immunological molecules. *Curr Opin Microbiol.* 7:262–266

Cao XH, Wang AH, Wang CL, Mao DZ, Lu MF, Cui YQ & Jiao RZ (2010) Surfactin induces apoptosis in human breast cancer MCF- 7 cells through a ROS/JNK-mediated mitochondrial/caspase pathway. *Chem Biol Interact* 183:357–362

Cao XH, Zhao SS, Liu DY, Wang Z, Niu LL, Hou LH & Wang CL (2011) ROS-Ca(2+) is associated with mitochondria permeability transition pore involved in surfactin-induced MCF-7 cells apoptosis. *Chem Biol Interact.* 190(1):16-27

Carrillo C, Teruel JA, Aranda FA & Ortiz A (2003) Molecular mechanism of membrane permeabilization by the peptide antibiotic surfactin. *Biochem Biophys Acta.* 1611: 91-97

Chassot F, Negri MF, Svidzinski AE, Donatti L, Peralta RM, Svidzinski TI & Consolaro ME (2008) Can intrauterine contraceptive devices be a *Candida albicans* reservoir? *Contraception.* 77:355-359

Chen J, Song X, Zhang H, Qu Y & Miao J (2006) Sophorolipid produced from the new yeast strain *Wickerhamiella domercqiae* induces apoptosis in H7402 human liver cancer cells. *Appl Microbiol Biotechnol.* 72:52–59

Chen ML, Penfold J, Thomas RK, Smyth TJP, Perfumo A, Marchant R, Banat IM, Stevenson P, Parry A, Tucker I & Grillo I (2010a) Solution self-assembly and adsorption at the air-water interface of the mono and di-rhamnose rhamnolipids and their mixtures. *Langmuir.* 26:18281-18292

Chen ML, Penfold J, Thomas RK, Smyth TJP, Perfumo A, Marchant R, Banat IM, Stevenson P, Parry A, Tucker I, & Grillo I (2010b) Mixing behaviour of the biosurfactant, rhamnolipid, with a conventional anionic surfactant, sodium dodecyl benzene sulfonate. *Langmuir.* 26:17958-17968

Chung YR, Kim CH, Hwang I & Chun J (2000) *Paenibacillus koreensis* sp. nov. A new species that produces an iturin-like antifungal compound. *Int J Syst Evol Microbiol.* 50:1495–1500

Cirigliano, MC & Carman, GM, (1985) Purification and characterization of liposan, a bioemulsifier from *Candida lipolytica*. *Appl Environ Microbiol.* 50:846–850

Cosson P, Zulianello L, Join-Lambert O, Faurisson F, Gebbie L, Benghezal M, van Delden C, Curty LK & Köhler T *(2002)* Pseudomonas aeruginosa *virulence analyzed in a* Dictyostelium discoideum *host system.* J Bacteriol. *184(11):3027-3033*

Cribby S, Taylor M & Reid G (2008) Vaginal microbiota and the use of probiotics. *Interdisciplinary perspectives on infectious diseases*. Article ID 256490, 9 pages.

Das P, Mukherjee S & Sen R (2008) Antimicrobial potential of a lipopeptide biosurfactant derived from a marine *Bacillus circulans*. *J Appl Microbiol.* 104:1675–1684

Dastgheib SMM, Amoozegar MA, Elahi E, Asad S & Banat IM (2008) Bioemulsifier production by a halothermophilic *Bacillus* strain with potential applications in microbially enhanced oil recovery. *Biotechnol Lett.* 30(2):263-270

Deleu M, Paquot M & Nylander T (2008) Effect of fengycin, a lipopeptide produced by *Bacillus subtilis*, on model biomembranes. *Biophys J.* 94:2667–2679

Dempsey KE, Riggio MP, Lennon A, Hannah VE, Ramage G, Allan D & Bagg J (2007) Identification of bacteria on the surface of clinically infected and noninfected prosthetic hip joints removed during revision arthroplasties by 16S rRNA gene sequencing and by microbiological culture. *Arthritis Res Ther.* 9:R46

Desai JD & Banat IM (1997) Microbial production of surfactants and their commercial potential. *Microbiol Mol Biol Rev.* 61:47–64

Donadio S, Monciardini P, Alduina R, Mazza P, Chiocchini C, Cavaletti L, Sosio M & Puglia AM (2002) Microbial technologies for the discovery of novel bioactive metabolite. *J Biotechnol.* 99:187–198

Donlan RM & Costerton JW (2002) Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev.* 15: 167-193

Eeman M, Francius G, Dufrêne YF, Nott K, Paquot M & Deleu M (2009) Effect of cholesterol and fatty acids on the molecular interactions of fengycin with stratum corneum mimicking lipid monolayers. *Langmuir.* 25:3029–3039

Faivre V & Rosilio V (2010) Interest of glycolipids in drug delivery: from physicochemical properties to drug targeting *Expert Opin Drug Deliv.* 7(9):1031-1048

Falagas ME, Betsi GE, Tokas T & Athanassiou S (2006a) Probiotics for prevention of recurrent urinary tract infections in women: a review of the evidence from microbiological and clinical studies. *Drugs*. 66:1253-1261.

Falagas ME, Betsi GI & Athanasiou S (2007) Probiotics for the treatment of women with bacterial vaginosis. *Clin Microbiol Infect.* 13: 657–664

Falagas ME, Betsi GI, Athanasiou S (2006b) Probiotics for prevention of recurrent vulvovaginal candidiasis: a review. *Antimicrob Chemother*. 58:266-272.

Falagas ME & Makris GC (2009) Probiotic bacteria and biosurfactants for nosocomial infection control: a hypothesis. *J Hosp Infect*. 71(4):301-306

Fassi FL, Wroblewski H & Blanchard A (2007) Activities of antimicrobial peptides and synergy with enrofloxacin against *Mycoplasma pulmonis. Antimicrob Agents Chemother*. 51:468-74

Federle MJ & Bassler BL (2003) Interspecies communication in bacteria. *J Clin Invest.* 112:1291–1299

Fracchia L, Cavallo M, Allegrone G & Martinotti MG (2010) A *Lactobacillus*-derived biosurfactant inhibits biofilm formation of human pathogenic *Candida albicans* biofilm producers, In: *Current Research, Technology and Education Topics in Applied*

Microbiology and Microbial Biotechnology (vol. 2), Mendez Vilas A Ed., pp. 827-837, FORMATEX, Spain

Francius G, Dufour S, Deleu M, Paquot M, Mingeot-Leclercq MP & Dufrêne YF (2008) Nanoscale membrane activity of surfactins: influence of geometry, charge and hydrophobicity. *Biochim Biophys Acta*. 1778:2058–2068

Franzetti A, Tamburini E & Banat IM (2010a) Applications of biological surface active compounds in remediation technologies, In: *Biosurfactants*, *BIOSURFACTANTS Book Series: Advances in Experimental Medicine and Biology*, Sen R Ed., Volume: 672, pp. 121-134, Landes Bioscience, Austin, TX

Franzetti A, Gandolfi I, Bestetti G, Smyth TJP & Banat IM (2010b) Production and applications of trehalose lipid biosurfactants. *Eur J Lipid Sci Technol*. 112:617–627

Gautam KK & Tiagi VK (2006) Microbial surfactants: A review. *J Oleo Sci.* 55:155-166

Grangemard I, Wallach J, Maget-Dana R & Peypoux F (2001) Lichenysin: a more efficient cation chelator than surfactin. *Appl Biochem Biotechnol.* 90:199–210

Gudiña EJ, Teixeira JA & Rodrigues LR (2010) Isolation and functional characterization of a biosurfactant produced by *Lactobacillus paracasei*. *Colloids Surf B Biointerfaces.* 76:298–304

Gupta V & Garg R (2009) Probiotics. *Indian J Med Microbiol.* 27:202–209

Hall-Stoodley L, Costerton JW & Stoodley P (2004) Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol*. 2:95-108

Han Y, Huang X, Cao M & Wang Y (2008) Micellization of surfactin and its effect on the aggregate conformation of amyloid β(1-40). *J Phys Chem B*. 112:15195–15201

Hancock, REW & Chapelle DS (1999) Pepedide antibiotics. *Antimicrob Agents Chemother.* 43:1317–1323

Hatakka K, Ahola AJ, Yli-Knuuttila H, Richardson M, Poussa T & Meurman JK (2007) Probiotics reduce the prevalence of oral *Candida* in the elderly—a randomized controlled trial. *J Dent Res.* 86:125–130

Hong HA, Duc LH & Cutting SM (2005) The use of bacterial spore formers as probiotics. *FEMS Microbiol Rev*. 29:813–835

Horowitz S, Gilbert JN & Griffin WM (1990) Isolation and characterization of a surfactant produced by *Bacillus licheniformis* 86. *J Ind Microbiol Biot.* 6(4):243-248

Huang X, Lu Z, Bie X, Lü F, Zhao H & Yang S (2007) Optimization of inactivation of endospores of *Bacillus cereus* by antimicrobial lipopeptides from *Bacillus subtilis* fmbj strains using a response surface method. *Appl Microbiol Biotechnol.* 74:454–461

Huang X, Lu Z, Zhao H, Bie X, Lü FX & Yang S (2006) Antiviral activity of antimicrobial lipopeptide from *Bacillus subtilis* fmbj against pseudorabies virus, porcine parvovirus, newcastle disease virus and infectious bursal disease virus in vitro. *Int J Pept Res Ther*. 12:373–377

Huang X, Suo J & Cui Y (2011) Optimization of antimicrobial activity of surfactin and polylysine against *Salmonella enteritidis* in milk evaluated by a response surface methodology. *Foodborne Pathog Dis.* 8(3):439-43

Igarashi S, Hattori Y & Maitani Y (2006) Biosurfactant MEL-A enhances cellular association and gene transfection by cationic liposome. *J Control Release.* 112:362–368

Imamura Y, Chandra J, Mukherjee PK, Lattif AA, Szczotka-Flynn LB, Pearlman E, Lass JH, O'Donnell K & Ghannoum MA (2008) *Fusarium* and *Candida albicans* biofilms on soft contact lenses: model development, influence of lens type, and susceptibility to lens care solutions. *Antimicrob Agents Chemother*. 52: 171-182

Imura T, Ito S, Azumi R, Yanagishita H, Sakai H, Abe M & Kitamoto D (2007) Monolayers assembled from a glycolipid biosurfactant from *Pseudozyma* (*Candida*) *antarctica* serve as a high-affinity ligand system for immunoglobulin G and M. *Biotechnol Lett*. 29:865–870

Imura T, Masuda Y, Ito S, Worakitkanchanakul W, Morita T, Fukuoka T, Sakai H, Abe M & Kitamoto D (2008) Packing density of glycolipid biosurfactant monolayers give a significant effect on their binding affinity toward immunoglobulin G. *J Oleo Sci*. 57:415-22

Inoh Y, Furuno T, Hirashima N & Nakanishi M (2009) Nonviral vectors with a biosurfactant MEL-A promote gene transfection into solid tumors in the mouse abdominal cavity. *Biol Pharm Bull*. 32:126—128

Inoh Y, Furuno T, Hirashima N, Kitamoto D & Nakanishi M (2010) The ratio of unsaturated fatty acids in biosurfactants affects the efficiency of gene transfection. *Int J Pharmaceut*. 398:225–230

Inoh Y, Kitamoto D, Hirashima N & Nakanishi M (2001) Biosurfactants of MEL-A increase gene transfection mediated by cationic liposomes. *Biochem Biophys Res Commun*. 289:57–61

Inoh Y, Kitamoto D, Hirashima N & Nakanishi M (2004) Biosurfactant MEL-A dramatically increases gene transfection via membrane fusion. *J Control Release* 94:423–431.

Ishigami Y & Suzuki S (1997) Development of biochemicals—functionalization of biosurfactants and natural dyes. *Prog Org Coatings*. 31:51–61.

Ito S, Imura T, Fukuoka T, Morita T, Sakai H, Abea M & Kitamoto D (2007) Kinetic studies on the interactions between glycolipid biosurfactant assembled monolayers and various classes of immunoglobulins using surface plasmon resonance. *Colloids Surf B Biointerfaces*. 58:165–171

Jaques P (2010) Surfactin and other lipopeptides from *Bacillus* spp. In: *Biosurfactants: From Genes to Applications,* Soberón-Chávez G Ed., pp. 57-91, Springer, Münster, Germany

Jung M, Lee S & Kim H (2000) Recent studies on natural products as anti-HIV agents. *Curr Med Chem*. 7:649-661

Kalishwaralal K, Deepak V, Pandiana SBRK, Kottaisamy M, BarathManiKanth S, Kartikeyan B & Gurunathan S (2010) Biosynthesis of silver and gold nanoparticles using *Brevibacterium casei. Colloids Surf B Biointerfaces.* 77:257–262

Kalmokoff ML, Austin JW, Wan XD, Sanders G, Banerjee S & Farber JM (2001) Adsoption, attachment and biofilm formation among isolates of *Listeria monocytogenes* using model conditions. *J Appl Microbiol.* 91:725–734

Kim BS, Lee JY & Hwang BK (2000) *In vivo* control and *in vitro* antifungal activity of rhamnolipid B, a glycolipid antibiotic, against *Phytophthora capsici* and *Colletotrichum orbiculare. Pest Manag Sci.* 56:1029-1035

Kim H, Ryu JH & Beuchat LR (2006) Attachment of and biofilm formation by *Enterobacter sakazakii* on stainless steel and enteral feeding tubes. *Appl Environ Microbiol*. 72:5846–5856

Kim K, Yoo D, Kim Y, Lee B, Shin D & Kim E-K (2002) Characteristics sophorolipid as an antimicrobial agent. *J Microbiol Biotechnol.* 12:235-241

Kim PI, Ryu J, Kim YH & Chi YT (2010) Production of biosurfactant lipopeptides iturin A, fengycin, and surfactin A from *Bacillus subtilis* CMB32 for control of *Colletotrichum gloeosporioides. J Microbiol Biotechnol*. 20(1):138–145

Kiran GS, Sabarathnam B & Selvin J (2010a) Biofilm disruption potential of a glycolipid biosurfactant from marine *Brevibacterium casei. FEMS Immunol Med Microbiol.* 59:432–438

Kiran GS, Sabu A & Selvin J (2010b) Synthesis of silver nanoparticles by glycolipid biosurfactant produced from marine *Brevibacterium casei* MSA19. *J Biotechnol.* 148:221–225

Kitamoto D, Isoda H, Nakahara T. (2002) Functions and potential applications of glycolipid biosurfactants-from energy-saving materials to gene delivery carriers. *J Biosci Bioeng*. 94(3):187-201.

Kitamoto D, Morita T, Fukuoka T, Konishi M & Imura T (2009) Self-assembling properties of glycolipid biosurfactants and their potential applications. *Curr Opin Colloid Interface Sci.* 14:315– 328

Kitamoto D, Toma K & Hato M (2005) Glycolipid-based nanomaterials, In: Handbook of Nanostructured Biomaterials and Their Applications in Nanobiotechnology, vol. 1, Nalwa HS Ed., p. 239–271, American Science Publishers, California, USA

Kitamoto D, Yanagishita H, Shinbo T, Nakane T, Kamisawa C & Nakahara T (1993) Surface active properties and antimicrobial activities of mannosylerythritol lipids as biosurfactants produced by *Candida antarctica*. *J Biotechnol*. 29:91–96

Kõll P, Mändar R, Marcotte H, Leibur E, Mikelsaar M & Hammarström L (2008) Characterization of oral lactobacilli as potential probiotics for oral health. *Oral Microbiol Immunol.* 23:139–147

Konishi M, Imura T, Fukuoka T, Morita T & Kitamoto D (2007) A yeast glycolipid biosurfactant, mannosylerythritol lipid, shows high binding affinity towards lectins on a self-assembled monolayer system. *Biotechnol Lett.* 29:473–480

Kumar A, Ali A & Yerneni LK (2007) Effectiveness of a mycoplasma elimination reagent on a mycoplasma-contaminated hybridoma cell line. *Hybridoma (Larchmt).* 26(2):104-106

Kumar CG, Mamidyala SK, Das B, Sridhar B, Devi GS & Karuna MS (2010) Synthesis of biosurfactant-based silver nanoparticles with purified rhamnolipids isolated from *Pseudomonas aeruginosa* BS-161R. *J Microbiol Biotechnol.* 20:1061-1068

Kuyukina MS, Ivshina IB, Gein SV, Baeva TA & Chereshnev VA (2007) *In vitro* immunomodulating activity of biosurfactant glycolipid complex from *Rhodococcus ruber*. *Bull Exp Biol Med.*144:326-30.

Landman D, Georgescu C, Martin DA & Quale J (2008) Polymyxins revisited. *Clin Microbiol Rev.* 21:449–465

Lang S (2002) Biological amphiphiles (microbial biosurfactants). *Curr Opin Coll Int Sci.* 7:12–20

Lang S & Philp JC (1998) Surface-active lipids in *Rhodococci*. *Anton Leeuw Int. J G*. 74:59–70.

Lin SC (1996) Biosurfactants: recent advances. *J Chem Tech Biotechnol.* 66:109-120

Litzler PY, Benard L, Barbier-Frebourg N, Vilain S, Jouenne T, Beucher E, Bunel C, Lemeland JF & Bessou JP (2007) Biofilm formation on pyrolytic carbon heart valves: influence of surface free energy, roughness, and bacterial species. *J Thorac Cardiovasc Surg*. 134:1025-1032

Lu JR, Zhao XB & Yaseen M (2007) Biomimetic amphiphiles: biosurfactants. *Curr Opin Colloid Interface Sci*. 12:60-67

Luna JM, Rufino RD, Sarubbo LA, Rodrigues LR, Teixeira JA & de Campos-Takaki GM (2011) Evaluation antimicrobial and antiadhesive properties of the biosurfactant lunasan produced by *Candida sphaerica* UCP 0995. *Curr Microbiol.* 62:1527-1534

Maier RM (2003) Biosurfactants: evolution and diversity in bacteria. *Adv Appl Microbiol.* 52: 101-121

Maitani Y, Yano S, Hattori Y, Furuhata M & Hayashi K (2006) Liposome vector containing biosurfactant-complexed DNA as herpes simplex virus thymidine kinase gene delivery system. *J Liposome Res.* 16:359-72.

Martinotti MG, Rivardo F Allegrone G, Ceri H, Turner R (2009) Biosurfactant composition produced by a new *Bacillus licheniformis* strain, uses and products thereof. International patent PCT/IB2009/055334, 25 November

Matsuyama T, Tanikawa T, & Nakagawa Y (2010) Serrawettins and other surfactants produced by *Serratia*. In: *Biosurfactants: From Genes to Applications,* Soberón-Chávez G Ed., pp. 93-120, Springer, Münster, Germany

McInerney MJ, Javaheri M & Nagle DP (1990) Properties of the biosurfactant produced by *Bacillus liqueniformis* strain JF-2. *I. J. Microbiol Biotechnol*. 5:95–102

Merk K, Borelli C & Korting HC (2005) Lactobacilli—bacteria–host interactions with special regard to the urogenital tract. *Int J Med Microbiol*. 295:9–18

Meurman JH (2005) Probiotics: do they have a role in oral medicine and dentistry? *Eur J Oral Sci.* 113:188–196

Meurman JH & Stamatova I (2007) Probiotics: contributions to oral health. *Oral Dis.* 13:443–445

Meylheuc T, Methivier C, Renault M, Herry JM, Pradier CM & Bellon-Fontaine MN (2006a) Adsorption on stainless steel surfaces of biosurfactants produced by gram-negative and gram-positive bacteria: consequence on the bioadhesive behavior of *Listeria monocytogenes*. *Colloids Surf B Biointerfaces.*52:128–137

Meylheuc T, Renault M & Bellon-Fontaine MN (2006b) Adsorption of a biosurfactant on surfaces to enhance the disinfection of surfaces contaminated with *Listeria monocytogenes*. *Int J Food Microbiol.* 109:71–78

Mimee B, Labbé C, Pelletier R & Bélanger RR (2005) Antifungal activity of flocculosin, a novel glycolipid isolated from *Pseudozyma flocculosa. Antimicrob Agents Chemother.* 49:1597–1599

Mimee B, Pelletier R & Bélanger RR (2009) In vitro antibacterial activity and antifungal mode of action of flocculosin, a membrane-active cellobiose lipid. *J Appl Microbiol.* 107:989–996

Mireles JR II, Toguchi A & Harshey RM (2001) *Salmonella enterica* serovar *typhimurium* swarming mutants with altered biofilm forming abilities: surfactin inhibits biofilm formation. *J Bacteriol.* 183:5848–5854

Mohammad Abdel-Mawgoud A, Hausmann R, Lépine F, Müller MM & Déziel E (2010) Rhamnolipids: detection, analysis, biosynthesis, genetic regulation, and bioengineering of production, In: *Biosurfactants: From Genes to Applications,* Soberón-Chávez G Ed., pp. 13-55, Springer, Münster, Germany

Morikawa M (2006) Beneficial biofilm formation by industrial bacteria *Bacillus subtilis* and related species. *J Biosci Bioeng.* 101:1–8

Morikawa M, Hirata Y, Imanaka T (2000). A study on the structure–function relationship of the lipopeptide biosurfactants. *Biochim Biophys Acta*. 1488:211-218

Morikawa M, Ito M & Imanaka T (1992) Isolation of a new surfactin producer *Bacillus pumilus* A-1, and cloning and nucleotide sequence of the regulator gene, psf-1. *J Ferment Bioeng*. 74:255-261

Morita T, Fukuoka T, Konishi M, Imura T, Yamamoto S, Kitagawa M, Sogabe A & Kitamoto D (2009a) Production of a novel glycolipid biosurfactant, mannosylmannitol lipid, by *Pseudozyma parantarctica* and its interfacial properties. *Appl Microbiol Biotechnol*. 83:1017–1025

Morita T, Kitagawa M, Suzuki M, Yamamoto S, Sogabe A, Yanagidani S, Imura T, Fukuoka T & Kitamoto D (2009b) A yeast glycolipid biosurfactant, mannosylerythritol lipid, shows potential moisturizing activity toward cultured human skin cells: the recovery effect of MEL-A on the SDS-damaged human skin cells. *J Oleo Sci*. 58:639–642

Morita T, Kitagawa M, Yamamoto S, Suzuki M, Sogabe A, Imura T, Fukuoka T & Kitamoto D (2010) Activation of fibroblast and papilla cells by glycolipid biosurfactants, mannosylerythritol lipids. *J Oleo Sci*. 59:451-5

Mukherjee S, Das P & Sen R (2006) Towards commercial production of microbial surfactants. *Trends Biotechnol*. 24:509-515

Muthusamy K, Gopalakrishnan S, Ravi TK & Sivachidambaram P (2008) Biosurfactants: Properties, commercial production and application. *Curr Sci*. 94:736-747

Nakanishi M, Inoh Y, Kitamoto D & Furuno T (2009) Nano vectors with a biosurfactant for gene transfection and drug delivery. *J Drug Delivery Sci Technol*. 19:165–169.

Narayanan J, Ramji R, Sahu H & Gautam P (2010) Synthesis, stabilisation and characterisation of rhamnolipid-capped ZnS nanoparticles in aqueous medium. *IET Nanobiotechnol*. 4:29-34.

Narayanan KB & Sakthivel N (2010) Biological synthesis of metal nanoparticles by microbes. *Adv Colloid Interfac*. 156:1–13

Naruse N, Tenmyo O, Kobaru S, Kamei H, Miyaki T, Konishi M & Oki T (1990) Pumilacidin, a complex of new antiviral antibiotics: production, isolation, chemical properties, structure and biological activity. *J Antibiot*. (Tokyo) 43:267–280

Neu TR (1996) Significance of bacterial surface-active compounds in interaction of bacteria with interfaces. *Microbiol Rev*. 60:151–166

Nguyen TT & Sabatini DA (2009) Formulating alcohol-free microemulsions using rhamnolipid biosurfactant and rhamnolipid mixtures. *J Surfact Deterg*. 12:109–115

Nguyen TTL, Edelen A, Neighbors B & Sabatini DA (2010) Biocompatible lecithin-based microemulsions with rhamnolipid and sophorolipid biosurfactants: formulation and potential applications. *J Colloid Interface Sci*. 348:498-504

Nicoli S, Eeman M, Deleu M, Bresciani E, Padula C & Santi, P (2010) Effect of lipopeptides and iontophoresis on aciclovir skin delivery. *J Pharm Pharmacol*. 62:702–708

Nithya C & Pandian SK (2010) The in vitro antibiofilm activity of selected marine bacterial culture supernatants against *Vibrio* spp. *Arch Microbiol*. 192:843–854

Nitschke M, Araújo LV, Costa SG, Pires RC, Zeraik AE, Fernandes AC, Freire DM & Contiero J (2009) Surfactin reduces the adhesion of food-borne pathogenic bacteria to solid surfaces. *Lett Appl Microbiol*. 49:241–247

Nitschke M, Costa SG & Contiero J (2010) Structure and applications of a rhamnolipid surfactant produced in soybean oil waste. *Appl Biochem Biotechnol.* 160(7):2066-74

Ortiz A, Teruel JA, Espuny MJ, Marqués A, Manresa Á & Aranda FJ (2008) Interactions of a *Rhodococcus* sp. biosurfactant trehalose lipid with phosphatidylethanolamine membranes. *Biochim Biophys Acta.* 1778:2806–2813

Ortiz A, Teruel JA, Espuny MJ, Marqués A, Manresa Á & Aranda FJ (2009) Interactions of a bacterial biosurfactant trehalose lipid with phosphatidylserine membranes. *Chem Phys Lipids.* 158:46–53

Osada, H (1998) Bioprobe for investigating mammlian cell cycles control. *J. Antibiotics* 51:973-982

Owen D & Fan L (2010) Polymeric Biosufactants. US Patent 20100144643, 6 October

Palanisamy P (2008) Biosurfactant mediated synthesis of NiO nanorods. *Mat Lett.* 62:743–746

Palanisamy P & Raichur AM (2009) Synthesis of spherical NiO nanoparticles through a novel biosurfactant mediated emulsion technique. *Mater Scie Eng C.* 29:199–204

Park SY & Kim Y (2009) Surfactin inhibits immunostimulatory function of macrophages through blocking NK-κB, MAPK and Akt pathway. *Int Immunopharmacol.* 9:886–893

Park SY, Kim YH, Kim EK, Ryu EY & Lee SJ (2010) Heme oxygenase-1 signals are involved in preferential inhibition of pro-inflammatory cytokine release by surfactin in cells activated with *Porphyromonas gingivalis* lipopolysaccharide. *Chem Biol Interact.* 188:437-45

Pecci Y, Rivardo F, Martinotti MG & Allegrone G (2010) LC/ESI-MS/MS characterisation of lipopeptide biosurfactants produced by the *Bacillus licheniformis* V9T14 strain *J Mass Spectrom.* 45:772–778

Perfumo A, Banat IM, Canganella F & Marchant R (2006) Rhamnolipid production by a novel thermotolerant hydrocarbon-degrading *Pseudomonas aeruginosa AP02-1. J Appl Microbiol.* 75:132-138

Perfumo A, Smyth TJP, Marchant R & Banat IM (2010a) Production and roles of biosurfactants and bioemulsifiers in accessing hydrophobic substrates, In: *Handbook of Hydrocarbon and Lipid Microbiology,* Timmis KN Ed., pp. 1501-1512, Springer-Verlag, Berlin Heidelberg, Germany

Perfumo A, Rancich I & Banat IM (2010b). Possibilities and challenges for biosurfactants uses in petroleum industry, In: *Biosurfactants, BIOSURFACTANTS Book Series: Advances in Experimental Medicine and Biology*, Sen R Ed., Volume: 672, pp. 135-145, Landes Bioscience, Austin, TX

Petrelli D, Zampaloni C, D'Ercole S, Prenna M, Ballarini P, Ripa S & Vitali LA (2006) Analysis of different genetic traits and their association with biofilm formation in *Staphylococcus epidermidis* isolates from central venous catheter infections. *Eur J Clin Microbiol Infect Dis.* 25:773-781

Peypoux F, Bonmatin JM & Wallach J (1999) Recent trends in the biochemistry of surfactin. *Appl Microbiol Biotechnol.* 51:553–563

Peypoux F, Guinand M, Michel G, Delcambe L, Das BC & Lederer E (1978) Structure of iturin A, a peptidolipid antibiotic from *Bacillus subtilis*. *Biochemistry.* 17:3992–3996

Peypoux F, Pommier MT, Marion D, Ptak M, Das BC & Michel G 1986. Revised structure of mycosubtilin, a peptidolipid antibiotic from *Bacillus subtilis*. *J. Antibiot.* (Tokyo) 39:636–641

Peypoux F, Pommier MT, Das BC, Besson F, Delcambe L & Michel G (1984) Structures of bacillomycin D and bacillomycin L peptidolipid antibiotics from *Bacillus subtilis*. *J. Antibiot*. (Tokyo) 77:1600–1604

Pierce D & Heilman TJ (2001) Germicidal composition. US Patent 6262038, 17 July

Pissuwan D, Niidome T & Cortie MB (2011) The forthcoming applications of gold nanoparticles in drug and gene delivery systems. *J Control Release*. 149:65–71

Pissuwan D, Valenzuela SM, Miller CM, Killingsworth MC & Cortie MB (2009) Destruction and control of *Toxoplasma gondii* tachyzoites using gold nanosphere/antibody conjugates. *Small*. 5:1030–1034.

Ramisse F, Delden C, Gidenne S et al. (2000) Decreased virulence of a strain of *Pseudomonas aeruginosa* O12 overexpressing a chromosomal type 1 β-lactamase could be due to reduced expression of cell-to-cell signalling dependent virulence factors. *FEMS Immunol Med Microbiol*. 28:241–245

Rasmussen TB & Givskov M (2006) Quorum-sensing inhibitors as antipathogenic drugs. *Int J Med Microbiol*. 296:149–161

Raya A, Sodagari M, Pinzon NM, He X, Newby BZ & Ju LK (2010) Effects of rhamnolipids and shear on initial attachment of *Pseudomonas aeruginosa* PAO1 in glass flow chambers. *Environ Sci Pollut Res*. 17:1529–1538

Raza ZA, Khalid ZM & Banat IM (2009) Characterization of rhamnolipids produced by a *Pseudomonas aeruginosa* mutant strain grown on waste oils. *J Environ Sci Health A Tox Hazard Subst Environ Eng*. 44:1367-1373

Raza ZA, Khalid ZM, Khan MS, Banat IM, Rehman A, Naeem A & Saddique MT (2010) Surface properties and sub-surface aggregate assimilation of Rhamnolipid surfactants in different aqueous system. *Biotechnol Lett*. 32:811-816

Reddy AS, Chen CY, Chen CC, Jean JS, Chen HR, Tseng MJ, Fan CW & Wang JC. (2010) Biological synthesis of gold and silver nanoparticles mediated by the bacteria *Bacillus subtilis*. *J Nanosci Nanotechnol*. 10:6567-74

Reddy AS, Chen CY, Chen CC, Jean JS, Fan CW, Chen HR, Wang JC & Nimje VR (2009) Synthesis of gold nanoparticles via an environmentally benign route using a biosurfactant. *J Nanosci Nanotechnol*. 9:6693–6699

Reid G, Bruce A & Smeianov V (1998) The role of Lactobacilli in preventing urogenital and intestinal infections. *Int Dairy J*. 8:555–562

Reid G, Bruce AW, Busscher HJ & Van der Mei HC (2000) *Lactobacillus* therapies. US Patent 6051552 18 April

Reid G, Bruce AW, Fraser N, Heinemann C, Owen J & Henning B (2001) Oral probiotics can resolve urogenital infections. *FEMS Immunol Med Microbiol*. 30:49–52

Remichkova M, Galabova D, Roeva I, Karpenko E, Shulga A & Galabov AS (2008) Anti-herpesvirus activities of *Pseudomonas* sp. S-17 rhamnolipid and its complex with alginate. *Z Naturforsch C*. 63:75–81

Rivardo F, Martinotti MG, Turner RJ & Ceri H (2010) The activity of silver against *Escherichia coli* biofilm is increased by a lipopeptide biosurfactant. *Can. J. Microbiol*. 56:272–278

Rivardo F, Martinotti MG, Turner RJ & Ceri H (2011) Synergistic effect of lipopeptide biosurfactant with antibiotics against *Escherichia coli* CFT073 biofilm. *Int J Antimicrob Agents*. 37:324-331

Rivardo F, Turner RJ, Allegrone G & Ceri H, Martinotti MG (2009) Anti-adhesion activity of two biosurfactants produced by *Bacillus* spp. prevents biofilm formation of human bacterial pathogens. *Appl Microbiol Biotechnol.* 83:541–553

Rodrigues L, Banat IM, Teixeira J & Oliveira R (2006a) Biosurfactants: potential applications in medicine. *J Antimicrob Chemother.* 57:609–618

Rodrigues L, van der Mei H, Banat IM, Teixeira J & Oliveira R (2006b) Inhibition of microbial adhesion to silicone rubber treated with biosurfactant from Streptococcus thermophilus A. *FEMS Immunol Med Microbiol.* 46:107–112

Rodrigues L, van der Mei HC, Teixeira J & Oliveira R (2004a) Biosurfactant from *Lactococcus lactis* 53 inhibits microbial adhesion on silicone rubber. *Appl Microbiol Biotechnol.* 66:306-311.

Rodrigues L, van der Mei HC, Teixeira J & Oliveira R (2004b) Influence of biosurfactants from probiotic bacteria on formation of biofilms on voice prosthesis. *Appl Environ Microbiol* 70:4408-4410.

Rodrigues L, Banat IM, van der Mei HC, Teixeira JA, Oliveira R & Oliveira R (2006c) Interference in adhesion of bacteria and yeasts isolated from explanted voice prostheses to silicone rubber by rhamnolipid biosurfactants. *J Appl Microbiol.* 100:470–480.

Rodrigues L, Banat IM, Teixeira J & Oliveira R (2007) Strategies for the prevention of microbial biofilm formation on silicone rubber voice prostheses. *J Biomed Mater Res B Appl Biomater.* 81B:358-370.

Rodrigues LR, Teixeira JA, van der Mei HC & Oliveira R (2006d) Physicochemical and functional characterization of a biosurfactant produced by *Lactococcus lactis* 53. *Colloids Surf B Biointerfaces*. 49:79–86

Rodrigues LR & Teixeira JA (2010) Biomedical and therapeutic applications of biosurfactants. *Adv Exp Med Biol*. 672:75-87

Ron EZ & Rosenberg E (2001) Natural roles of biosurfactants. *Environ Microbiol.* 3:229–236

Rosenberg E & Ron EZ (1997) Bioemulsans: microbial polymeric emulsifiers. *Curr Opin Biotechnol.* 8:313–316

Rosenberg E & Ron EZ (1999) High- and low-molecular-mass microbial surfactants. *Appl Microbiol Biotechnol* 52:154–162

Rosenberg E, Zuckerberg A, Rubinovitz C & Gutnick DL (1979) Emulsifier of *Arthrobacter* RAG-1:isolation and emulsifying properties. *Appl Environ Microbiol.* 37:402–408

Rosenberg M (2006) Microbial adhesion to hydrocarbons: twenty-five years of doing MATH. *FEMS Microbiol Lett.* 262:129–134

Rufino RD, Luna JM, Sarubbo LA, Rodrigues LR, Teixeira JA & Campos-Takaki GM (2011) Antimicrobial and anti-adhesive potential of a biosurfactant Rufisan produced by *Candida lipolytica* UCP 0988. *Colloids Surf B Biointerfaces.* 84:1-5

Saini HS, Barragán-Huerta BE, Lebrón-Paler A, Pemberton JE, Vázquez RR, Burns AM, Marron MT, Seliga CJ, Gunatilaka AA & Maier RM (2008) Efficient purification of the biosurfactant viscosin from *Pseudomonas libanensis* strain M9-3 and its physicochemical and biological properties. *J Nat Prod.* 71:1011–1015

Sánchez M, Aranda FJ, Teruel JA, Espuny MJ, Marqués A, Manresa Á & Ortiz A (2010) Permeabilization of biological and artificial membranes by a bacterial dirhamnolipid produced by *Pseudomonas aeruginosa*. *J Colloid Interface Sci.* 341:240–247

Sánchez M, Aranda FJ, Teruel JA & Ortiz A (2009) Interaction of a bacterial dirhamnolipid with phosphatidylcholine membranes: a biophysical study. *Chem Phys Lipids.* 161:51–55

Saravanakumari P & Mani K (2010) Structural characterization of a novel xylolipid biosurfactant from *Lactococcus lactis* and analysis of antibacterial activity against multi-drug resistant pathogens. *Bioresour Technol.* 101:8851–8854

Satpute SK, Banat IM, Dhakephalkar PK, Banpurkar AG, Chopade BA (2010a) Biosurfactants, bioemulsifiers and exopolysaccharides from marine microorganisms. *Biotechnol Adv.* 28:436-450

Satpute SK, Banpurkar AG, Dhakephalkar PK, Banat IM & Chopade BA (2010b) Methods for investigating biosurfactants and bioemulsifiers: a review. *Crit Rev Biotechnol.* 30:127-144

Selvam R, Maheswari P, Kavitha P, Ravichandran M, Sas B & Ramchand CN (2009) Effect of *Bacillus subtilis* PB6, a natural probiotic on colon mucosal inflammation and plasma cytokines levels in inflammatory bowel disease. *Indian J Biochem Biophys.* 46:79–85

Seydlová G & Svobodová J (2008) Review of surfactin chemical properties and the potential biomedical applications. *Cent Eur J Med.* 3:123–133

Shah V, Doncel GF, Seyoum T, Eaton KM, Zalenskaya I, Hagver R,Azim A & Gross R (2005) Sophorolipids, microbial glycolipids with anti-human immunodeficiency virus and sperm-immobilizing activities. *Antimicrob Agents Chemother.* 49:4093–4100

Shaligram NS & Singhal RS (2010) Surfactin – A review, on biosynthesis, fermentation, purification and applications. *Food Technol Biotechnol.* 48:119–134

Shao Z (2010) Trehalolipids, In: *Biosurfactants: From Genes to Applications,* Soberón-Chávez G Ed., pp. 121-143, Springer, Münster, Germany

Sharma VK., Yngard RA & Lin Y (2009) Silver nanoparticles: Green synthesis and their antimicrobial activities. *Adv Colloid Interfac.* 145:83–96

Shete AM, Wadhava G, Banat IB & Chopade BA (2006) Mapping of patents on bioemulsifiers and biosurfactants : A review. *J Sci Ind Res* (India). 65:91-115.

Singh A, van Hamme JD & Ward OP (2007) Surfactants in microbiology and biotechnology: part 2. Application aspects. *Biotechnol Adv.* 25:99–121

Singh BR, Dwivedi S, Al-Khedhairy AA & Musarrat J (2011) Synthesis of stable cadmium sulfide nanoparticles using surfactin produced by *Bacillus amyloliquifaciens* strain KSU-109. *Colloids Surf B Biointerfaces.* In press

Singh P & Cameotra SS (2004) Potential applications of microbial surfactants in biomedical sciences. *Trends Biotechnol.* 22:142–146

Singh S, D'Britto V, Prabhune AA, Ramana CV, Dhawan A & Prasad BLV (2010) A Cytotoxic and genotoxic assessment of glycolipid-reduced and -capped gold and silver nanoparticles. *New J Chem.* 34:294-301

Singh S, Patel P, Jaiswal S, Prabhune AA, Ramana CV & Prasad BLV (2009) A direct method for the preparation of glycolipid–metal nanoparticle conjugates: sophorolipids as reducing and capping agents for the synthesis of water re-dispersible silver nanoparticles and their antibacterial activity. *New J Chem.* 33:646-652

Sivapathasekaran, C, Das P, Mukherjee S, Saravanakumar J, Mandal M & Sen R (2010) Marine bacterium derived lipopeptides: characterization and cytotoxic activity against cancer cell lines. *Int J Pept Res Ther.* 16:215–222

Smyth TJP, Perfumo A, Marchant R & Banat IM (2010a) Isolation and analysis of low molecular weight microbial glycolipids, In: *Handbook of Hydrocarbon and Lipid Microbiology*, Timmis KN Ed., pp. 3705-3723, Springer, Berlin,

Smyth TJ, Perfumo A, Marchant R, Banat IM, Chen M, Thomas RK, Penfold J, Stevenson PS & Parry NJ (2010b) Directed microbial biosynthesis of deuterated biosurfactants and potential future application to other bioactive molecules. *Appl Microbiol Biotechnol*. 87:1347-1354

Smyth TJP, Perfumo A, McClean S, Marchant R & Banat IM (2010c) Isolation and analysis of lipopeptides and high molecular weight biosurfactants, In: *Handbook of Hydrocarbon and Lipid Microbiology*, Timmis KN Ed., pp. 3689–3704, Springer, Berlin, Germany

Soberón-Chávez G & Maier RM (2010) Biosurfactants: a general overview, In: *Biosurfactants: From Genes to Applications,* Soberón-Chávez G Ed., pp. 1-11, Springer, Münster, Germany

Sotirova AV, Spasova DI, Galabova DN, Karpenko E & Shulga A (2008) Rhamnolipid–biosurfactant permeabilizing effects on gram-positive and gram-negative bacterial strains. *Curr Microbiol.* 56:639–644

Stepanovic S, Cirkovic I, Ranin L & Svabic-Vlahovic M (2004) Biofilm formation by *Salmonella* spp. and *Listeria monocytogenes* on plastic surface. *Lett Appl Microbiol.* 38:428–432

Stickler DJ (2008) Bacterial biofilms in patients with indwelling urinary catheters. *Nat Clin Pract Urol.* 5:598-608

Suzuki M, Kitagawa M, Yamamoto S, Sogabe A, Kitamoto D, Morita T, Fukuoka T & Imura T (2010) Activator including biosurfactant as active ingredient, mannosyl erythritol lipid, and production method publication. Patent application number: 20100168405, 7 January

Tanaka Y, Tojo T, Uchida K, Uno J, Uchida Y & Shida O (1997) Method of producing iturin A and antifungal agent for profound mycosis. *Biotechnol Adv.* 15:234–235

Tang JS, Zhao F, Gao H, Dai Y, Yao ZH, Hong K, Li J, Ye WC & Yao XS (2010) Characterization and online detection of surfactin isomers based on HPLC-MS^n analyses and their inhibitory effects on the overproduction of nitric oxide and the release of TNF-α and IL-6 in LPS-induced macrophages. *Mar Drugs*. 8:2605-2618.

Thavasi R, Jayalakshmi S, Balasubramanian T & Banat IM (2008) Production and characterization of a glycolipid biosurfactant from *Bacillus megaterium* using economically cheaper sources. *World J Microbiol Biotechnol.* 24:917-925

Thavasi R, Jayalakshmi S & Banat IM (2011) Effect of biosurfactant and fertilizer on biodegradation of crude oil by maring isolates of *Bacillus megaterium* and *Corynebacterium kutscheri* and *Pseudomonas aeruginosa*. *Bioresouce Technol.* 102:772-778

Ueno Y, Hirashima N, Inoh Y, Furuno T & Nakanishi M (2007a) Characterization of biosurfactant-containing liposomes and their efficiency for gene transfection. *Biol Pharm Bull.* 30:169–172

Ueno Y, Inoh Y, Furuno T, Hirashima N, Kitamoto D & Nakanishi M (2007b) NBD-conjugated biosurfactant (MEL-A) shows a new pathway for transfection. *J Control Release*. 123:247–253

Valle J, Da Re S, Henry N, Fontaine T, Balestrino D, Latour-Lambert P & Ghigo JM (2006) Broad-spectrum biofilm inhibition by a secreted bacterial polysaccharide. *Proc Natl Acad Sci USA.* 103:12558–12563

Van Bogaert INA & Soetaert W (2010) Sophorolipids. In: *Biosurfactants: From Genes to Applications,* Soberón-Chávez G Ed., pp. 179-210, Springer, Münster, Germany

Van Bogaert INA, Saerens K, De Muynck C, Develter D, Wim S & Vandamme EJ (2007) Microbial production and application of sophorolipids. *Appl Microbiol Biotechnol.* 76:23–34

Van der Mei HC, Free RH, Elving GJ, van Weissenbruch R, Albers FWJ & Busscher HJ (2000) Effect of probiotic bacteria on prevalence of yeasts in oropharyngeal biofilms on silicone rubber voice prostheses *in vitro*. *J Med Microbiol.* 49:713-718.

Van Hamme JD, Singh A & Ward OP (2006) Physiological aspects Part 1 in a series of papers devoted to surfactants in microbiology and biotechnology. *Biotechnol Adv.* 24:604–620

Van Hoogmoed CG, Dijkstra RJB, van der Mei HC & Busscher HJ (2006) Influence of biosurfactant on interactive forces between mutans streptococci and enamel measured by atomic force microscopy. *J Dent Res.* 85:54–58

Van Hoogmoed CG, van Der Kuijl-Booij M, van der Mei HC & Busscher HJ (2000) Inhibition of *Streptococcus mutans* NS adhesion to glass with and without a salivary conditioning film by biosurfactant-releasing *Streptococcus mitis* strains. *Appl Environ Microbiol*. 66:659-663.

Van Hoogmoed CG, Van der Mei HC & Busscher HJ (2004) The influence of biosurfactants released by *S. mitis* BMS on the adhesion of pioneer strains and cariogenic bacteria. *Biofouling,* 20:261–267

Vanittanakom, N, Loeffler W, Koch U & Jung G (1986) Fengycin–a novel antifungal lipopeptide antibiotic produced by *Bacillus subtilis* F-29-3. *J. Antibiot.* (Tokyo). 39:888–901

Velraeds MMC, van de Belt-Gritter B, Busscher HJ, Reid G & van der Mei HC (2000) Inhibition of uropathogenic biofilm growth on silicone rubber in human urine by lactobacilli - a teleologic approach. *World J Urol*. 18:422-426.

Velraeds MMC, Van der Mei HC, Reid G & Busscher HJ (1996) Inhibition of initial adhesion of uropathogenic *Enterococcus faecalis* by biosurfactants from *Lactobacillus* isolates. *Appl Environ Microbiol*. 62:1958-1963.

Velraeds MMC, Van der Mei HC, Reid G & Busscher HJ (1997) Inhibition of initial adhesion of uropathogenic *Enterococcus faecalis* to solid substrata by an adsorbed biosurfactant layer from *Lactobacillus* acidophilus. *Urology*. 49:790-794.

Vollenbroich D, Ozel M, Vater J, Kamp RM & Pauli G (1997a) Mechanism of inactivation of enveloped viruses by the biosurfactant surfactin from *Bacillus subtilis*. *Biologicals.* 25:289–297

Vollenbroich D, Pauli G, Ozel M & Vater J (1997b) Antimycoplasma properties and application in cell culture of surfactin, a lipopeptide antibiotic from *Bacillus subtilis. Appl. Environ. Microbiol*. 63:44-49

von Eiff C, Kohnen W, Becker K & Jansen B (2005) Modern strategies in the prevention of implant-associated infections. *Int J Artif Organs*. 28:1146–1156

Walencka E, Różalska S, Sadowska B & Różalska B (2008) The influence of *Lactobacillus acidophilus* derived surfactants on staphylococcal adhesion and biofilm formation. *Folia Microbiol.* 53:61–66

Worakitkanchanakul W, Imura T, Fukuoka T, Morita T, Sakai H, Abe M, Rujiravanit R, Chavadej S, Minamikawa H & Kitamoto D (2008) Aqueous-phase behavior and vesicle formation of natural glycolipid biosurfactant, mannosylerythritol lipid-B. *Colloids Surf B Biointerfaces.* 65:106–112

Xie Y, Li Y & Ye R (2005) Effect of alcohols on the phase behavior of microemulsions formed by a biosurfactant—rhamnolipid. *J Dispers Sci Technol.* 26:455–461

Xie Y, Ye R & Liu H (2006) Synthesis of silver nanoparticles in reverse micelles stabilized by natural biosurfactant. *Colloid Surf A-Physicochem Eng Asp.* 279:175–178

Xie YW, Ye RQ & Liu HL (2007) Microstructure studies on biosurfactant-rhamnolipid/*n*-butanol/water/*n*-heptane microemulsion system. *Colloid Surf A-Physicochem Eng Asp.* 292:189–195

Yakimov MM, Timmis KN, Wray V & Fredrickson HL (1995) Characterization of a new lipopeptide surfactant produced by thermotolerant and halotolerant subsurface *Bacillus licheniformis* BAS 50. *Appl Environ Microbiol.* 61:1706–1713

Yoo DS, Lee BS & Kim EK (2005) Characteristics of microbial biosurfactant as an antifungal agent against plant pathogenic fungus *J Microbiol Biotechnol.* 15:1164-1169

Zaragoza A, Aranda FJ, Espuny MJ, Teruel JA, Marqués A, Manresa Á & Ortiz A (2009) Mechanism of membrane permeabilization by a bacterial trehalose lipid biosurfactant produced by *Rhodococcus* sp. *Langmuir.* 25:7892–7898

Zaragoza A, Aranda FJ, Espuny MJ, Teruel JA, Marqués A, Manresa Á & Ortiz A (2010) Hemolytic activity of a bacterial trehalose lipid biosurfactant produced by *Rhodococcus* sp.: evidence for a colloid-osmotic mechanism. *Langmuir. 26*(11):8567–8572

Zeraik AE & Nitschke M (2010) Biosurfactants as agents to reduce adhesion of pathogenic bacteria to polystyrene surfaces: effect of temperature and hydrophobicity. *Curr Microbiol.* 61:554–559

15

Contact Lenses Characterization by AFM MFM, and OMF

Dušan Kojić, Božica Bojović, Dragomir Stamenković,
Nikola Jagodić and Đuro Koruga
University of Belgrade, Faculty of Mechanical Engineering,
Department of BioMedical Engineering
Serbia

1. Contact lense industry

The contact lens (CL) industry and market have displayed a high level of dynamism in the past few decades, and have evolved into a rapidly changing field in which science and everyday practice constantly interact, not only through broadening of material and product portfolio, but through innovative therapeutic and diagnostic solutions as well.

Stable market growth with numerous rearrangements in different product segments is constantly taking place, mainly stirred by innovative material and optical design. The standardly used hydrogel materials are being rapidly replaced by silicone doped hydrogel materials. The analyses of customer CL usage and satisfaction indicate continued market growth in future, however with many changes in product profile and significant increase in multifocal and daily disposable lenses market share.

The main impulse behind the dynamism of CL industry stems from results of scientific and technological improvements, which are enhancing medical field and reminding us that the focal point of sustainable development lies in scientific investigations.

2. Contact lenses in present, past and future

The technology of materials used in CL production has improved vastly in the past decades starting from glass and moving to polymer based materials (PMMA) with, eventually, major steps being taken in including hydrogel and doped-hydrogel based materials, shifting the functionality of CLs from rigid gas-impermeable (RGP) to soft gas-permeable materials represented by silicone hydrogel materials that are now in use.

This chapter will focus on multimodal applied research of rigid gas-permeable contact lenses (CL) that are manufactured from fluorosilicone acrylate based material. Our multimodal research comprises measurement of intermolecular interactions on the basis of optical, mechanical, morphological and magnetic properties of CL material.

The role of our research in such a complex system of CL industry was introduction of new diagnostic modalities through improved material characterization.

In the course of last decade, scientists have developed different possibilities of "on eye"CL application that are not related to its optical capabilities for which they were invented in the first place – correcting eye's refractive error. Furthermore, improvement in CL material manufacturing, both soft and rigid gas permeable, are mostly directed towards increasing oxygen permeability and wearing comfort. Rarely today, CL producers are dedicated to improving CL material properties for the purpose of enhancing the quality of vision, on the contrary, by doping them with silicone, for example, the optical properties become even worse.

Apart from it's properties to correct eye's refractive anomaly (dioptric power), the most frequent factors influencing quality of vision while wearing RGP CL are those related to the fact that visible light, on it's way to the eye's "perception area"– the macula, must pass through CL material itself, and all it's characteristics can seriously modify it.

Geometrical optics and related functions of the eye (vision acuity etc.) should not be considered as the only one mechanism of light interaction with human organism. We also consider physical inputs that influence the functioning of the central nervous system (CNS) on the basis of optical-neuronal interactions, and point to perspectives of investigating the role of CLs in modifying the influence of light for therapeutic purposes, or using the CLs as a potential diagnostic tool in monitoring the state of other systemic parameters (such as serum glucose level etc). As a rule, when comparing "naked eye"vision with the vision aided by CL it is inevitable to conclude that vision with CLs is of lower quality, due to reduced contrast sensitivity, sub-normal color's perception, spherical and chromatic light aberrations. All these are considered mostly consequence of CL material imperfectness. The aim of our research is to organize a setup for development of a novel material for RGP CL production which should, after adequately lathe cut and polished, improve it's optical properties in transmitting visible and "near visible"light, while increasing contrast sensitivity for "on eye"usage and improving color perception with simulaneous reduction in both higher and lower order light aberrations. New, advanced CL materials, are still needed because, regarding biocompatibility and oxygen permeability, the advantages offered by high oxygen delivery have solved many hypoxia-related clinical problems, but the complications related to inflammation, infection and mechanical insult to the cornea still occur. The good news is that the industry continues to work on the next innovative materials and designs while our patients, and us, enjoy the silicone hydrogel lenses and the big step they represent toward safer and more effective extended (toward continuous) contact lens wear.

3. Physiological considerations related to contact lenses wearing

3.1 Tear layer considerations

Tear layer is constantly renewed tripartite film that covers conjunctival and corneal surface. All layers are separated:

- **Mucin layer** is the innermost and it is anchored to the corneal and conjunctival epithelial glycocalix of the microplicae.
- **Aqueous layer** is the middle one and represents 90% of the tear film thickness. It consists of water and salts dissolved in it, together with dissolved glucose, lysosime, for tear's specific pre-albumin, lactoferin and other.
- **Lipid layer** is the outermost, secreted by the Meibomian glands and it retards evaporation of the tears.

The tear layers have many different functions:

- Keeping the surface of cornea smooth and that way making it optically clear.
- During blinking tear film lubricates the friction area between lids and ocular surface.
- Many of its constituents prevents ocular infection.

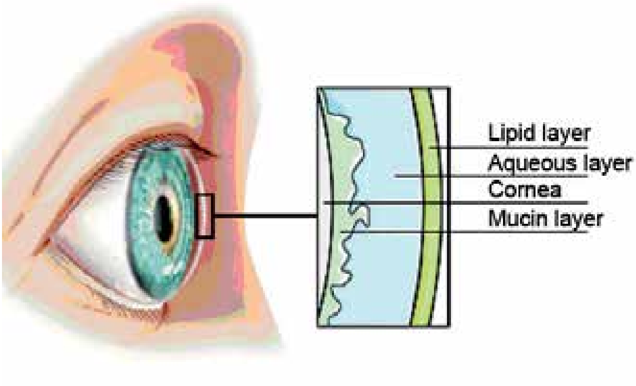


Fig. 1. Structure of sublayers that constitute tear layer.

Presence of any contact lens on the front surface of the eye influences physical and chemical properties of the tears and disturbs tear film stability. These changes in the healthy tear film are most likely caused by the following effects on the eye:

- Contact lens surface
- Contact lens design (mostly edges)
- Contact lens material
- Instillation of contact lens solutions with the lens

In comparison to the normal pre-ocular tear film, conventional (hydrogel) soft contact lens wearers have pre-lens tear film with all three layers of the tear film reduced. Tear film quickly deteriorates and renders poor surface wettability capability (weak attachment of ocular surface epithelium to the mucus layer). Pre-rigid gas permeable lens tear film is even thinner and more unstable then that of soft lens. Its lipid layer is often absent so the aqueous layer quickly dehydrates. In silicone-hydrogels, pre-lens tear film is something like combination of the previous two tear films although lipid layer is always present. As in the case of pre-ocular tear layer, good quality pre-lens tear film layer is *conditio sine qua non* for adequate optical properties of the eye.

Each eye consists of light receptor layer in the retina, optically active tissues that focuses the light on the receptor layer and nerve fiber system (retinal ganglion cell's fibers and optic nerve) which conducts electrical impulses created during electro-chemical reaction in the receptor layer (provoked by light absorption) all the way to the visual cortex and other parts of the brain.

Optically, eye functions very similar to photo-camera, inside the eye it is totally dark and this is provided by its outermost protection layer – the *sclera*, and pigmented parts of the middle layer – the *choroid*, which also provides most of the eye's blood supply. Light transmitting into the eye is absolutely controlled by the very agile diaphragm – the *pupil*. In order for light to be focused at the center of the eyes' reception layer it has to be refracted by *cornea*: absolutely transparent convex-concave lens and further refracted by another agile tissue in the eye - *the crystalline lens*. In order to prevent light reflection from behind the receptor layer to interfere with and disturb the primary light stimulation of the receptor cells, a pigmented cell layer is situated just behind the receptor layer and absorbs all the residual light.

Emmetropisation of the eye is a very sensitive process of dosing the eye's optical activity in relation to its axial growth. Any disturbances in this results in refractive errors of the eye (*myopia*, *hyperopia* and *astigmatism*). Eye, just as any other optical device, is also prone to higher order light aberrations that influence the stimulation of the receptor cells therefore influencing the quality of vision.

3.2 Light perception and visual signal transmission

Eye's light receptor cells (rods and cones) have large amounts of photo-sensitive pigments (*opsin* and *retinen*). Once stimulated by visible light (397 – 723 nm) these photo-sensitive pigments are changing their structure which leads to the chain of events that are ending in nerve activity. Rods are much more sensitive to the smallest amounts of light stimulation and they are most active in the dark environment (*scotopic* conditions) when pupil (the diaphragm) is dilated in order to receive as much light as possible. On the other hand, cones are much less light-sensitive and are active only in *photopic* conditions when the pupil shrinks in order to prevent too much light entering the eye which can disturb highly sophisticated visual functioning like color vision and small detail discrimination. Light stimulation blocks Na^+/K^+ channels in the receptor cell and disturbs the balance of ions in the sense of hyper-polarization of the cell. This results in reduction of the amount of the synaptic neuro-transmitter (normally released in certain amounts without light stimulation) that triggers electrical activity in the retinal ganglion cell which is conducted to the brain. Before it reaches the optical nerve, electrical activity created in the receptor cells is changed by the activity of the modulation cells present in different retinal layers.

4. Structure-function relationship of conventional and novel CL materials

Newer, soft lenses (hydrogel) are made from polymers that are inherently flexible or may become flexible through absorption of fluid into the polymer matrix. In general, when speaking of rigid CLs, most commonly used polymers are poly(methyl methacrylate) or PMMA, polyacrylamide (PAA), cellulose acetate butyrate (CAB), and in order to make them gas-permeable various mixtures based on PMMA doped with silicone or fluorine are used. Also, many modification of PMMA based materials are also present in the market (such as poly (2-hydroxyethyl methacrylate (PHEMA), poly (2-hydroxypropyl methacrylate) (PHPMA) etc.).

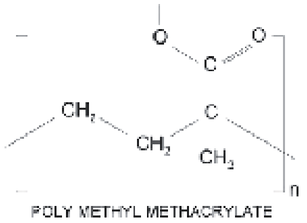


Fig. 2. Structure of poly methyl methacrylate.

The functions that contemporary contact lense materials are most often required to fulfill, encompass:

- **Dimensional stability** rendered through sufficient strenght and stiffnes – these properties are obtained by matrix of methylmethacrylate, which provides hardness and strength, and ethylene glycol dimethacrylate (EGDMA) which acts as a cross-linking agent adding to dimensional stability and stiffness but reducing water content.
- **Normoxia** rendered through material flexibility and gas permeability – these propeties are obtained by addition of several components, such as silicone (increases flexibility and gas

permeability through the material's silicon-oxygen bonds ($Si - O - Si$); however, it also brings in the disadvantage of poor wettability), fluorine (improves gas-permeability (less than Si) and improves wettability and deposit resistance in Si-containing lenses).

- **Proper adhesion** through controlled wettability – adequate level of adhesion is controlled by inclusion of hydroxyethyl-methacrylate – the basic water-absorbing monomer of most hydrogel-based soft lenses; methacrylic acid and n-vinyl pyrolidone (NVP) monomers are also added, both of which absorb high amounts of water and are usually adjuncts to hydroxyethyl methacrylate to increase lens water content.

Rigid contact lenses made of PMMA are recognized by excellent mechanical properties of dimensional stability (flexure resistance, stiffness and resistance to breakage) but practically have no oxygen permeability. Inclusion of Si and F has introduced gas-permeability properties but have compromised surface characteristics and stiffness. By chemically balancing silicone acrylate (SA) with stiff crosslinking monomers (esters) manufacturers have been able to achieve sufficieng gas-permeability without significantly compromising stiffness. Most contemporary gas-permeable lenses are composed of fluorosilicone acrylate that, due to oxygen's preference to dissolve into fluorinated materials, practically draws in the oxygen from the atmosphere and transports it towards the corneal surface, utilizing its $Si - O - Si$ component.

Hydrogel based lenses draw their advantage from the fact that hydrogels are materials that absorb and hold water inside their polymer matrices causing the spaces between the polymer chains to expand. Anything dissolved in the water can potentially enter the hydrogel matrix, depending on the molecular size and the matrix pore size. The pore size ranges from $0.5\mu m$ to $3.5\mu m$ for low and high-water content lenses, respectively. The oxygen permeability of hydrogel based contact lenses originates from their water content. Because various hydrogel polymers can greatly alter its chemical and physical properties, they may react differently to changes in pH, osmolarity, temperature and the components of the various lens care products that are used.

Silicone hydrogel based lenses use Si-doping which simultaneously decreases water content and increases permeability. Having in mind that the solubility of oxygen is very high in Si when compared with water, silicone doped hydrogel lenses permeability dramatically increases with lowering water content – which is a logical step towards improving functionality of CL materials. However optimal transport conditions for water and ion through a silicone lens, require an adequate amount of water. Synthesizing Si-hydrogel lenses is challenging since involves mixing non-polar (oxygen-rich silicone) and polar component (water) to produce lenses that renders high oxygen permeability, good wettability, flexibility, good optics, reasonable lens movement on the eye and general biocompatibility.

4.1 Conventional materials used in our investigation

We used two types of CL materials: gas-permeable CLs that are manufactured from fluorosilicone acrylate based material (Soleko $SP40^{TM}$) and standard non-doped PMMA CLs. The aim was to test the response of this material's surface roughness quality on the nano-level using standard nanotechnological methods and new nano-photonic method.

5. Characterization of contact lenses

Introduction of effective improvements in production quality as well as therapy efficiency requires the application of non-destructive surface analysis methods on the nanometer scale with minimal sample preparation. Many interesting studies have been performed on CL (Lim et al., 2001), (Kim et al., 2002), (Bruinsma et al., 2003), so that two research approaches can be perceived from the literature: one aimed at surface modification and the other that was concerned with varying or simulating exploitation parameters. In both cases morphology of CL surface served as the main parameter used to describe behavior and quality of CL material under deterioration influences.

Production engineering aspects. Conformation states of polymers constituting CL surface are changed during final stages of manufacturing process (polishing) which presents a complex problem because surface molecules and their orientation influence the final state of surface quality. In general, over-polished internal surface renders a dysfunctional product since it prevents adequate adherence to corneal tissue. CL with over-polished surface slides over the cornea and cannot maintain its initial position therefore disrupting the geometry and function of the optical system (human eye + CL). However, too rough a surface will eventually lead to irritation and possibly damage of corneal tissue.

Exploitation aspects. During CL exploitation a lacrymal film is formed between inner CL surface and cornea. A good fit of CL and cornea depends on the adequacy of CL geometry and surface roughness. Moreover, every single RGP CL wearer provides a unique ambient conditions in which these CL biosurfaces have to function properly. Since CL surfaces become significantly rougher after prolonged wear, they become more prone to bacterial adhesion and protein and lipid deposits. CL can lose functionality due to accumulated proteins, lipids, and other tear components on CL surface, despite routine cleaning activities. The loss of RGP CL functionality needs to be investigated and related to measurable parameters in order to recommend replacement based on significant changes in surface properties.

5.1 Multimodal, complementary and non-invasive approach

Answering the questions of optimal processing parameters directly influences surface quality of CLs which in turn reaches all aspects of its functionality: optical, medical and patient-comfort. Since our measurement confirmed that surface roughness of end-product ranges in the order of several tens of nanometers, the adequate tools for such samples must be derived from applied nanotechnology.

Experimental design, aimed at quantifying surface quality on the level of nano scale, utilized three methods of characterization that are meant to complement each other in interpreting and quantifying the measurement results. Moreover, we have applied a new method that is concerned with obtaining magnetic properties of samples from the interaction of material with visible light, named Opto-Magnetic Fingerprint of matter (OMF), with the aim to improve the speed and accuracy of analysis. Testing the product quality, reliabely and accurately, in production environment is an essential characteristic of maintaining high level of product quality control.

We have chosen to investigate near-surface magnetic and optical properties of chosen CL materials since they are the physical quantities that most closely correlate with subtle modifications in material structure and composition. Therefore, we investigated five types of CL materials (two conventional and three nanophotonic materials) by selected three approaches (techniques):

1. **Classical methods**: spectroscopic examination in the region of ultraviolet and visible light (UV-Vis spectroscoscopy).
2. **Nanotechnology based methods**: Phase-Contrasted Atomic Force Microscopy (PC-AFM) with extended mode of Magnetic Force Microscopy (MFM).
3. **Nanomedicine based methods**: Novel method named Opto-Magnetic Fingerprint (OMF) that obtains magnetic properties of materials on the basis of interaction with visible light.

The experiment was designed to provide proofs for three interrelated phenomena:

- Change in topographical and conformation states of surface and polymer via PC-AFM,
- Related change in nanomagnetic behavior of near surface layers via MFM,
- The change of paramagnetic/diamagnetic and optomagnetic state that is recorded by OMF.

All three methods possess very high sensitivity (nanonewton for forces, nanotesla for magnetization), which is a necessity for this type of measurements.

5.2 Ultraviolet-visible spectroscopy

In order to investigate optical properties of RGP CLs, we have performed spectroscopic measurements in the range of ultraviolet-visible (UV-Vis) light in the range of wavelengths between 280–800 nm, using the UV-Vis scanner produced by Horiba JobinYvon, USA. The measurements were conducted in diffuse reflectance regime and the result is displayed as a graph of reflected energy vs. wavelength of emitted radiation. The spectroscopic measurements are used as a reference guide and were performed on two locations: the central point on the outer surface and the same point measured from inner surface in order to determine the difference and hence the degree of eye protection to different ranges of wavelengths and radiation intensities.

5.3 Phase-contrasted atomic force microscopy – PC-AFM

Morphology of CL surfaces was obtained by atomic force microscope (PC-AFM) that can measure sample surface roughness with high precision (less than $10^{-12}m$ or $\approx 1pm$ and confirm sample surface state as belonging to either group adequate or inadequate roughness. Topography of lens surface is important for determining the connection between surface morphology (conformational states of surface layer polymers) and corresponding optical properties that are influenced by the processing parameters.

Basically, AFM is a scanning probe microscopy technique based on point-to-point examination of the specimen made by a sharpened tip probe (Binnig et al., 1986). AFM probe is a micro-cantilever with sharpened conical or pyramidal tip whose radius can range from 2 – 90 nm, depending on the application. All samples were imaged using phase-contrast technique in tapping mode and in ambient air. The AFM system used in this study was JSPM-5200, JEOL, Japan. The cantilever were type PPP-MFMR, produced by Nanosensors, Switzerland.

The principle of operation of AFM is shown in figure 3. In AFM, the probe (cantilever tip) is vibrated at near-resonant frequency and brought ino interaction with the sample by the mechanism of intermolecular attractive/repulsive forces that are distance-dependent. The cantilever is maintained in the close proximity from the sample so that probe tip is within reach of attractive/repulsive forces. A typical AFM system is able to detect intermolecular forces in the order of $10^{-11} - 10^{-13}N$ which makes it an extremely sensitive device. The intermolecular interactions belong to the class of van der Waals type forces, usually modeled by Lennard-Jones potentials. These forces cause the cantilever to deflect from the initial equilibrium position making it possible to derive the distance from the sample on the basis of force field gradient change that modulates the vibration of the cantilever.

A diode laser is directed to the back sided surface of the tip and is reflected to photo-diode detector. During the scanning movement, the angle of reflected beam is changed due to deflection of the cantilever (that is, in turn, due to interactions with the sample) and this movement is precisely recorded via photo-diode detector. A feedback loop is used to adjust the z-position of the sample so that a constant interaction is maintained as the tip is scanned across the sample in x and y directions. The tip-sample interaction equilibrium is constantly disrupted by changes in sample profile height which generates the control signal applied

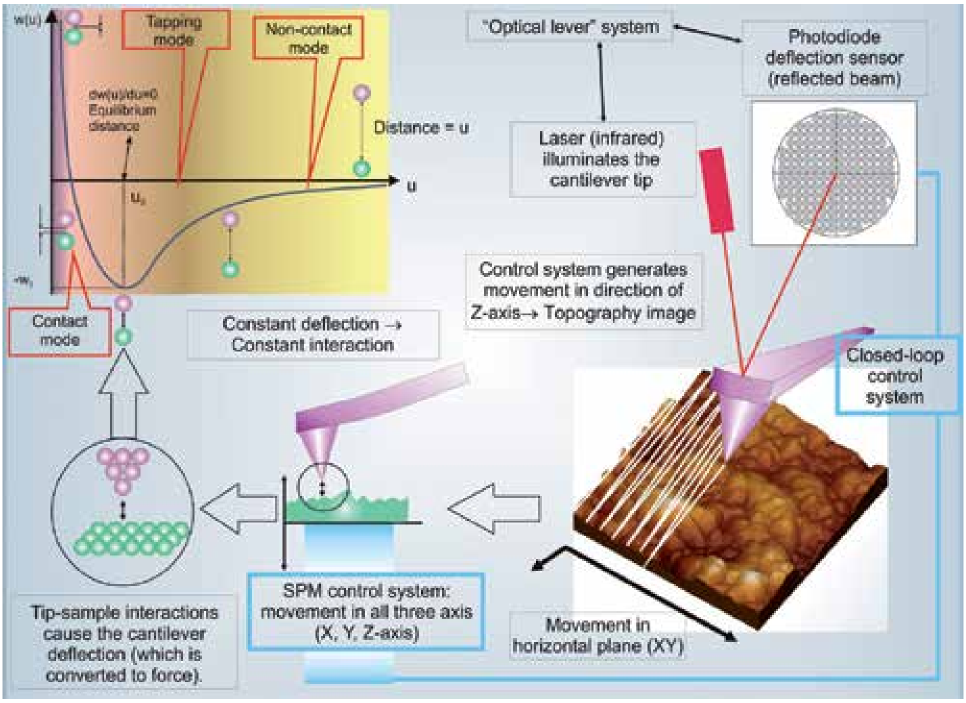

w(u)
Tapping mode
Non-contact mode
dw(u)/du=0
Equilibrium distance
Distance = u
u
u_0
$-w_0$
Contact mode
Constant deflection → Constant interaction
"Optical lever" system
Photodiode deflection sensor (reflected beam)
Laser (infrared) illuminates the cantilever tip
Control system generates movement in direction of Z-axis→ Topography image
Closed-loop control system
Tip-sample interactions cause the cantilever deflection (which is converted to force).
SPM control system: movement in all three axis (X, Y, Z-axis)
Movement in horizontal plane (XY)

Fig. 3. Principle of operation of atomic force microscope.

to the z-scanner stage. This disturbance represents the changes in surface topography and the action of the control system is recorded and represents the actual data represented on topography image of the sample.

The PC-AFM imaging technique is based on the fact that intermolecular force gradients have certain physico-chemical specificity. The variations in the sample generate variations in the slopes of Lennard-Jones potential curves because of the different intermolecular forces acting on the AFM sensor tip. These differences modulate the vibration amplitude of the AFM cantilever, creating a higher-harmonics in the feedback signal of AFM system. By recording higher harmonics in the oscillation signal (see 3) we can detect different force-distance relationships, hence, different components of a material can be visualized. However, the PC-AFM signal correlation to separate chemical species is still unknown, which disables the AFM from making exact chemical characterization of the sample. The data obtained by PC-AFM are rather of a qualitative nature and are suitable for two or three component systems.

5.4 Magnetic Force Microscopy – MFM

Magnetic Force Microscopy is an extended operation mode of AFM that was used to obtain the distribution of magnetic properties of the surface that are previously imaged with topography mode. This tecnique was used to measure the magnetic properties in para- and diamagnetic range because we are also probing the magnetic properties in the same range with a novel technique that is introduced in the next section (Opto-Magnetic Fingerprint). MFM will is used as a comparative method for OMF. The MFM technique utilizes special cantilever sensors that are coated with a thin film of cobalt ($\approx 50nm$) that renders its ferromagnetic properties and ensures magnetic interaction with the sample.

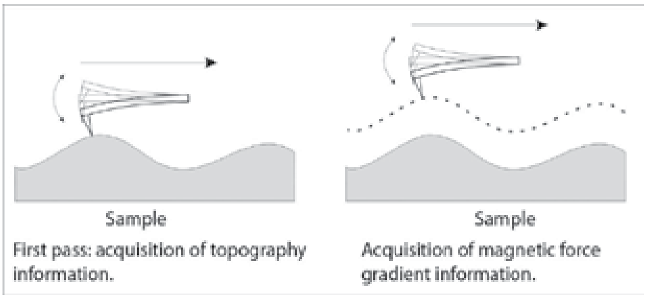

Sample
First pass: acquisition of topography information.
Sample
Acquisition of magnetic force gradient information.

Fig. 4. Principle of operation of atomic force microscope.

This thin film creates magnetic interaction with the sample that is recorded via two-pass technique or "lift scanning"(see 4) during which the sample is scanned twice, with second scan being performed with a gap distance in order to physically filter the slowly decaying magnetic from all other more rapidly decaying intermolecular forces. The sensitivity of magnetic force measurements goes around $10^{-11}N$ which is able to detect very weak magnetic interaction that encompasses para- and diamagnetic range. The cantilevers which we used in this study are produced by Nanosensors (Switzerland), with a lateral resolution under 50 nm and coercivity of 300 Oe.

The cantilever has been magnetized previously to the measurement in order to ensure its magnetic interaction with the sample. Although MFM, in its current stage of development, still represents a qualitative measurement of magnetic properties, we attempted to improve the accuracy of this analysis by measuring the remnant magnetization of the cantilever, using JR-5 spinner magnetometer (AGICO, Brno, Czech Republic), and identified that our series of measurements fitted within the range of $72 \pm 2nT$, with a standard deviation of 0.3%. The measured axis of cantilever remnant magnetization was positioned in the vertical direction, so it was perpendicular to sample surface. The data on components of vectors of remnant magnetization enabled us to confirm that we have measured the magnetic field component that is perpendicular to the sample surface.

5.5 Opto-Magnetic Fingerprint – OMF

OptoMagnetic Fingerprint (OMF) is a novel method in investigating optical and magnetic properties of materials that is based on electron properties of matter (covalent bonds, hydrogen bonds, ion-electron interaction, van der Walls interaction) and its interaction with light (Koruga et al., 2008). The method was originally developed for early skin cancer and melanoma detection by MySkin, Inc., USA (Bandić et al., 2002). Bearing in mind that the orbital velocity of valence electron in atoms is about $10^6 m/s$, this gives the ratio between magnetic force (F_M) and electrical force (F_E) of matter, of around $\frac{F_M}{F_E} \approx 10^{-4}$. Since force (F) is directly related to quantum action – Planck's action – defined as:

$$h = F \times d \times t = 6.626 \times 10^{-34} Js \qquad (1)$$

where F is force, d is displacement and t is time of action) this means that the action of magnetic forces is four orders of magnitude closer to quantum action than the electrical ones. Since quantum state of matter is primarily responsible for conformational changes on the molecular level, this means that detecting differences between matter states is by far more likely to give greater sensitivity on the level of magnetic forces than it would be on the level of measurement of electrical forces (Koruga et al., 2002).

Picture of surface that is taken by classical optical microscope is based on electromagnetic property of light, while OMF is based on difference between diffuse white light (like that of daily light) and reflected polarized light. Reflected polarized light is produced when source of diffuse light irradiates the surface of matter under certain angle (Brewster's angle, see figure 5). Each type of matter has different angle value of light polarization. Since reflected polarized light contains electrical component of light-matter interaction, taking the difference between white light (electromagnetic) and reflected polarized light (electrical) yields magnetic properties of matter based on light-matter interaction. Because such measurement can identify the conformational state and change in tissue on molecular level we named this method the opto-magnetic fingerprint of matter (OMF).

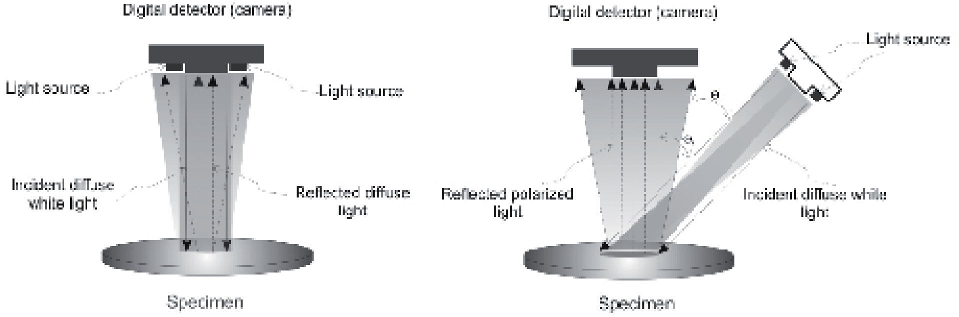

Digital detector (camera)
Light source
Light source
Incident diffuse white light
Reflected diffuse light
Specimen
Digital detector (camera)
Light source
θ
Reflected polarized light
Incident diffuse white light
Specimen

Fig. 5. Incident white light can give different information about thin layer of matter (surface) properties of sample depending on the angle of light incidence. When the incident white light is diffuse, the reflected white light is then composed of electrical and magnetic components, whereas diffuse incident light that is inclined under certain angle will produce reflected light which contains only electrical component of light. For each type of matter there is a characteristic angle of incidence (Bandić et al., 2002) for obtaining the appropriate reflected polarized light.

We used digital images in RGB (R-red, G-green, B-blue) system in our analysis, therefore we chose basic pixel data in red and blue channels for white diffuse light (W) and reflected polarized white light (P). Algorithm for data analysis is based on chromaticity diagram called "Maxwell's triangle"and spectral convolution operation according to ratio of (R-B)& (W-P) (Koruga et al., 2008). The abbreviated designation means that Red minus Blue wavelength of White light and reflected Polarized light are used in spectral convolution algorithm to calculate data for opto-magnetic fingerprint of matter. Therefore, method and algorithm for creating a unique spectral fingerprint are based on the convolution of RGB color channel spectral plots generated from digital images that have captured single and multi-wavelength light-matter interaction (Koruga et al., 2008). Preparation of digital pictures for OMF was made by usage of dermoscopic imaging device (MySkin, USA) that has previously been successfully used in biophysical skin characterization (skin photo type, moisture, conductivity, etc) (Bandić et al., 2002).

The final purpose of our research in applying OMF is the construction of quality control method which would be purely optical and able to, on the basis of digital image analysis and processing, detect both the morphology and functionality parameters in a quicker and more accurate manner. In order to do so we need to construct quantification parameters but, in this stage of research, primarily integrate results from morphological research and opto-magnetic properties.

6. Results

6.1 Topography measurement using Atomic Force Microscope – AFM

Topography measurements were routinely conducted in tapping mode in ambient air using uniform scanning surface of $5 \times 5\mu m$. The CL inner surface has been examined as shown in figure 6. The curvature of the surface prevents the AFM probe to reach its center, unless the sample is destroyed. Nevertheless, a good approach was able to the area that is approximately on the half distance of CLs radii. A total of four points were selected for scanning on each CL in two perpendicular directions. In each point two scans were conducted: one with $60 \times 60\mu m$ surface size and the other with the $5 \times 5\mu m$ surface size. The purpose of larger area scans were to confirm the uniform character of surface morphology while smaller scans were further analysed by fractal analysis.

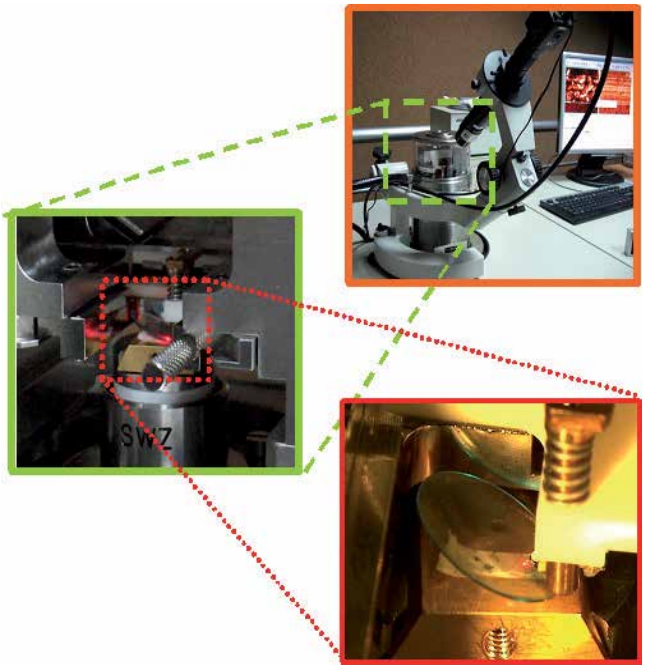


Fig. 6. Experimental setup used in AFM imaging of CL surfaces.

6.2 Phase-contrasted AFM

Phase contrast images are combined with topography images since these two measurements are performed simultaneously. Topography image yields information about surface shape and relative positions and dimensions. Phase contrast image enrichens that information by differentiating between different force gradients which in turn point to different conformational states of polymers, under the assumption that material constituents are homogenously distributed throughout the sample. Phase-Contrasted image of the CL is shown on figure 7.

6.3 Magnetic force microscopy

The AFM/MFM measurements display the features of surface morphology and magnetism on the nanometer scale and are shown on two images in figure 8. Since the sensitivity of

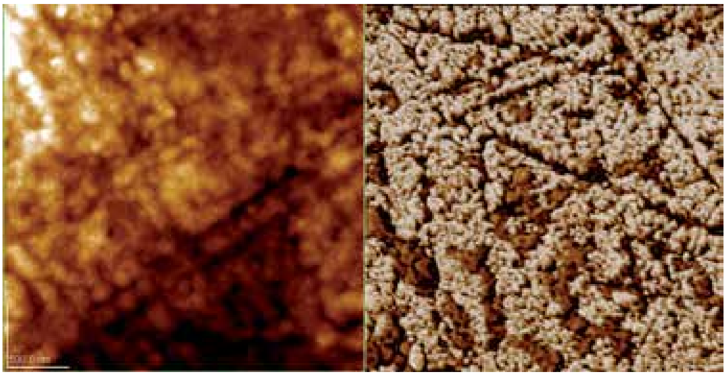


Fig. 7. Phase-Contrasted Atomic Force Microscopy: scan size is $5 \times 5\mu m$. Left: topography image with maximum profile height of 150.9 nm. Right: phase contrasted image for the same portion of the sample that is shown on the left image. Phase contrast image shows granular inhomogeneity in the sample that is synthesized as homogenous (on the nano-scale). The inhomogeneities may originate only as a consequence of processing, showing that certain parts of surface have polymers with altered conformation states, thereby expressing different slope of intermolecular interactions.

forces measured by cantilever go well below nanonewton range, this means that we are able to record paramagnetic and diamagnetic behavior of material. The brighter image areas mark more highly responsive magnetic behavior or paramagnetic areas while than darker areas correspond to diamagnetic areas of the sample.

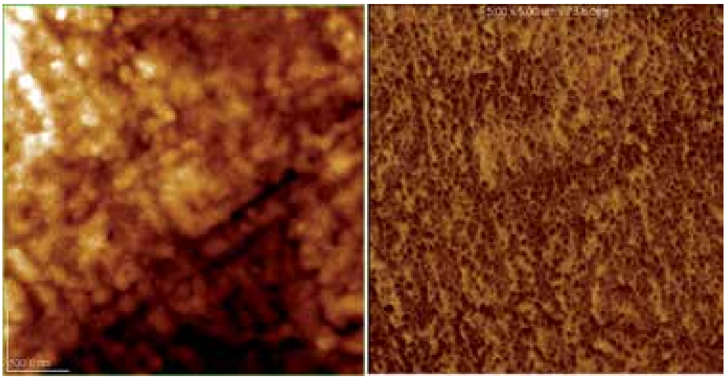


Fig. 8. Magnetic Force Microscopy: scan size is $5 \times 5\mu m$. Left: topography image with maximum profile height of 150.9 nm. Right: the magnetic force gradient image of the same area. This image shows changes in magnetic force gradient, exhibiting that magnetic behaviour exist on the para- and diamagnetic levels and that they are inhomogenous in our samples (that are chemically homogenous).

6.4 Opto-magnetic fingerprint and UV-vis spectroscopy

Digital images of contact lenses were analyzed in terms of their separate color channels (red, blue and green color components) and subsequently processed by spectral convolution algorithm to give the final result – OMF diagram – which shows the intensity values of paramagnetic (positive) and diamagnetic (negative), properties in comparison to wavelength difference. The diagrams on figure 9 show comparison of OMF diagrams with classical UV-Vis-NIR for two samples with different surface qualities.

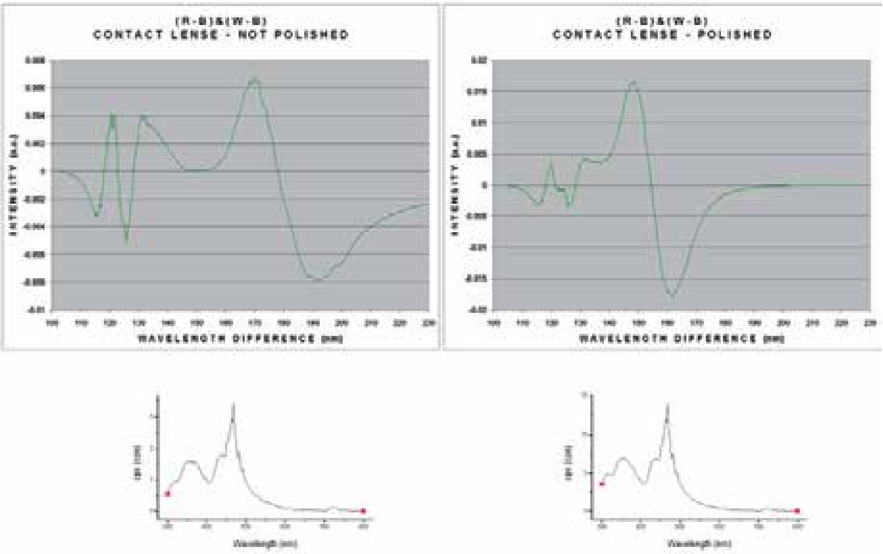


Fig. 9. Opto-Magnetic Fingerprint: comparisons of OMF (upper two diagrams) and UV-Vis spectras (lower two diagrams) for differently surface-treated contact lenses that were produced from the same, standard, material. The characteristic diagrams for lathe processed (left) and polished (right) CL surfaces. UV-Vis spectras show almost identical patterns of absorption while OMF diagrams show visible differences in the wavelength difference range between 100 and 150 nm. These differences are due to different processing and can be quantized by ratio of paramagnetic to diamagnetic properties.

The OMF diagrams show the distribution of energies of reflected light that are distributed over a wavelength difference range. We can observe a qualitatively same pattern, except in the subregion between 100–150 nm, which shows variations in wavelengths and intensities. These diagrams show very high sensitivity of OMF imaging and software analysis, which is the exact purpose this method was intended to achieve.

Moreover, we have tested the OMF response for two inner surfaces of RGP CLs of type $Boston^{TM}$ with respect to changes in the surface qualities. One pair of lenses was processed differently (one was while the other was not polished) while the other pair was processed identically (same polishing times).

6.5 Fractal analysis of contact lenses' roughness profiles

Currently we have no clear answer to the question what surface standard parameter should be used for critical limit determination. The result of the study, reported in (Bruinsma et al., 2003), also confirms the need for quantitatively establishing the replacement schedule of RGP CL. The water contact angle, percentage of elemental surface composition and deposit rate of bacteria was related to standard average roughness parameter R_a. It was stated in (Kim et al., 2002) that surface roughness was the most influential lens surface property after 10 days of wear.

Our research has utilized fractal analysis of roughness profiles to quantify the texture properties of CL inner surface. Standard surface parameters (roughness descriptors) fails

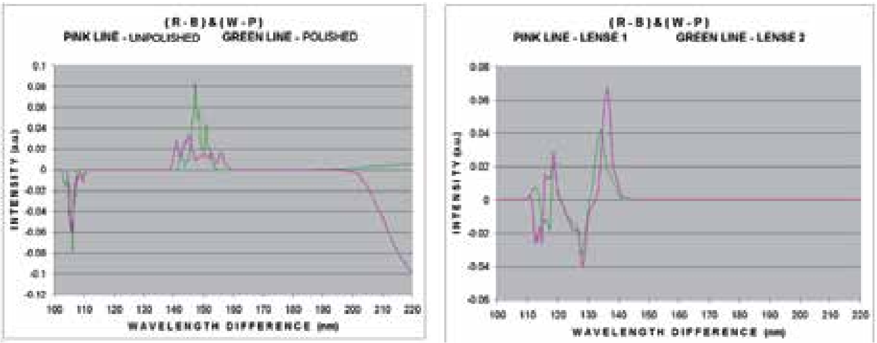


Fig. 10. Opto-Magnetic Fingerprint diagrams. Image on the left: overlapped OMF diagrams of two differently processed CLs that show different behaviour in the wavelength difference range of 130–150nm and sensitivity to changes in surface qualities. Image on the right: overlapped OMF diagrams of two identically processed CLs that show minor change in diagrams and surface qualities.

to describe functional nature of the surface. Moreover, use of more than one roughness parameter exhibits more shortcomings. This is mainly due to the partial information contained in each descriptor.

Fractal analysis of biomedical surface topography is influenced by growing interest in biomaterials surface technology. Fractal geometry provides a useful tool for the analysis of complex and irregular structures such as biomedical surface topography based on image analysis methods that consider an image as a 3D surface.

Fractal dimension calculation is based on "slit-island"and "skyscrapers"methods that were proposed in (Bojović, 2008). This method stipulates that surface recording data is obtained as an image, by using scanning probe microscope. Fractal analysis consists of the following steps:

1. Conversion of AFM image to numerical data in the form of matrix with subsequent conversion of matrix with 256 levels to matrixs with 2^{16} levels of intensity $[0, 65535]$ needed for further calculations.

2. Calculation of image surface area by well known method called "skyscrapers"method. This method approximates surface area of image A with sum of top squares that represent skyscrapers' roofs and the sum of exposed lateral sides of skyscrapers, according to (Chappard et al., 1998). The roof of skyscrapers are increased subsequently by grouping of adjacent pixel grouping. Thus, the intensities of grey scale are averaged. The square size ε is 2^n and the formula is as follows:

$$A(\varepsilon) = \sum \varepsilon^2 + \sum \varepsilon[|z(x,y) - z(x+1,y)| + |z(x,y) - z(x,y+1)|] \qquad (2)$$

3. According to (Bojović, 2008), the fractal dimension D can be generated from relation 3 for Hausdorff-Besicovitch dimension where N(ε) is the number of self-similar structures of linear size ε needed to cover the entire structure. Number N(ε) can be represented as shown in 4 and used for the area vs. square size relationship 5 resulting in equation 6. Using of logarithmic rules on relation 6 result in a linear equation, expressed as 7. Fractal dimension D is obtained as the slope of fitted line, determined by using relation 7 in the

custom-made procedure for calculation 11.

$$D = \lim_{\varepsilon \to \infty} \frac{\log N(\varepsilon)}{\log \frac{1}{\varepsilon}} \tag{3}$$

$$N(\varepsilon) = c_1 \varepsilon^{-D} \tag{4}$$

$$A(\varepsilon) = N(\varepsilon)\varepsilon^2 \tag{5}$$

$$A(\varepsilon) = c_1 \varepsilon^{2-D_s} \tag{6}$$

$$\log A = (2 - D_s) \log \varepsilon + c \tag{7}$$

4. The log-log graph of surface vs. square size ε is then produced and line is fitted for each image (see figure 11).
5. Finally the fractal dimension is generated by skyscrapers method for AFM topography image as slope of the fitted line. This slope is steeper for rougher contact lens surface.

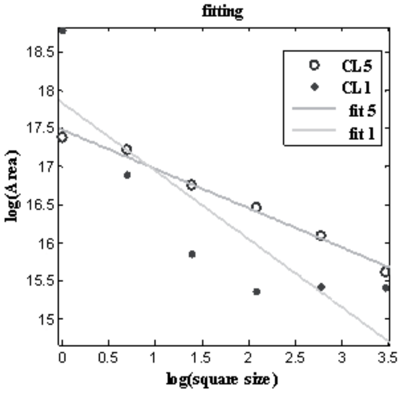


Fig. 11. The log-log graph of image area vs. square size for two CLs. Samples are two RGP CLs made of ML 92 Siflufocon A. The first lens (CL5) was worn for about 3 years while the second lens was worn over a period of more than 5 years (CL1). We can see that fractal dimensions can easily distinguish between two levels of surface roughness created by wear of CL surface.

The procedure is schematically presented in the figure 12 .
Mandelbrot claimed that nature has a fractal face and scholars proved that engineering surfaces have fractal geometry. Compilations of a man-made surface with a tear component on it also show fractal behaviour, proven by power law of area vs. scale relationship. According to (Russ, 1998) a surface with fractal dimension 2.5 would be the optimum as an engineering surface for certain applications.
The fractal dimension generated by skyscrapers method for topography image offers additional and appropriate information about surface roughness. Fractal dimension, as roughness parameter, adequately explains surface functional behaviour. Fractal dimension for new contact lens surface could be an adequate behaviour prediction parameter.

7. Discussion

By performing UV-Vis spectroscopy we have shown that UV-Vis spectra do not change with respect to changes in surface quality (see figure 9) because they measure bulk response of

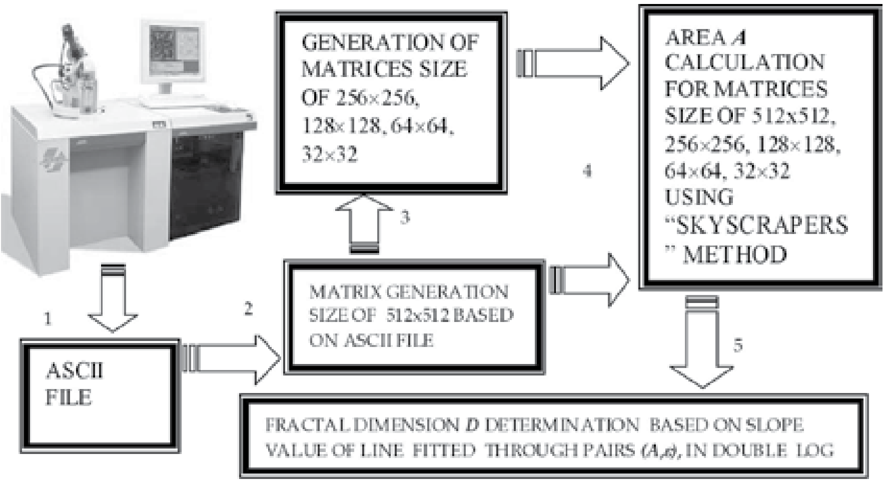

GENERATION OF MATRICES SIZE OF 256×256, 128×128, 64×64, 32×32
AREA A CALCULATION FOR MATRICES SIZE OF 512x512, 256×256, 128×128, 64×64, 32×32 USING "SKYSCRAPERS" METHOD
4
3
MATRIX GENERATION SIZE OF 512x512 BASED ON ASCII FILE
1
2
ASCII FILE
5
FRACTAL DIMENSION D DETERMINATION BASED ON SLOPE VALUE OF LINE FITTED THROUGH PAIRS (A,a), IN DOUBLE LOG

Fig. 12. Diagram of steps involved in fractal analysis of AFM scans.

contact lense materials while opto-magnetic diagrams (OMF) displayed informations that are more sensitive to changes in near-surface properties of material (which are primarily altered during contact lense production). OMF has shown as method that can detect very sensitive to differences in topographical features of contact lenses.

Since, conformation changes in near-surface polymer molecules generate quantum effects they might influence magnetic properties as well. Because of that, we have investigated near-surface regions of contact lenses samples from magnetic and optomagnetic point of view to see whether is there exist a measurable difference in surface magnetic and optical properties. Our aim was to explore the relationship between surface morphology on one side and optical and magnetic properties on the other side.

Measurements on the nano-scale have shown that phase-contrasted atomic force microscopy (PC AFM) and magnetic force microscopy imaging carry additional information that is not contained in morphology scans. However, PC AFM and MFM data need to be integrated with results of OMF in order to obtain quick quantitative assessment of changes in nano- and pico-magnetism that can be related to change in surface structure and its optical properties. Elucidating the origin of these kinds of surface behavior requires further investigations and inclusion of other polishing process parameters on one side and quantitative MFM measurements on the other side. It is our opinion that this kind of analysis will be able to precisely determine parameters of final shape and performance of CL surface.

Since conformation states of RGP CL surface determine paramagnetic diamagnetic properties that can be detected by novel OMF technique, we consider the OMF method and molecular level approach to investigation of optical properties of CL quality as a very promising field for both basic research and technology, with direct influence on application in biomedicine.

New methods of investigation require novel data processing techniques. Fractal analysis has offered more sophisticated tool that, on the basis of nano-scale precision information, generates finer estimate of surface roughness quality.

8. Conclusion

Light has influence on brain activity with very complex pathway (see figure 13). Since light is composed of electrical and magnetic spectra it is very important to know how light interact with contact lenses. These aspects of brain functioning have been thoroughly investigated;

vast amount of informations is available but the most subtle biophysical aspects are still not completely understood.

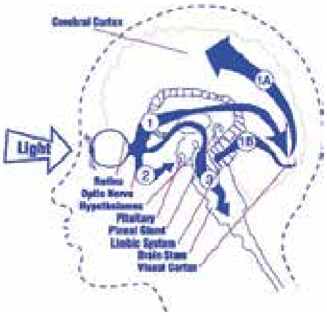


Fig. 13. Nerve pathways from the eyes to the brain goes not only to the visual cortex, but also to deeper brain areas, concerned with neurotransmiters, neurohormones, emotions, etc.

Visual perception is the ability to interpret information from information contained in visible light that reaches the eye. The act of seeing starts when the lens of the eye focuses an image of its surroundings onto a light-sensitive membrane in the back of the eye, called the retina. Since visible light is composed of electrical and magnetic spectra, which have different influence on brain activity (EEG and MEG signals, see figure 14), we investigated magnetic property of contact lenses, as optical material, which have influence on electrical and magnetic light signals properties.

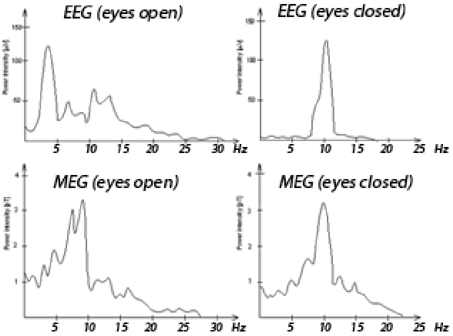


Fig. 14. Brain activity (EEG and MEG) under light influence when eyes are open and closed.

The findings of our measurements enable us to couple optical and magnetic behavior analysis in determination of mechanical properties of surfaces. The dynamical structure of CL materials and its behavior under dynamical mechanical and thermal load is expressed in changes in paramagnetic/diamagnetic behavior. Our results show that these changes are measurable and can be quantified by a simple, quick and accurate method of OMF.

We believe that this combination of intertwining methods could yield an optimal approach towards investigating phenomena in material synthesis and its behavior under mechanical and thermal stress that is still not well understood. The grounds for our propositions are proven relationships between optical, magnetic and mechanical properties of matter. It is our intention to further improve the application of all three used methods and customize their parameters in order to combine them into a new device for opto-magneto-spectroscopic characterization of matter.

Furthermore, our future research will involve nanomaterials as a new doping material influencing CL physical properties such as light transmission, these changes will be investigated by by UV-Vis-NIR spectroscopy as well as optomagnetism. The potential application of nanomaterials might bring significant biophysically based implications for contact lenses industry, biomedical application industry and applied optical science.

9. References

Lim, H., Lee, Y., Han, S., Cho J., Kim, K.J., Surface treatment and characterization of PMMA, PHEMA, and PHPMA, *Journal of Vacuum Science and Technology A*, Vol. 19, No. 4, 2001, pp. 1490–1496, ISSN 0734–2101.

Kim S.H., Opdahl. A., Marmo, C., Somorjai, G.A., AFM and SFG studies of pHEMA-based hydrogel contact lens surfaces in saline solution: adhesion, friction, and the presence of non-crosslinked polymer chains at the surface, *Biomaterials*, Vol. 23, No. 7, 7 September 2001, pp 1657–1666, ISSN 0142–9612.

Bruinsma, G.M., Rustema-Abbinga M., de Vriesa, J., Busschera, H.J., van der Lindenb, M.L., Hooymansc, J.M.M., van der Meia, H.C., Multiple surface properties of worn RGP lenses and adhesion of Pseudomonas aeruginosa, *Biomaterials*, Vol. 24, No. 7, 2003, pp 1663–1670, ISSN 0142–9612.

Binnig, G., Quate, C.F., Gerber, C. (1986), Atomic Force Microscope, *Physical Review Letters*, Vol. 56, No. 9, 3 March 1986, pp. 930–933, ISSN 0031–9007.

Koruga, Đ., Tomić, A., Ratkaj, Z., Matija, L., Classical and Quantum Information Channels in Protein Chain,*Material Science Forum*, Vol. 518, 2006, pp. 491–496, ISSN 1662–9752.

Koruga Đ., Tomić A, Method and algorithm for analysis of light-matter interaction based on spectral convolution, US Pat. App. No.61/061,852, 2008, PCT/US2009/030347, Publication No: WO/2009/089292, Publication Date: 2009-07-16.

Bandić, J., Koruga, Đ., Mehendale, R., Marinkovich, S., System, device and method for dermal imaging, US Pat. App. No. PCT/US2008/050438, Publication No: WO/2008/086311, Publication Date: 2008-07-17.

Bojović, B., Miljković, Z., Babić, B., Fractal Analysis of AFM Images of Worn-Out Contact Lens Surface, *Faculty of Mechanical Engineering Transaction*, 36, 4, 175–180, 2008, ISSN 1451–2092.

Bojović, B., Investigation of Interaction of Engineering Surfaces Condition and Fractal Geometry, PhD Thesis in Serbian, University of Belgrade, 2009.

Chappard, D., Degasne, I., Hure, G., Legrand, E., Audran, M., Basle, M.F., Image analysis measurements of roughness by texture and fractal analysis correlate with contact profilometry, *Biomaterials*, Vol. 24, 2003, pp. 1399–1407, ISSN 0142–9612.

Russ, J.C., Fractal Dimension Measurement of Engineering Surface, *International Journal of Machine Tools and Manufacturing*, Vol. 38, No. 5-6, 1998, pp. 567–571, ISSN 0890-6955.

16

Synthesis and Characterization of Amorphous and Hybrid Materials Obtained by Sol-Gel Processing for Biomedical Applications

Catauro Michelina and Bollino Flavia
Department of Aerospace and Mechanical Engineering,
Second University of Naples, Aversa,
Italy

1. Introduction

An interesting research field with medical applications is represented today by ceramics, as they can be used to obtain useful biomaterials for the production of implants (Vallet-Regí, 2001, 2006a, 2006b, 2006c); many parts of the human body, in fact, can be replaced or repaired with biomaterials and more specifically with bioceramics (Black & Hastings, 1998). Regardless of the ceramic type and the application procedure, the introduction of an implant in a living body always causes inflammation phenomena and frequently infection processes as well. Those problems can be overcome by using local drug delivery methods to confine pharmaceuticals such as antibiotics, anti-inflammatory, anti-carcinogens, etc. (Arcos et. al., 2001; Ragel & Vallet-Regì, 2000; Vallet-Regì et al., 2000). The possibility of introducing certain drugs into the ceramic matrices employed for bone and teeth repair is undoubtedly an added value to be taken into account.

The traditional use of high temperature procedures to model glasses and ceramics to the desired shape is very well known; on the other hand, the degradation temperature of a pharmaceutical compound is usually around 100°C, which is very low if compared with the high ones needed to compact the components (around 1000°C). Consequently, the main problem is how to include pharmaceuticals in conventional glass and ceramic implants. The scientific community is currently investigating new procedures to incorporate drugs into implantable biomaterials.

The sol-gel process, among others, has proved to be a versatile one and has been widely used in the preparation of amorphous and or hybrid materials (Hench & West, 1990; Judeinstein & Sanchez, 1996; Novak, 1993), with applications, for example, in non-linear optical materials (Hsiue et al., 1994) and mesomoporous materials (Wei et al., 1999). The family of organic-inorganic hybrid materials has attracted considerable attention because of its interesting properties such as molecular homogeneity, transparency, flexibility and durability. A key issue that remains unresolved in these organic-modified materials is the degree of mixing of the organic-inorganic components, i.e., phase homogeneity. The high optical transparency to visible light indicates that the organic-inorganic phase separation, if any, is on a scale of ≤ 400nm. Such hybrids are promising materials for various applications, e.g.: solid state lasers (optical components), replacements for silicon dioxide as insulating materials in the microelectronic industry, anti-corrosion and scratch resistant coatings, contact lenses or host materials for chemical sensors. In the recent years interest in those materials is connected to their possible applications as biomaterials (Gigant et al., 2002; Joshua et al., 2001; Klukowska et al., 2002; Mackenzie & Bescher, 1998; Matsuura et al., 2001; Spanhel et. al., 1995). One indirect advantage of including polymers is that it is possible to obtain synergistic effects that combine the best properties of polymers with the best properties of inorganic materials. These materials are considered as biphasic materials, where the organic and inorganic phase is mixed at the nm to sub-µm scales. Nevertheless, it is obvious that the properties of these materials are not just the sum of the individual contributions from both phases; the role of the inner interfaces could be predominant. The nature of the interface has recently been used to divide these materials into two distinct classes (Sanchez & Ribot, 1994). In class I, organic and inorganic compounds are embedded and only the weak bonds (hydrogen, van der Waals bonds) give the cohesion to the whole structure. In class II materials, the phases are linked together through strong chemical bonds (covalent or ionic-covalent bonds). Both class I and class II hybrids were prepared by sol-gel technique (Young, 2002).

The aim of the present chapter is to summarize the synthesis via sol-gel and the characterisation methods of amorphous and hybrid materials for biomedical applications. Therefore, the emphasis of our discussion will be focussed on the science, rather than on the technology, of sol-gel processing. The controlled release of pharmaceuticals such as anti-inflammatory agents and antibiotics from strong and biocompatible hosts has relevant applications: they include implantable therapeutic systems, filling materials for bone or teeth repair, which curtail inflammatory or infectious side effects of implant materials when coatings of biocompatible materials containing anti-inflammatory or antibiotic drugs are applied.

2. General processing methods

Different types of colloids can be used to produce polymers or particles from which we can obtain a ceramic material: for example, sols (suspensions of solid particles in a liquid), aerosols (suspensions of particles in a gas) or emulsions (suspensions of liquid droplets in another liquid). The sol-gel chemistry is based on the hydrolysis and polycondensation of molecular precursors such as metal alkoxides M(OR)x, where M = Si, Sn, Ti, Zr, Al, Mo, V, W, Ce and so forth. The following sequence of reactivity is usually found as $Si(OR)_4 << Sn(OR)_4 = Ti(OR)_4 < Zr(OR)_4 = Ce(OR)_4$ (Novak, B.M.,1993). Fig. 1 presents a schema of the procedures which one could follow within the scope of sol-gel processing. In the sol-gel process, the precursors for the preparation of a colloid consist of a metal or metalloid element surrounded by various ligands. A list of the most commonly used alkoxy ligands is presented in Tab. 1.

Alkyl		**Alkoxy**	
methyl	$\bullet CH_3$	Methoxy	$\bullet OCH_3$
ethyl	$\bullet CH_2CH_3$	Ethoxy	$\bullet OCH_2CH_3$
n-propryl	$\bullet CH_2CH_2CH_3$	n-propoxy	$\bullet OCH_2CH_2CH_3$
Iso-propyl	$H_3C(\bullet C)HCH_3$	Iso-propyl	$H_3C(\bullet O)CHCH_3$
n-butyl	$\bullet CH_2(\bullet CH_2)_2CH_3$	n-butoxy	$\bullet O(CH_2)_3CH_3$
Sec-butyl	$H_3C(\bullet C)HCH_2CH_3$	Sec- propoxy	$H_3C(\bullet O)CHCH_2CH_3$
Iso- butyl	$\bullet CH_2CH(CH_3)_2$	Iso- propoxy	$\bullet OCH_2CH(CH_3)_2$
Tert-butyl	$\bullet C(CH_3)_3$	Tert- propoxy	$\bullet OC(CH_3)_3$

Table 1. Commonly used ligands in sol-gel process.

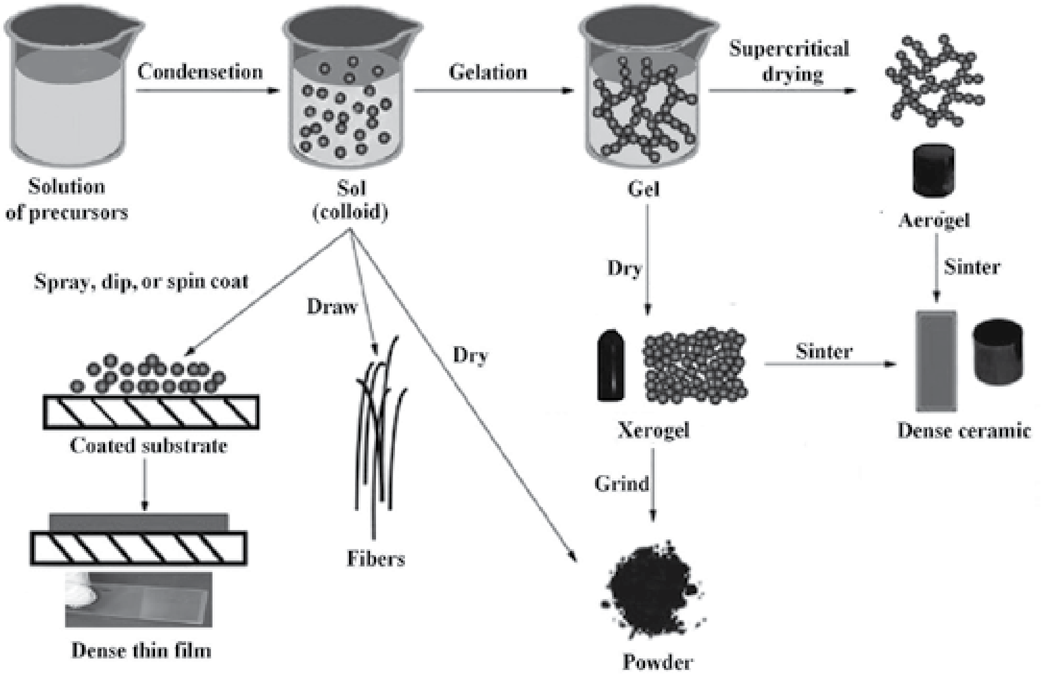

Condensetion
Gelation
Supercritical drying
Solution of precursors
Sol (colloid)
Gel
Aerogel
Spray, dip, or spin coat
Draw
Dry
Dry
Sinter
Xerogel
Sinter
Dense ceramic
Coated substrate
Grind
Fibers
Dense thin film
Powder

Fig. 1. Schematic of sol-gel processing.

2.1 Hydrolysis and condensation

The alkoxydes used as ligands can be organometallic compounds, where direct metal-carbon bonds are present, or also members of the family or metalloid atoms, the so called metal alkoxydes, among which the most widely known, as it has been extensively studied, is the silicon tetraethoxide (or tetrathoxy-silane, or tetraethyl orthosilicate, TEOS), $Si(OC_2H_5)_4$. Silicate gels are most often synthesized by hydrolyzing monomeric, tetrafunctional alkoxide precursors employing a mineral acid (e.g., HCl) or base (e.g. NH_3) as a catalyst. At the functional group level, the sol-gel process starts with the following reaction:

$$Si(OR)_4 + H_2O \xrightarrow[\text{Esterification}]{\text{Hydrolysis}} OH\text{-}Si(OR)_3 + ROH \qquad (1)$$

which can even be stopped while the metal is only partially hydrolyzed, $Si(OR)_{4-n}(OH)_n$. Then, two partially hydrolyzed molecules can link together in a condensation reaction, such as one of the following:

$$(OR)_3Si\text{-}OR + OH\text{-}Si(OR)_3 \xrightarrow[\text{Alcoholysis}]{\text{Alcohol condensation}} (OR)_3Si\text{-}O\text{-}Si(OR)_3 + ROH \qquad (2)$$

$$(OR)_3Si\text{-}OH + OH\text{-}Si(OR)_3 \xrightarrow[\text{Hydrolysis}]{\text{Water condensation}} (OR)_3Si\text{-}O\text{-}Si(OR)_3 + H_2O \qquad (3)$$

where R is an alkyl group, C_xH_{2x+1}. The hydrolysis reaction (eq. 1) replaces alkoxide group (OR) with hydroxyl group (OH). Subsequent condensation reactions involving the silanol group produce siloxane bonds (Si-O-Si) and the by-products alcohol (ROH) (eq. 2) or water (eq. 3). Under most conditions, condensation starts (eqs. 2 and 3) before hydrolysis (eq. 1) is complete. A solvent such as an alcohol is normally used as a homogenizing agent, as water and alkoxysilanes are immiscible (Fig. 2). However, a gel can be prepared from silicon alkoxide-water mixtures without adding a solvent (Avnir & Kaufman, 1987), since the alcohol produced as the by-product of the hydrolysis reaction is sufficient to homogenize the initially phase separated system. It should be noted that the alcohol is not simply a solvent. As indicated by the reverse of eqs. 1 and 2, it can participate in esterification or alcoholysis reactions.

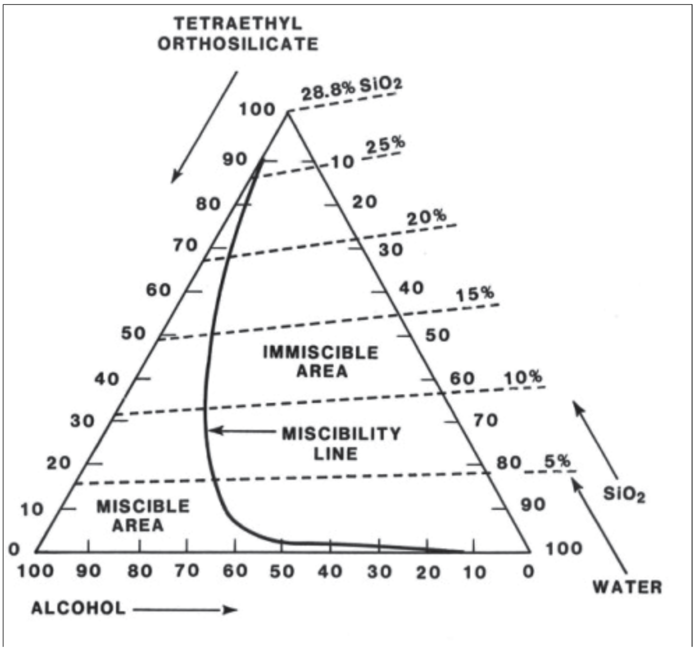


Fig. 2. TEOS , H_2O, Synasol (95% EtOH, 5% water) ternary-phase diagram at 25°C . For pure ethanol the miscibility line is slightly shifted to the right (Cologan & Setterstrom, 1946).

The H_2O:Si molar ratio (r) in eq. 1 has been made to vary from less than one to over 50, and the concentration of acid or bases from less than 0.01 (Brinker et al., 1982) to 7M (Stober et al., 1968) depending on the desired end product. Typical gel-synthesis procedures used to produce bulk gels, films, fibres, and powders are listed in Tab 2. Hydrolysis occurs by the nucleophilic attack of the oxygen contained in water on the silicon atom as shown by the reaction of isotopically labelled water with TEOS that produces only unlabelled alcohol in both acid-base-catalyzed systems (Voronkov & et al., 1978). Hydrolysis is facilitated in the presence of homogenizing agents (alcohols, dioxane, THF, acetone, etc.) that are especially beneficial in promoting the hydrolysis of silanes containing bulk organic or alkoxy ligands. It should be emphasized, however, that the addition of solvents may promote esterification or depolymerization reactions according to the reverse of eqs. 1 and 2.

SiO_2 Gel Types	**% Mole**					
	TEOS	**EtOH**	**H_2O**	**HCl**	**NH_3**	**H_2O/Si(r)**
Bulk	6.7	25.8	67.3	0.2	-	10
Fibers	11.31	77.26	11.31	0.11	-	1.0
Films	5.32	36.23	58.09	0.35	-	10.9
Monodisperse Spheres	0.83	33.9	44.5	-	20.75	53.61

Table 2. Sol-gel Silicate compositions for bulk gels, fibres, film and powder.

The Hydrolysis is more rapid and complete when catalysts are employed (Voronkov et al., 1978). Although mineral acids or ammonia are most generally used in sol-gel processing, other known catalysts are acetic acid, KOH, amines , KF, HF, titanium alkoxides, and vanadium alkoxides and oxides (Voronkov et al , 1978). In the literature mineral acids are reported to be more effective catalysts than the equivalent base concentrations. However, neither the increasing acid of silanol groups with the extent of hydrolysis and condensation (Keefer, 1984) nor the generation of unhydrolyzed monomers via base-catalyzed alcoholic or hydrolytic depolymerization processes have generally been taken into account. Aelion et al., (1950a, 1950b) investigated the hydrolysis of TEOS under acid and basic conditions using several cosolvents: ethanol, methanol, and dioxane. The extent of hydrolysis (eq. 1) was determined by distillation of the ethanol by-product. Karl Fischer titration was used to follow the consumption of water by hydrolysis (eq.1) and its production by condensation (eq.3). Aelion et al. observed that the rate and extent of the hydrolysis reaction was mostly influenced by the strength and concentration of the acid or base catalyst. As under acid conditions, the hydrolysis of TEOS in base media was a function of the catalyst concentration (Aelion et al., 1950a, 1950b).

Steric (spatial) factors exert the greatest effect on the hydrolytic stability of organoxysilanes (Voronkov et al., 1978). Any complication of the alkoxy group delays the hydrolysis of alkoxysilanes, but the hydrolysis rate is lowered at most by the branched alkoxy group (Voronkov et al., 1978). The effects of alkyl length and the degree of branching observed by (Aelion et al., 1950a, 1950b) are illustrated in Tab. 3 for the hydrolysis of tetralkoxysilanes.

R	**k** **10^2 (1 mol^{-1} s^{-1} $[H^+]^{-1}$**
C_2H_5	5.1
C_4H_9	1.9
C_6H_{13}	0.83
$(CH_3)_2CH(CH_2)_3CH(CH_3)CH_2$	0.30

Table 3. Rate constant k for acid hydrolysis of tetralkoxysilanes $(RO)_4Si$ at 20°C

Fig. 3 compares the hydrolysis of TEOS and TMOS under acid and basic conditions. The delaying effect of the bulkier ethoxide group is clearly evident. According to (Voronkov et al., 1978) in the case of mixed alkoxides, $(RO)_x(R'O)_{4-x}Si$ where R'O is a higher (larger) alkoxy group than RO, if the R'O has a normal (i.e. linear) structure, its retarding effect on the hydrolysis rate is manifest only when x= 0 or 1. If R'O is branched, its delaying effect is evident even when x= 2. The hydrolysis of the n-propoxide group was observed to be slower than the ethoxide group during the second hydrolysis step under both acid and basic conditions. This result suggests that a delaying effect of a higher, normal alkoxide group is realized regardless of the extent of substitution.

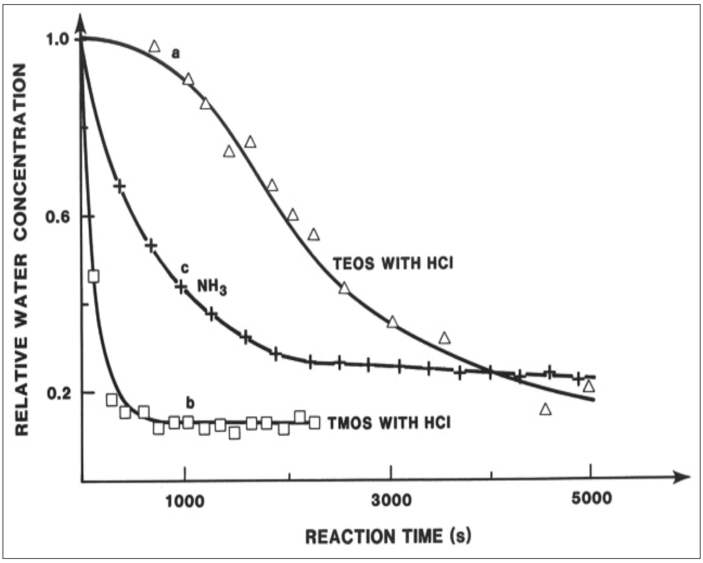


Fig. 3. Relative water concentration versus time during acid- or base-catalzed hydrolysys of □: TMOS with HCl ; X: TEOS and TMOS with NH_3. Δ:TEOS with HCl (Shih et al., 1987)

The substitution of one alkyl group with alkoxy groups increases the electron density on the silicon. Conversely, hydrolysis (substitution of OH for OR) or condensation (substitution of OSi for OR or OH) decreases the electron density on the silicon Fig. 4. Inductive effects are evident from investigations on the hydrolysis of methylethoxysilanes (Schimdt et al., 1984), $(CH_3)_x(C_2H_5O)_{4-x}Si$ where x varies from 0 to 3. Fig. 5 shows that under acidic (HCl) conditions, the hydrolysis rate increases with the degree of substitution x, of electron-providing alkyl group, whereas under basic (NH_3) conditions the reverse trend is clearly observed.

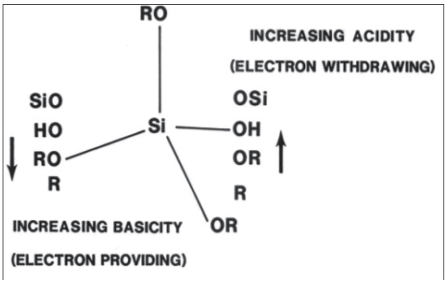


Fig. 4. Inductive effects of substituents attached to silicon,R, OR, OH or OSi (Brinker, 1988)

Fig. 5 also shows the accelerating effect of methoxide substitution on the hydrolysis rate (TMOS versus TEOS). The acceleration and retardation of hydrolysis with increasing x under acid and basic conditions respectively, suggest that the hydrolysis mechanism is sensitive to inductive effects and is apparently unaffected by the extent of alkyl substitution. Because increased stability of the transition state will increase the reaction rate, the inductive effects are evident for positively and negatively charged transition states or intermediates under acid and basic conditions respectively. This reasoning leads to the hypothesis that under acid conditions, the hydrolysis rate decreases with each subsequent hydrolysis step (electron withdrawing), whereas under basic conditions the increased electron-withdrawing capabilities of OH (and OSi) compared to OR may establish a condition in which each subsequent hydrolysis step occurs more quickly as hydrolysis and condensation proceed.

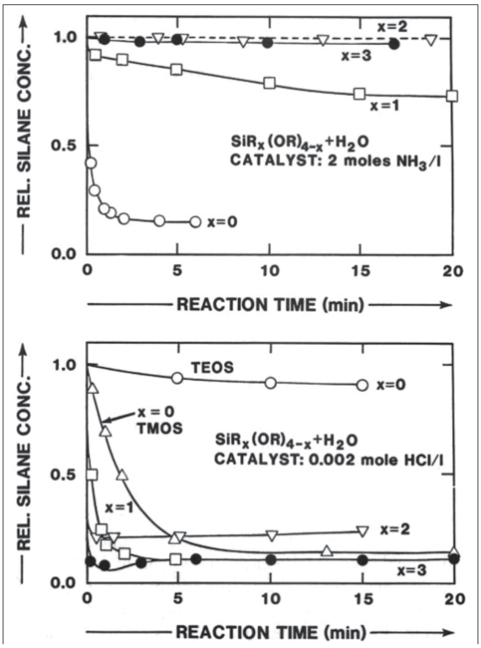
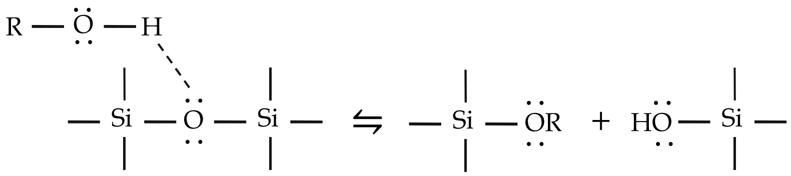


Fig. 5. Relative silane concentration versus time during acid- and base-catalyzed hydrolysis of different silanes in ethanol (volume ratio to EtOH=1:1). ●: $(CH_3)_3SiOC_2H_5$. ∇:$(CH_3)_2Si(OC_2H_5)_2$. □: $(CH_3)_2Si(OC_2H_5)_3$. ○:$Si(OC_2H_5)_4$. Δ:$Si(OCH_3)_4$. (Shih et al., 1987)

From the standpoint of organically modified alkoxysilanes, $R_xSi(OR)_{4-x}$, the inductive effects indicate that acid-catalyzed conditions are preferable (Schimdt et al., 1984), since acids are effective in promoting hydrolysis both when x=0 and x>0. As indicated in Tab. 2, the hydrolysis reaction has been performed with r values ranging from <1 to over 25 depending on the desired polysilicate product, for example, fibers, bulk gel or colloidal particles. From eq. 1, an increased value of r is expected to promote the hydrolysis reaction. (Aelion et al., 1950a, 1950b) found the acid-caltayzed hydrolysis of TEOS to be first-order in $[H_2O]$; however, they observed an apparent zero-order dependence of the water concentration under base-catalyzed conditions. As explained, this is probably due to the production of monomers by siloxane bond hydrolysis and redistribution reactions.

Solvents are usually added to prevent liquid-liquid phase separation during the initial stages of the hydrolysis reaction and to control the concentrations of silicate and water that influence the gelation kinetics. More recently, the effects of solvents have been studied primarily in the context of drying control chemical additives (DCCA) used as cosolvents with alcohol in order to facilitate rapid drying of monolithic gels without cracking (Hench et al., 1986). Solvents can be classified as polar or nonpolar and as protic or aprotic . The dipole moment of a solvent determines the length over which the charge of one species can be "felt" by surrounding species. The lower the dipole moment, the larger this length becomes. This is important in electrostatically stabilized systems and when considering the distance over which a charged catalytic species, for example an OH^- nucleophile or H_3O^+ electrophile, is attracted to or repelled from potential reaction sites, depending on their charge. The availability of labile protons determines whether anions or cations are solvated more strongly through hydrogen bonding. Because hydrolysis is catalyzed either by hydroxyl (pH>7) or hydronium ions (pH<7), solvent molecules that hydrogen bonds to hydroxyl or hydronium ions reduce the catalytic activity under basic or acid conditions respectively. Therefore, aprotic solvents that do not form a hydrogen bond to hydroxyl ions have the effect of making hydroxyl ions more nucleophilic, whereas protic solvents make hydronium ions more electrophilic (Morrison & Boyd, 1966).

Hydrogen bonding may also influence the hydrolysis mechanism, hydrogen bonding with the solvent can sufficiently activate weak leaving group to realize a bimolecolar, nucleophilic (S_N2-Si) reaction mechanism (Voronkov et al., 1978). The availability of labile protons also influences the extent of the reverse reactions, reesterification (reverse eq. 1) or siloxane bond alcoholysis or hydrolysis (reverse of eqs. 2 and 3). Aprotic solvents do not participate in reverse reactions such as reesterification or hydrolysis, because they lack a sufficiently electrophilic proton and are unable to be deprotonated to form sufficiently strong nucleophiles (OH^- or OR^-) necessary for reaction 4. Therefore compared to alcohol or water, aprotic solvents such as THF or dioxane are considerably more "inert" (they do not formally take part in sol-gel processing reactions), they may influence reaction kinetics by increasing the strength of nucleophiles or decreasing the strength of electrophiles.



$$\overset{\mathrm{R}-\ddot{\mathrm{O}}-\mathrm{H}}{-\mathrm{Si}-\ddot{\mathrm{O}}-\mathrm{Si}-} \rightleftharpoons -\mathrm{Si}-\ddot{\mathrm{O}}\mathrm{R} + \mathrm{H}\ddot{\mathrm{O}}-\mathrm{Si}- \qquad (4)$$

2.2 Gelation

Clusters resulting from the hydrolysis and condensation reactions eventually collide and link together into a gel, which is often defined as "strong" or "weak" according to whether the bonds connecting the solid phase are permanent or reversible; however, as noted by (Flory, 1974), the difference between weak and strong ones is a matter of time scale. Even covalent siloxane bonds in silica gel can be cleaved, allowing the gel to exhibit a slow and irreversible (viscous) deformation. Thus the chemical reactions that bring about gelation continue long beyond the gelation point, permitting flow and producing gradual changes in the structure and properties of the gel. The outline of the aging of a gel is as follows:

- Phenomenology
- Classical theory
- Percolation theory
- Kinetic model

The simplest picture of gelation is that clusters grow by condensation of polymers or aggregation of particles until the clusters collide; then, links form between the clusters to produce a single giant cluster that is called gel. When the gel forms, many clusters will be present in the sol phase, entangled in but not attached to the spanning cluster; with time, they progressively become connected to the network and the stiffness of the gel will increase. The gel appears when the last link is formed between two large clusters to create the spanning cluster. This bond is no different from innumerable others that form before and after the gel point, except that it is responsible for the onset of elasticity by creating a continuous solid network. The sudden change in rheological behaviour is generally used to identify the gel point in a crude way.

The classic theory explains the theory developed by Flory (1953) and Stockmayer (1945) to account for the gel point and the molecular-weight distribution in the sol. The most important deficency of this model is that it neglects the formation of closed loops within the growing clusters, and this leads to unrealistic predictions about the geometry of the polymers. The percolation theory offers a description that does not exclude the formation of closed loops and so does not predict a divergent density for large clusters. The disadvantage of the theory is that it generally does not lead to analytical solutions for such properties as the percolation threshold or the size distribution of polymers. However, these features can be determined with great accuracy from computer simulations, and the results are often quite different from the predictions of the classical theory. Excellent reviews of percolation theory and its relation to gelation have been written by Zallen (1983) and Stauffer et al. (1982); and the kinetic models are based on Smoluchowski' s analysis of the growth and aggregation of clusters.

The Smoluchowski equation describes the rate at which the number, n_s, of clusters of size s changes with time t, during an aggregation process :

$$\frac{dn_s}{dt} = \frac{1}{2}\sum_{i\,+\,j\,=\,s} K\,(i,j)n_i n_j - n_s\,\sum_{j=1}^{\infty} K\,(s,j)n_j \tag{5}$$

The coagulation kernel, K(i,j) is the rate coefficient for aggregation of a cluster of size i with another and of size j. The first term in eq. 5 gives the rate of creation of size s by aggregation of two smaller clusters, and the second term gives the rate at which clusters of size s are eliminated by further aggregation. For this equation to apply, the sol must be so diluted that collisions between more than two clusters can be neglected, and the clusters must be free to diffuse so that the collisions occur at random. Further, since K depends only on i and j, ignoring the range of structures that could be present in a cluster of a given size, this is a mean-field analysis that replaces structural details with averages.

2.3 Drying

The drying of a porous material is a process which can be divided into several stages. At first the body shrinks by an amount equal to that volume of the evaporated liquid and the liquid-vapor interface remains at the exterior surface of the body. The second stage begins when the body becomes too stiff to shrink and the liquid recedes into the interior, leaving air-filled pores near the surface. Even as air invades the pores, a continuous liquid film supports flow to the exterior, so evaporation continues to occur from the surface of the body. Eventually, the liquid becomes isolated into pockets and drying can proceed only by evaporation of the liquid within the body and diffusion of the vapor to the outside. In the specialized literature the factors affecting stress development are discussed and various strategies to avoid warping and cracking are described. The outline is as follows:

- Phenomenology
- Drying stress
- Avoiding fracture

The first stage of drying is called the constant rate period (CRP), because the rate of evaporation per unit area of the drying surface is uniform (Fortes & Okos, 1980; Macey, 1942; Moore, 1961). The evaporation rate is close to that of an open dish of liquid, as indicated by the data for the drying of alumina gel (Dwivedi, 1986), shown in Fig. 6. The rate may differ slightly, depending on the texture of the surface. For example, as sand beds dry, the water conforms to the shapes of the particles, so the wet area is larger than the planar one pertaining to the surface of the body, and the rate of evaporation is correspondingly higher (Ceaglske & Hougen, 1937). The distribution of a spreading liquid is illustrated schematically in Fig.6. The chemical potential, μ, of the liquid in the adsorbed film is equal to the one under the concave meniscus, otherwise liquid would flow from one to the other to balance the potential. The chemical potential μ is lower than bulk liquid because of disjoining and capillary forces, therefore the vapour pressure (p_v) decreases according to:

$$\frac{p_v}{p_0} = Exp\ (\Delta\mu/R_gT) \tag{6}$$

where p_0 is the vapour pressure of bulk liquid, R_g is the ideal gas constant, T is the temperature and Δμ is the increment of the chemical potential. The rate of evaporation, V_E, is proportional to the difference between p_v and the ambient vapour pressure, p_A:

$$V_E = k\ (p_v - p_A) \tag{7}$$

where k is a coefficient that depends on the design of the drying chamber, draft rate, etc. It appears reasonable to conclude that the surface of the body must be covered with a film of liquid (as in Fig. 6a), because the rate would decrease as the body shrinks if evaporation occurrs only from the menisci, Fig. 6b.

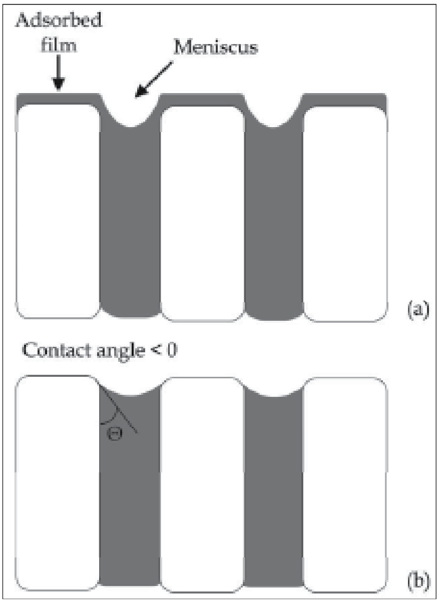


Fig. 6. Distribution of liquid at the surface of a drying porous body, when liquid is (a) spreading (contact angle θ=0°) or (b) wetting, but not spreading (90°> θ >0°). The chemical potential of the liquid in the adsorbed film is equal to that under the meniscus.

3. Experimental procedures

3.1 Sol-gel synthesis of organic – inorganic hybrid materials

Hybrid organic-inorganic biomaterials were prepared by means of a sol-gel process from an analytical reagent grade of metal alkoxides $M(OR)_x$, in an ethanol, organic polymer like poli-ε-caprolactone (PCL Mw = 65,000), water and solvent ($CHCl_3$) mixture. Water, diluted with ethanol was added to the solution under a vigorous stirring. A flow-chart of hybrids (MO_2 + PCL x wt%) can show the synthesis by the sol-gel method. MO_2/PCL, all mixed with drugs (y wt%), were also prepared by using an analytical reagent grade as a precursor material.

In this study SiO_2/PCL (PCL 0, 6, 12, 50 wt%) materials were used as support matrices for controlled drug release. Silica gel, originally developed for engineering applications, is also currently being studied as a polymer for the entrapment and sustained release of drugs (Teoli et al., 2006). In the present study the sol-gel method was applied to encapsulate Ketoprofen (5, 10, 15 wt%) as a model drug. The drug loaded amorphous bioactive materials were studied in terms of their drug release kinetics

The hybrid inorganic-organic materials (PCL 0, 6, 12, 50 wt%) were prepared by means of sol-gel process from an analytical reagent grade of tetraethyl orthosilicate (TEOS) in an ethanol, poly-ε-caprolactone (PCL), water, and chloroform ($CHCl_3$) mixture. Water, diluted with ethanol was added to the solution under vigorous stirring. Fig. 7 shows the flow chart of hybrid (SiO_2 + %PCL + %Ketoprofen) synthesis by the sol-gel method. As it is shown in the same Fig. 7, SiO_2/PCL (PCL 0, 6, 12, 50 wt%) all mixed with ketoprofen (5, 10, 15 %) were prepared by using an analytical reagent grade as precursor material.

After the addition of each reactant the solution was stirred and the resulting sols were uniform and homogeneous. The gelification time was controlled by varying the concentration of PCL, as shown in Tab. 4. After gelification the gels were air dried at 50°C for 24h to remove the residual solvent; as this treatment does not modify the stability of ketoprofen, glassy pieces were obtained (Fig. 8). Discs with a diameter of 13 mm and a thickness of 2 mm were obtained by pressing a fine (<125 μm) gel powder into a cylindrical holder.

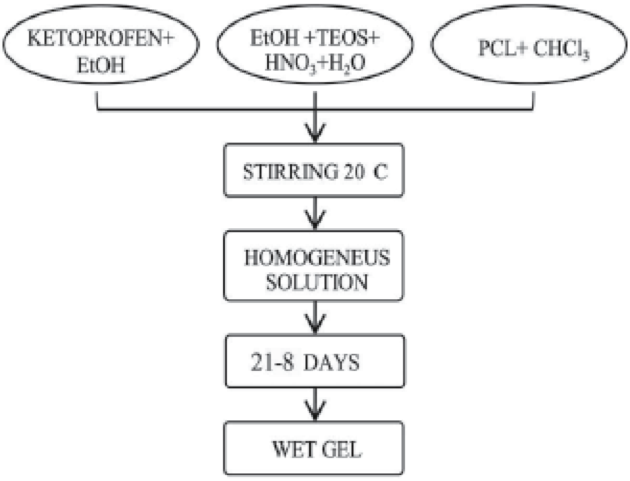


Fig. 7. Flow chart of SiO_2/PCL gel synthesis.

Materials prepared	Gelation time (20-25°C)
SiO_2 + 0%PCL + 5%Ketoprofen	21
SiO_2 + 6%PCL + 5%Ketoprofen	16
SiO_2 + 12%PCL + 5%Ketoprofen	14
SiO_2 + 50%PCL + 5%Ketoprofen	8
SiO_2 + 0%PCL + 10%Ketoprofen	19
SiO_2 + 6%PCL + 10%Ketoprofen	15
SiO_2 + 12%PCL + 10%Ketoprofen	12
SiO_2 + 50%PCL + 10%Ketoprofen	9
SiO_2 + 0%PCL + 15%Ketoprofen	18
SiO_2 + 6%PCL + 15%Ketoprofen	14
SiO_2 + 12%PCL + 15%Ketoprofen	12
SiO_2 + 50%PCL + 15%Ketoprofen	10

Table 4. Variation in the gelification time, controlled by changing the concentration of PCL.

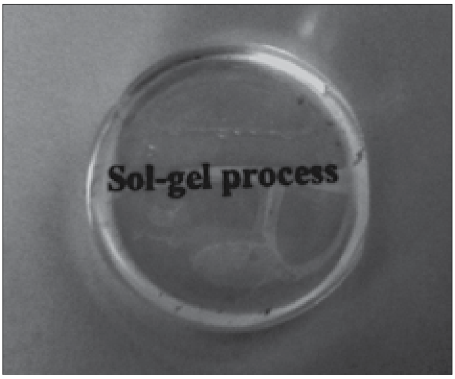


Fig. 8. SiO_2/PCL gel after drying.

Chromatographic experiments were carried out on a Shimadzu HPLC system, equipped with a Class-VP 5.0 software, an UV spectrophotometric detector SPD-10AVvp and two pumps LC-10ADvp, with low-pressure gradient systems. Samples of solutions were injected by a syringe via a Rheodyne loop injector; the loop volume was 20 µl, the analytical column was a Phenomenex C18 (150 × 4.60 mm; 5 µ); the flow rate of the mobile phase A (water) was set at 0.8 ml/min and that of the mobile phase B (methanol) was set at 0.2 ml/min. The total runtime was 10 minutes. HPLC grade methanol was obtained by Sigma-Aldrich. HPLC grade water was prepared using a Millipore (0.22 µm) system. A standard solution of ketoprofen 3 mM in a simulated body fluid (SBF) was prepared and the samples were taken at the end of the release from the materials.

The nature of SiO_2 gel, poly-ε-caprolactone (PCL) and PCL/SiO_2 hybrid materials were ascertained by X-ray diffraction (XRD) analysis using a Philips diffractometer. The presence of hydrogen bonds between organic-inorganic components of the hybrid materials was ascertained by FTIR analysis. Fourier transform infrared (FTIR) transmittance spectra were recorded in the 400-4000 cm^{-1} region using a Prestige 21 Shimatdzu system, equipped with a DTGS KBr (Deuterated Tryglycine Sulphate with potassium bromide windows) detector, with resolution of 2 cm^{-1} (45 scans). KBr pelletized disks containing 2 mg of sample and 200 mg KBr were made. The FTIR spectra were elaborated by IR solution software. The microstructure of the synthesized gels was studied by a scanning electron microscopy (SEM) Cambridge model S-240 on samples previously coated with a thin Au film and by a Digital Instruments Multimode atomic force microscopy (AFM) in contact mode in air.

3.2 Study of in vitro bioactivity

In order to study their bioactivity, samples of the studied hybrid materials were soaked in a simulated body fluid (SBF) with ion concentrations nearly equal to those of the human blood plasma, as reported elsewhere and shown in Tab. 5 (Hench & Clark, 1978; Ohtsuki et al., 1992; Paul, 1992). During soaking, the temperature was kept fixed at 37°C. Taking into account that the ratio of the exposed surface to the volume solution influences the reaction, a constant ratio of $50\ mm^2\ ml^{-1}$ of solution was respected (Hutmacher et al., 2001).

	Ions concentration (mM)	
Ion	**Human blood plasma**	**SBF**
Na^+	142.0	142.0
K^+	5.0	5.0
Mg^{2+}	1.5	1.5
Ca^{2+}	2.5	2.5
Cl^-	103.0	148.0
HCO_3^-	27.0	4.2
HPO_4^{2-}	1.0	1.0
SO_4^{2-}	0.5	0.5
pH	7.2 - 7.4	7.4

Table 5. Simulated body fluid (SBF) ionic concentration (mM).

It was shown that SBF reproduces *in vivo* bonelike apatite formation on bioactive glass and ceramic (Hench, 1991). *In vitro* studies using SBF have suggested that bioactive glass and glass-ceramics form bonelike apatite on their surface (see Fig. 9) by providing surface functional groups of silanol (Si—OH), which are effective for apatite nucleation. These groups combine with Ca^{2+} ions present in the fluid imposing an increase of positive charge on the surface. In addition, Ca^{2+} ions combine with the negative charge of the phosphate ions to form amorphous phosphate, which spontaneously transforms into hydroxyl-apatite $[Ca_{10}(PO_4)_6(OH)_2]$ where the atomic ratio Ca/P is 1.60 (Ohtsuki et al., 1992). The SBF is already supersaturated with respect to the apatite under normal conditions. Once the apatite nuclei are formed, they can grow spontaneously by consuming the calcium and phosphate ions from the surrounding body fluid. It is known from literature (Hench, 1991; Ohtsuki et al., 1992) that CaO, SiO_2-based glasses, CaO, P_2O_5- based glasses, sodium silicate glasses are more bioactive then ion free glass and ceramic. That is due to the dissolution of appreciable amounts of calcium and phosphate ions, which increase the degree of supersaturation of the surrounding body fluid with respect to the apatite. Moreover, the released ions are exchanged with H_3O^+ ions in SBF forming silanol groups on their surface; this reaction causes a pH increase of SBF solution and, consequently, Si–OH groups are dissociated into negatively charged units Si–O^- that interacts with the positively charged calcium ions to form the calcium silicate.

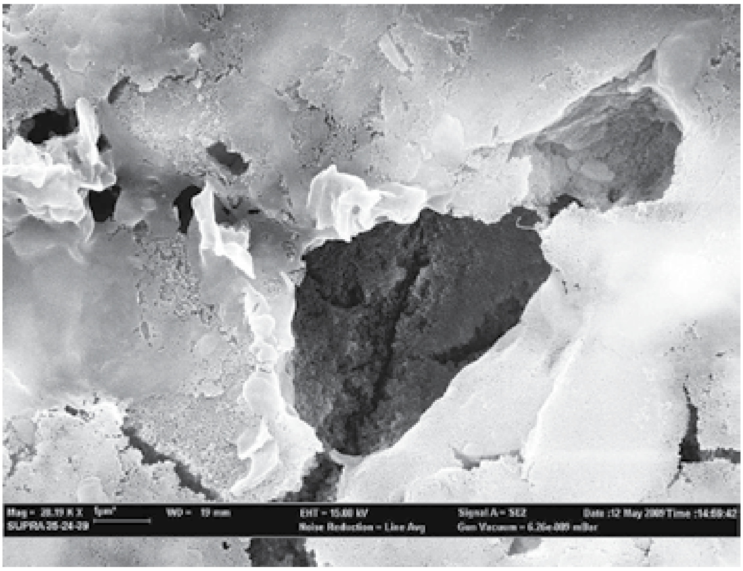


Fig. 9. Apatite layer on SiO_2/PCL surface.

3.3 Study of in-vitro release

For the study of drug release, the discs of the investigated material were soaked in 15 ml of SBF, continuously stirred, at 37°C. The SBF was previously filtered with a Millipore (0.22 μm) system, to avoid any bacterial contamination. Drug release measurements were carried out by means of UV-VIS spectroscopy. Absorbance values were taken at a wavelength λ corresponding to an absorbance maximum value. The calibration curve was determined by taking absorbance versus drug concentration between 0 and 3 mM as parameters. For that interval the calibration curve fits the Lambert and Beers' law (Wang & Pantano, 1992):

$$A = 1{,}26 \cdot C \tag{8}$$

where A is the absorbance and C is the concentration (mM).

4. Characterization

4.1 Sol-gel characterization

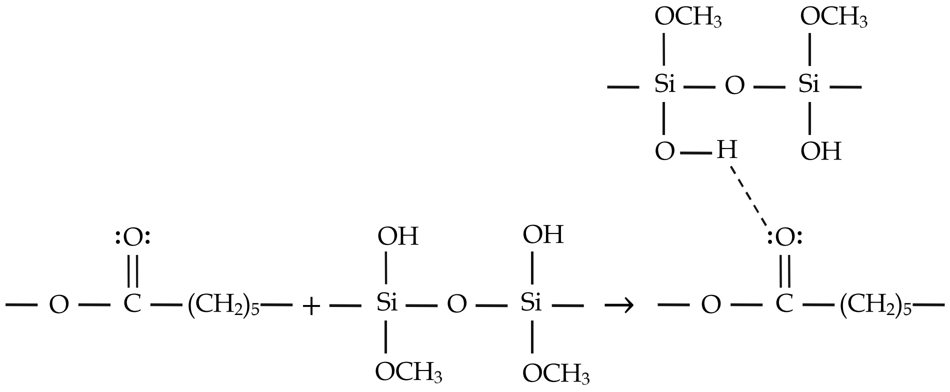
Gelification is the result of hydrolysis and condensation reactions according to the following reactions:

$$Si(OCH_3)_4 + nH_2O \Rightarrow Si(OCH_3)_{4-n}(OH)_n + nCH_3OH \tag{9}$$

$$-SiOH + CH_3O\text{-}Si- \Rightarrow -Si\text{-}O\text{-}Si- + CH_3OH \tag{10}$$

$$-Si\text{-}OH + OH\text{-}Si \Rightarrow -Si\text{-}O\text{-}Si- + H_2O \tag{11}$$

Reaction 12 shows the formation of hydrogen bonds between the carbossilic group of organic polymer and the hydroxyl group of inorganic matrix.



(12)

The existence of hydrogen bonds was proved by FTIR measurements. Fig. 10 shows the infrared spectrum of the SiO_2 gel (Fig. 10a), the SiO_2/PCL (6, 12, 50%wt) gels (Fig. 10b, 10c, 10d) and PCL (Fig. 10e). In Fig. 10a the bands between 3400 and 1600 cm^{-1} are attributed to water (Sanchez & Ribot, 1994). The bands at 1080 and 470 cm^{-1} are due to the stretching and bending modes of SiO_4 tetrahedra (Sanchez & Ribot, 1994). In the Fig. 10b, 10c, 10d and 10e, the bands at 2928 and 2840 cm^{-1} are attributed to a symmetric stretching of $-CH_2-$ of policaprolattone. The band at 1730 cm^{-1} is due to the characteristic carboxylic group shifted to low wave numbers. The broad band at 3200 cm^{-1} is the characteristic O-H group of hydrogen bonds.

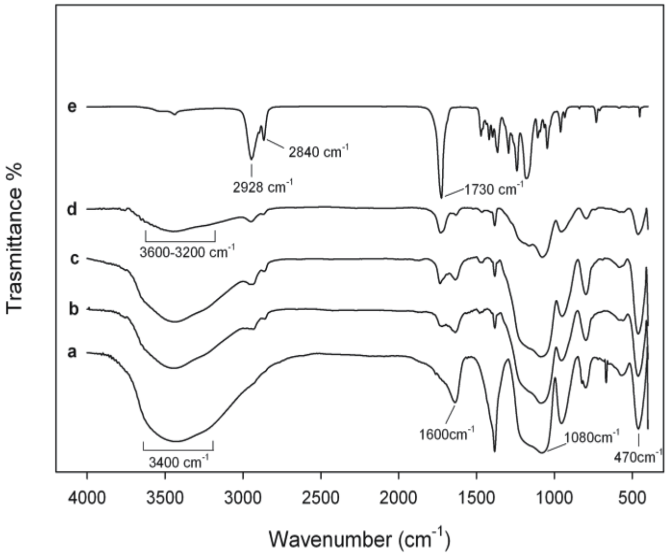


Fig. 10. FTIR of (a) SiO_2 gel, (b) SiO_2+ PCL 6wt% (c) SiO_2+ PCL 12wt% (d) SiO_2+ PCL 50wt% gels and (e) PCL.

The nature and the microstructure of the hybrid materials have been studied by X-ray diffraction (XRD), scanning electron microscopy (SEM) and atomic force microscopy (AFM). The diffractograms in Fig. 11a show that SiO_2 gel exhibits the broad humps which are characteristic for amorphous materials, while the sharp peaks that can be detected on the diffractogram of poly-ε-caprolactone and ketoprofen are typical of crystalline materials (Fig. 11b and 11c). On the other hand the XRD spectrum of hybrid SiO_2/PCL exhibits the broad humps characteristic of amorphous materials (Fig. 11d), as well that of SiO_2 gel.

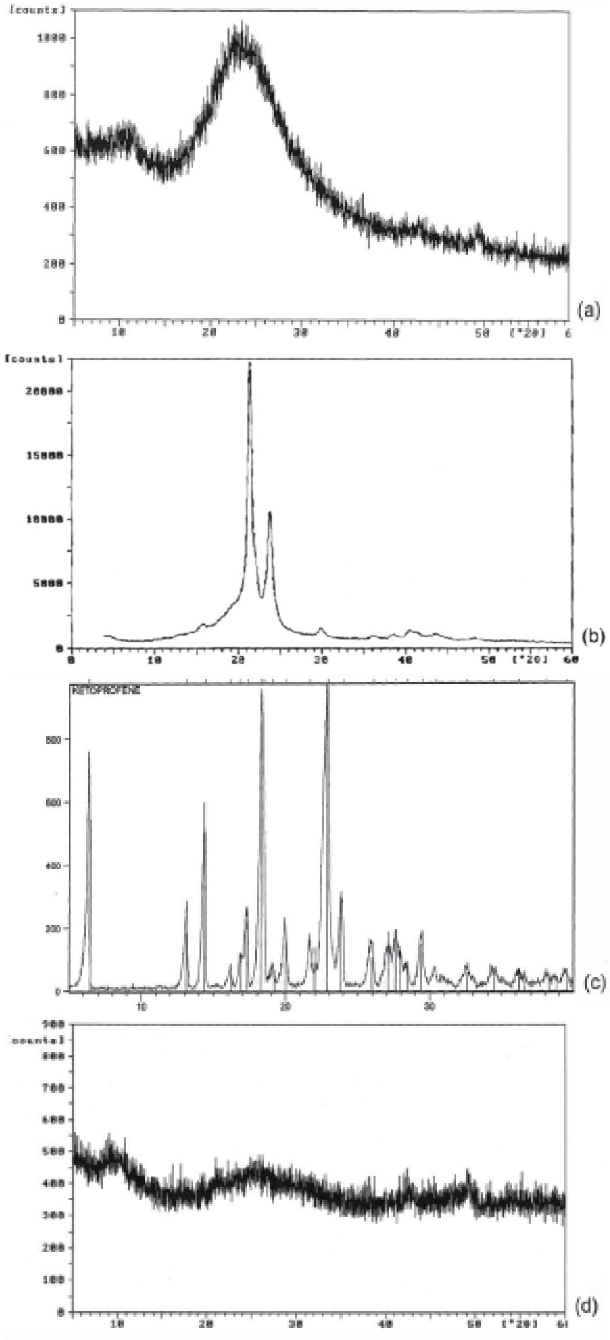


Fig. 11. XRD diffractogram of (a) SiO_2 gel, (b) PCL, (c) Ketoprofen, (d) SiO_2/PCL gel.

SEM micrographs show that no appreciable difference can be observed between the morphology of the four amorphous materials. The samples appear as shown in Fig. 12.
The degree of mixing of the organic-inorganic components, i. e. the phase homogeneity, has been ascertained by applying the atomic force microscopy (AFM) in the analysis of the sol-gel hybrid material.

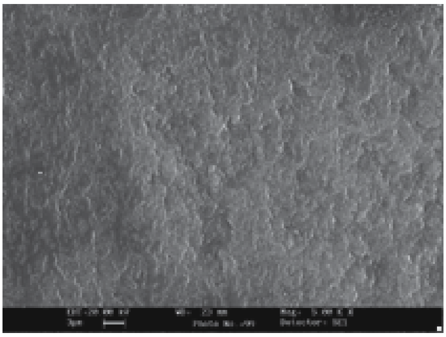
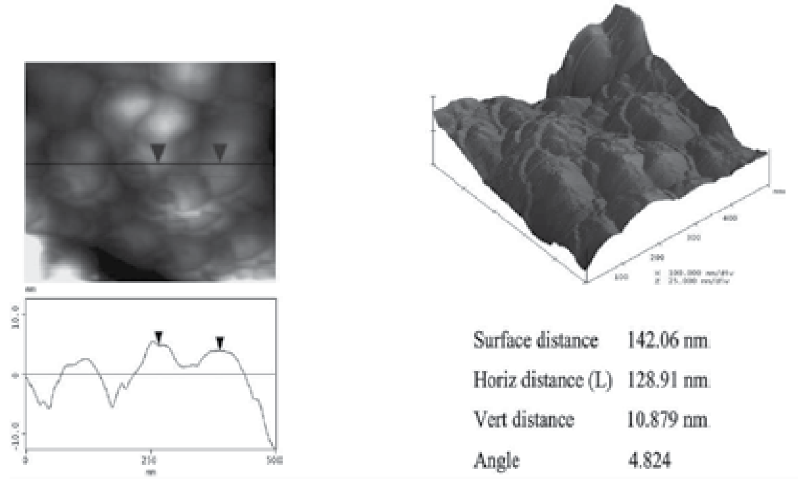


Fig. 12. SEM micrograph of SiO_2/PCL gel.

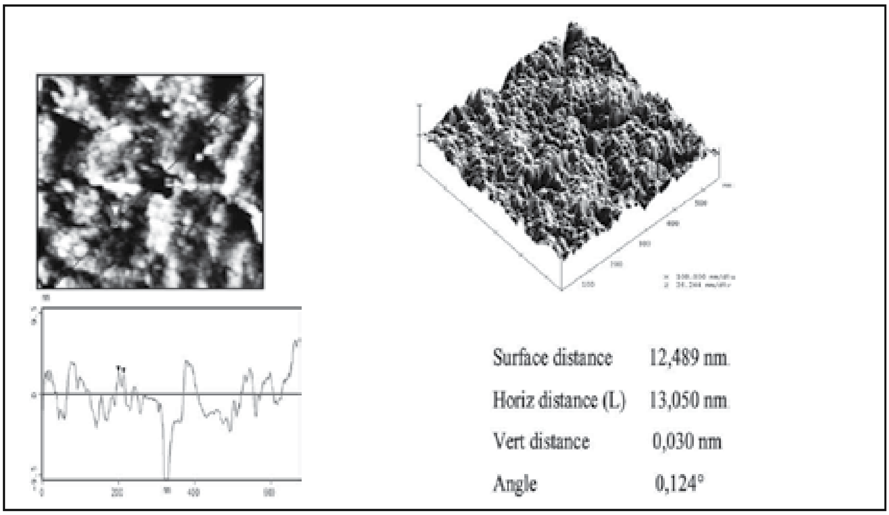
The AFM contact mode image can be measured in the height mode or in the force mode. Force images (z range in nN) have the advantage that they appear sharper and richer in contrast and that the contours of the nanostructure elements are clearer. In contrast, height images (z range in nm) show a more exact reproduction of the height itself. In this work the height mode has been adopted to evaluate the homogeneity degree of the hybrid materials.



	SiO_2	SiO_2 + 6%PCL	SiO_2 + 12%PCL	SiO_2 + 50%PCL
Surface distance	142,06 nm	18,813 nm	15,446 nm	12,489 nm
Horiz distance (L)	128,91 nm	20,507 nm	16,779 nm	13,050 nm
Vert distance	10,079 nm	0,0529 nm	0,0697 nm	0,030 nm
Angle	4,824°	0,143°	0,221°	0,124°

Fig. 13. AFM image showing the microstructure of SiO_2 gel .

The AFM topographic images of SiO_2 and SiO_2/PCL (0, 6, 12, 50 wt%) gel samples are shown in Figs. 13 and 14. It can be observed that the average domain size is less than 130 nm. This result confirms that the synthesized PCL/SiO_2 gels can be considered organic-inorganic hybrid materials as suggested by literature data (Hench & Clark, 1978).



Surface distance	12,489 nm
Horiz distance (L)	13,050 nm
Vert distance	0,030 nm
Angle	0,124°

Fig. 14. AFM image showing the microstructure of SiO_2/PCL gel.

4.2 Biological characterisation

The hybrid materials were soaked in SBF, as indicated by Ohtsuki et al. (1992), for in vitro bioactivity tests. The FTIR spectra after several exposures to SBF, 7, 14 and 21 days are shown in Fig. 15b, 15c and 15d. Evidence of the formation of an hydroxyapatite layer is given by the appearance of the 1116 and 1035 cm^{-1} bands, usually assigned to P-O stretching (Teoli et al., 2006) and of the 580 cm^{-1} band usually assigned to the P-O bending mode (Teoli et al., 2006). The splitting, already after a 7 day soaking, of the 580 cm^{-1} band into two others at 610 and 570 cm^{-1} can be attributed to formation of crystalline hydroxyapatite (Ohtsuki et al.). Finally the band at 800 cm^{-1} can be assigned to the Si-O-Si band vibration between two adjacent tetrahedral, characteristic of silica gel (Teoli et al., 2006). These considerations support the hypothesis that a surface layer of silica gel forms as supposed in the mechanism proposed in the literature for hydroxyapatite deposition (Allen et al., 2000; Khor et al., 2002). Moreover an evaluation of the morphology of the apatite deposition and a qualitative elemental analysis were also carried out by electron microscopy observations on pelletized discs previously coated with a thin Au film. The EDS reported in Tab. 5 confirm that the surface layer observed in the SEM micrographs (Fig. 16) consists of calcium phosphate and which increases as the PCL.

Materials soaked in SBF for 21 days	**Contents of Ca Atomic %**	**Contents of P Atomic %**
SiO_2 + 0%PCL + Ketoprofen	2.65	1.64
SiO_2 + 6%PCL + Ketoprofen	3.13	1.94
SiO_2 + 12%PCL + Ketoprofen	5.04	3.23
SiO_2 + 50%PCL + Ketoprofen	7.82	4.86

Table 5. The EDS analyses of hybrid materials 21 days after immersion in SBF.

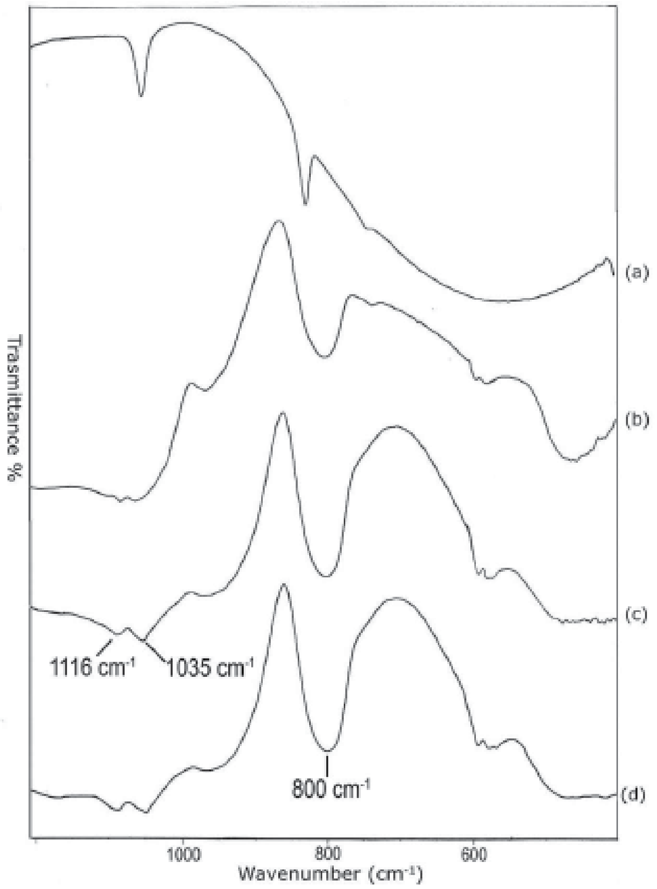


Fig. 15. FTIR spectra of SiO_2/PCL gel samples after different times of exposure to SBF: (a) not exposed; (b) 7 days exposed;(c) 14 days exposed; (d) 21 days exposed.

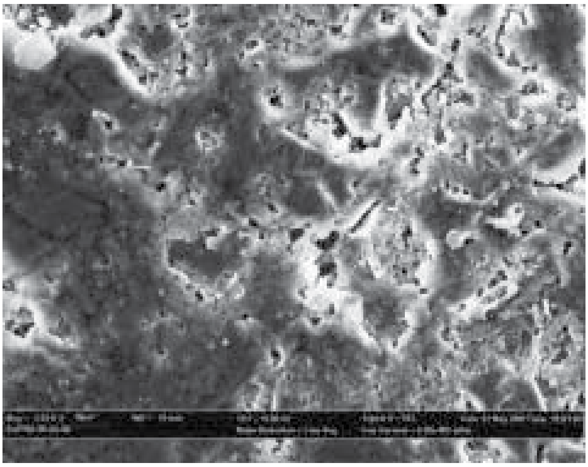


Fig. 16. SEM micrograph of SiO_2/PCL gel after being exposed to SBF 21 days.

4.3 Release kinetic characterization

Kinetic measurements of release from the studied materials were carried out in 15 ml of SBF incubated at 37±0.1°C and under continuous magnetic stirring at 150 rpm. Sink conditions were maintained throughout all studies. The discs used were obtained with particle size between 63-125μm compressed at 3 tons and aliquots of 600 μl were withdrawn at 1 h interval and replaced with an equal volume of release medium pre-equilibrated to temperature. Release was essayed by measuring the photometrical absorbance at 259.5 nm. In order to establish the relationship between the UV absorbance of at λ = 259.5 nm and the concentration of the solutions a calibration curve (r^2 =0.9907) was drawn for a standard solution with 4 levels of concentration: 0.0 mM, 1.0 mM, 2.0 mM and 3.0 mM (Fig. 17). All the standard solutions were prepared in SBF.

Fig. 18a, 19a, 20a and 18b, 19b, 20b show the drug release rates expressed as a percentage of the drug delivered, related to the drug-loading value, as a function of time. It was observed that from the SiO_2+PCL (0, 6, 12, 50 wt%)+ ketoprofen 5wt% gels about 60wt% of the drug was released in a relatively fast manner during the initial 2 hrs and it seems to be completed within 7 hrs without any evident difference in the time of release. For the SiO_2+PCL(0, 6, 12, 50 wt%)+ ketoprofen 10 and 15wt % gels about 60wt% of the drug was released during the initial 1 hr and 0,5 hr respectively and it is complete in about 3 hr and 4 hr respectively.

The differences observed in the release behaviour between SiO_2+ PCL (0, 6, 12, and 50 wt%) + ketoprofen might be due to the different networks of the four gels that are determined by the different content percentage of PCL. The two stage release observed in all cases suggests that the initial stage of release occurs mainly by dissolution and diffusion of the drug entrapped close to or at the surface of the samples. The second and slower release stage involves the diffusion of the drug entrapped within the inner part of the clusters. An interesting observation is the general presence of an early lag period, which indicates the need for the penetration of the solvent into the structure. Fig. 18b, 19b and 20b show this particular kinetic describing the changes of the release speed during the two stages.

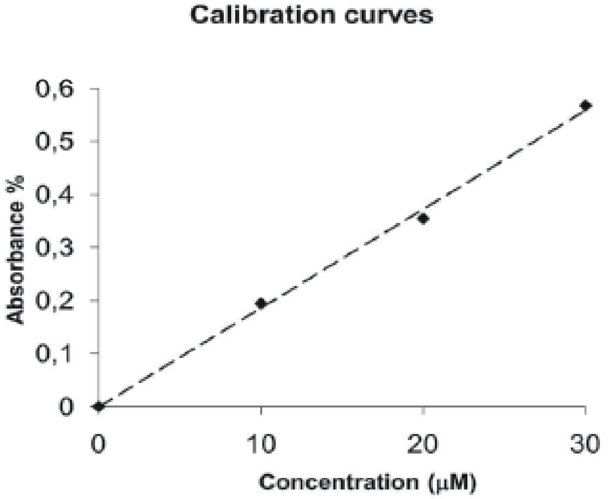


Fig. 17. Calibration curve (259.5 nm) depending on the concentration of Ketoprofen.

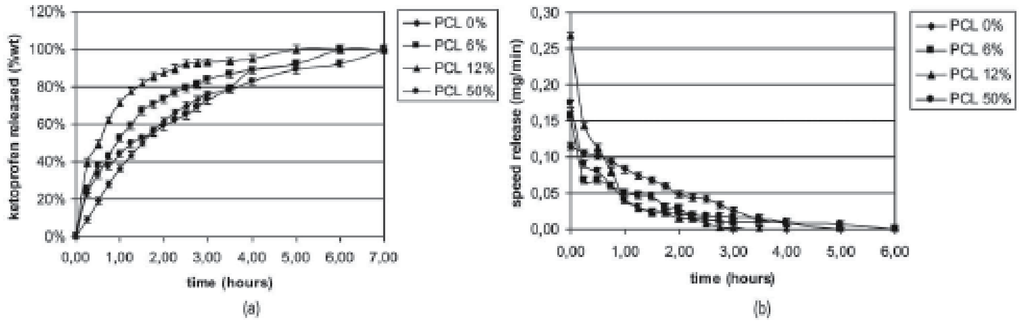


Fig. 18. (a) Time-dependent drug release plot for SiO_2 + PCL (0, 6, 12, 50%wt) + ketoprofen 5% at 37°C in SBF solution; (b) Time-dependent drug release rate plot for SiO_2 + PCL (0, 6, 12, 50%wt) + ketoprofen 5% at 37°C in SBF solution.

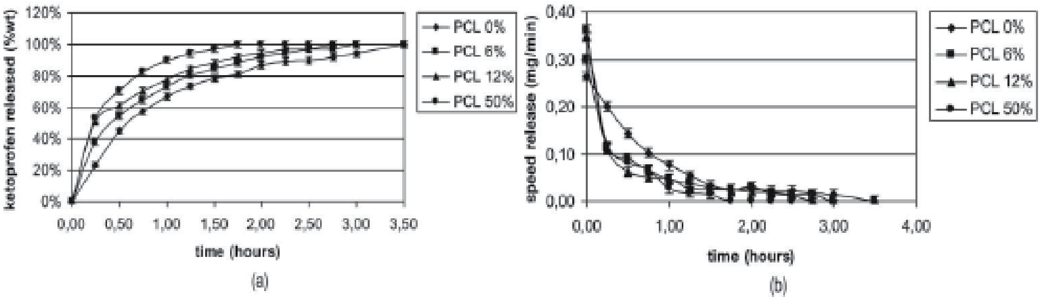


Fig. 19. (a) Time-dependent drug release plot for SiO_2 + PCL (0, 6, 12, 50%wt) + ketoprofen 10% at 37°C in SBF solution; (b) Time-dependent drug release rate plot for SiO_2 + PCL (0, 6, 12, 50%wt) + ketoprofen 10% at 37°C in SBF solution.

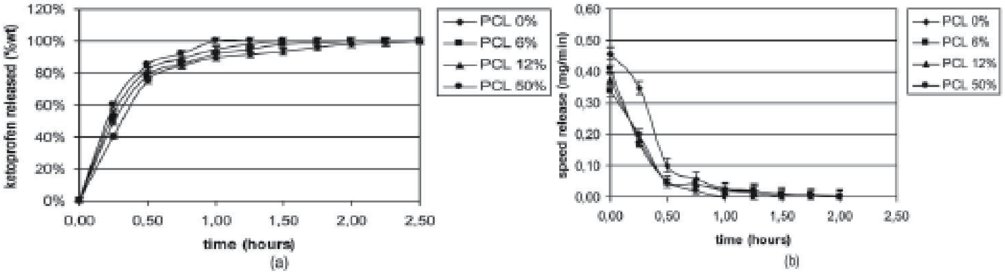


Fig. 20. (a) Time-dependent drug release plot for SiO_2 + PCL (0, 6, 12, 50%wt) + ketoprofen 15% at 37°C in SBF solution; (b) Time-dependent drug release rate plot for SiO_2 + PCL (0, 6, 12, 50%wt) + ketoprofen 15% at 37°C in SBF solution.

5. Applications

Applications for sol-gel processing derive from the various special shapes obtained directly from the gel state (e.g. monoliths, films, fibers, and monisized powders) combined with composition and microstructural control and low processing temperatures. Compared to conventional sources of ceramic raw materials, often minerals dug from the earth, synthetic chemical precursors are a uniform and reproducible source of raw materials than can be made extremely pure through various synthetic means. Low processing temperatures, which result from microstructural control (e.g. high surface areas and small pore sizes), expand glass-forming regions by avoiding crystallization or phase separation, making new materials available to the technologist. The advantages of the sol-gel process (for preparing glass) are shown in Tab. 6 (Brinker & Scherer 1990). The disadvantages of sol-gel processing include the cost of the raw materials, shrinkage that accompanies drying and sintering, and processing time, as it is shown in Tab 7.

1. Better homogeneity from raw materials.
2. Better purity from raw materials.
3. Low temperature of preparation:
    a. Saving energy;
    b. Minimizing evaporation losses;
    c. Minimizing air pollution ;
    d. No reactions with containers, thus purity;
    e. Bypassing crystallization.
4. New noncrystalline solids outside the range of normal glass formation.
5. New crystalline phases from new noncrystalline solids.
6. Better glass products from special properties of gel.
7. Special products such as film.

Table 6. Some advantages of the Sol-Gel Methods over Conventional Melting for Glass

1. High cost of raw materials.
2. Large shrinkage during processing.
3. Residual fine pores.
4. Residual hydroxyl.
5. Residual carbon
6. Health hazards of organic solution.
7. Long process time

Table 7. Some disadvantages of the Sol-Gel Methods

There are books (Klein, 1988) and a number of review papers (Dislich, 1986; Johnson, 1985; Klein & Garvey, 1982; Mackenzie, 1988; Uhlmann et al., 1984; Ulrich, 1988a) which discuss this topic in detail and whose primary purpose is to provide a source of references to current technology, and at the same time to analyze critical issues associated with the various classes of applications that must be addressed in order to advance the sol-gel technology; for example, a short outline can be as follows:

1. Thin films and coatings, which can be applied to optical, electronic, protective, and porous thin films or coatings. That represents the earliest commercial application of sol-gel technology.
2. Monoliths, i.e. applications for cast bulk shapes dried without cracking in such areas as optical components, transparent superinsulation, and ultralow-expansion glasses.
3. Powders, which can be used as ceramic precursors or abrasive grains and applications of dense or hollow ceramic or glass spheres.
4. Fibers, which are drawn directly from viscous sols and are used primarily for reinforcement or fabrication of refractory textiles.

5. Gels can also be used as matrices for fiber-, whisker-, or particle-reinforced composites and as host for organic, ceramic, or metallic phases.
6. Porous Gel; many applications exist which result from the ability to tailor the porosity of thin free-standing membranes, as well as bulk xerogels or aerogels.

In recent years interest in bioactive and biocompatibility of surface-active, biomaterials has grown.

Biomaterials have been used extensively in medical, personal care and food applications, with many similar polymers being used across disciplines. This perspective will emphasize hybrid materials used in medicine and specifically those designed as scaffolds for use in tissue engineering and regenerative medicine. The areas of active research in tissue engineering include: biomaterials design (incorporation of the appropriate chemical, physical, and mechanical/structural properties to guide cell and tissue organization); cell/scaffold integration (inclusion into the biomaterial scaffold of either cells for transplantation or biomolecules to attract cells, including stem cells, from the host to promote integration with the tissue after implantation); and biomolecule delivery (inclusion of growth factors and/or small molecules or peptides that promote cell survival and tissue regeneration). While a significant and growing area of regenerative medicine involves the stimulation of endogenous stem cells, this perspective will emphasize hybrid materials scaffolds used for delivery of cells and biomolecules. The challenges and solutions pursued in designing polymeric biomaterial scaffolds with the appropriate 3-dimensional structure are curently studied.

Ceramics and hybrid dioxide-based materials for the repairing of muscle–skeletal tissues are being increasingly applied over the last half century (Hench, 1991; Li, et al., 1996). Orthopaedic and maxillo–facial prosthesis provide evidence for the enhanced biomechanical performance of titanium and its alloys among metallic prosthetic components (Kitsugi et al., 1996). TiO_2-based bioactive ceramic suggests that bone grafting is achieved by supporting the precipitation of calcium (Ca) and phosphorus (P) into a structure similar to the mineral phase of bone. Accordingly, titanium is very promising to develop biomedical materials and devices designed as hard tissue substitutes with improved interface properties (Coreno & Coreno, et al. 2005; Hench, 1991; Li, et al., 1996; Kitsugi et al., 1996).

6. Conclusions

Sol-gel processing has attracted much attention, for the possibility that the method offers to new materials. We define sol-gel rather broadly as the preparation of glass, glass-ceramic and hybrid materials by a sol, its gelation and removal of the solvent. The sol-gel chemistry is based on the hydrolysis and polycondensation of molecular precursors such as metal alkoxides $M(OR)x$, where M = Si, Sn, Ti, Zr, Al, Mo, V, W, Ce and so forth.

There are many potential applications of sol-gel derived materials in the form of films, fibers, monoliths, powders, composites, and porous media. The most successful applications are those that utilize the potential advantages of sol-gel processing such as purity, homogeneity, and controlled porosity combined with the ability to form shaped objects at low temperatures, avoiding inherent disadvantages such as costs of raw materials, slow processing times, and high shrinkage.

The SiO_2 + PCL (0, 6, 12 and 50 %wt) materials, prepared via sol-gel process, were found to be organic - inorganic hybrid materials. The polymer (PCL) can be incorporated into the network by hydrogen bonds between the carboxylic groups of organic polymer and the hydroxyl groups of inorganic matrix. The release kinetics demonstrates that the investigated materials supply high doses of the anti-inflammatory during the first hours when soaked in SBF and then a slower drug release allows a maintenance dose until the end of the experiment.

7. References

Aelion, R., Loebel, A. & Eirich, F. (1950). Hydrolysis and polycondensation of tetraalkoxysilanes, *Recueil des Travaux Chimiques des Pays-Bas et de la Belgique*, Vol. 69, pp. 61-75, ISSN 0370-7539

Aelion, R., Loebel, A. & Eirich, F. (1950). Hydrolysis of ethyl silicate, *Journal of the American Chemical Society*, Vol. 72, pp. 5705-5712, ISSN 0002-7863

Allen, C., Han, J., Yu, Y., Maysinger, D. & Eisenberg, A. (2000). Polycaprolactone-*b*-poly(ethylene oxide) copolymer micelles as a delivery vehicle for dihydrotestosterone. *Journal of Controlled Release,* Vol. 63, No 3, pp. 275-286, ISSN 0168-3659

Arcos, D., Ragel, C.V. & Vallet-Regí, M. (2001). Bioactivity in glass/PMMA composites used as drug delivery system, *Biomaterials*, Vol. 22, No. 7, pp. 701-708, ISSN 0142-9612

Avnir, D. & Kaufman, V.R. (1987). Alcohol is an unnecessary additive in the silicon alkoxide sol-gel process, *Journal of Non-Crystalline Solids*, Vol. 92, No. 1, pp. 180-182, ISSN 0022-3093

Black, J. & Hastings, G. (1998). *Handbook of Biomaterial Properties*, Chapman & Hall, ISBN 0412603306, New York, U.S.A.

Brinker, C.J., Keefer, K.D., Schaefer, D.W. & Ashley, C.S. (1982). Sol-gel transition in simple silicates, *Journal of Non-Crystalline Solids,* Vol. 48, No. 1, pp. 47-64, ISSN 0022-3093

Brinker, C.J. (1988). Hydrolysis and condensation of silicates: effects on structure, *Journal of Non-Crystalline Solids*, Vol. 100, No. 1-3, pp. 31-50, ISSN 0022-3093

Ceaglske, N.H. & Hougen, O.A. (1937). Drying granular solids, *Journal of Industrial and Engineering Chemistry (Washington, D. C.)*, Vol. 29, pp. 805-813, ISSN 0095-9014

Cogan, H.D. & Setterstrom, C.A. (1946). Properties of ethyl silicate, *Chemical & Engineering News*, Vol. 24, pp. 2499-2501, ISSN 0009-2347

Coreno, J. & Coreno, O. (2005). Evaluation of calcium titanate as apatite growth promoter. *Journal of Biomedical Materials Research*, Vol. 75A, No. 2, pp. 478–84, ISSN 478-484

Dislich, H. (1986). Sol-gel: science, processes and products, *Journal of Non-Crystalline Solids*, Vol. 80, No 1-3, pp.115-121, ISSN 0022-3093

Dwivedi, R.K. (1986). Drying behavior of alumina gels, *Journal of Materials Science Letters*, Vol. 5, No. 4, pp. 373-376, ISSN 0261-8028

Flory P.J. (1953). *Principles of Polymer Chemistry*, Cornell University Press, ISBN 0-8014-0134-8, Ithaca, New York

Flory, P.J. (1974). Gels and gelling process, *Faraday Discussions of the Chemical Society*, Vol. 57, pp. 7-18, ISSN 0301-7249

Fortes, M. & Okos, M. (1980). Drying Theories: Their Bases and Limitations as Applied to Foods and Grains, In: *Advances in Drying Vol 1,* A. Mujumdar (Ed.), 119-154, Hemisphere Publishing, ISBN 0070439753, New York, U.S.A.

Gigant, K., Posset, U. & Schottner, G. (2002). FT-Raman spectroscopic study of the structural evolution in binary UV-curable vinyltriethoxysilane/tetraethoxysilane mixtures from the sol to the xerogel state, *Applied Spectroscopy*, Vol. 56, No. 6, pp. 762-769, ISSN 0003-7028

Hench, L.L. & Clark, D.E. (1978). Physical chemistry of glass surfaces, *Journal of Non-Crystalline Solids*, Vol. 28, No. 1, pp. 83-105, ISSN 0022-3093

Hench, L.L., Orcel, G. & Nogues, J.L. (1986). The role of chemical additives in sol-gel processing, *Materials Research Society Symposium Proceedings*, Vol. 73, No. Better Ceramic Chemistry 2, pp. 35-47, ISSN 0272-9172

Hench, L.L. & West, J.K. (1990). The sol-gel process, *Chemical Reviews*, Vol. 90, No. 1, pp. 33-72, ISSN 0009-2665

Hench, LL. (1991). Bioceramics: from concept to clinic, *Journal of American Ceramic Society*, Vol. 74, No. 7, pp. 1487–1510, ISSN 1551-2916.

Hsiue, G.H., Kuo, J.K., Jeng, R.J., Chen, J.I., Jiang, X.L., Marturunkakul, S., Kumar, J. & Tripathy, S.K. (1994). Stable second-order nonlinear optical polymer network based on an organosoluble polyimide, *Chemistry of Materials*, Vol. 6, No. 7, pp. 884-887, ISSN 0897-4756

Hutmacher, D.W., Schantz, T., Zein, I., Woei N.K., Hin T.S. & Cheng T.K. (2001). Mechanical properties and cell cultural response of polycaprolactone scaffolds designed and fabricated via fused deposition modeling, *Journal of Biomedical Materials Research,* Vol. 55, No. 2, pp. 203-216 ISSN 0021-9304

Johnson, D.W.J. (1985). Sol-gel processing of ceramics and glass, *American Ceramic Society Bulletin*, Vol. 64, No. 12, pp. 1597-1602, ISSN 0002-7812

Joshua, D.Y., Damron, M., Tang, G., Zheng, H., Chu, C.-J. & Osborne, J.H. (2001). Inorganic/organic hybrid coatings for aircraft aluminum alloy substrates. *Progress in Organic Coatings*, Vol. 41, No. 4, pp. 226-232, ISSN 0300-9440

Judeinstein, P. & Sanchez, C. (1996). Hybrid organic-inorganic materials: a land of multidisciplinarity, *Journal of Materials Chemistry*, Vol. 6, No. 4, pp. 511-525, ISSN 0959-9428

Keefer, K.D. (1984). The effect of hydrolysis conditions on the structure and growth of silicate polymers, *Materials Research Society Symposium Proceedings,* Vol. 32, No. Better Ceramic Chemistry, pp. 15-24, ISSN 0272-9172

Khor, H.L., Ng, K.W., Schantz, J.T., Phan, T.-T., Lim, T.C., Teoh, S.H. & Hutmacher, D.W. (2002). Poly(e-caprolactone) films as a potential substrate for tissue engineering an epidermal equivalent. *Materials Science and Engineering: C,* Vol. 20, No. 1-2, pp. 71-75, ISSN 0928-4931

Kitsugi, T., Nakamura, T., Oka, M., Yan, W.Q., Goto, T., Shibuya, T., Kokubo, T. & Miyaji, S. (1996). Bone bonding behavior of titanium and its alloys when coated with titanium oxide (TiO_2) and titanium silicate (Ti_5Si_3), *Journal of Biomedical Materials Reserch*, Vol. 32, No. 2, pp. 149–56, ISSN 1097-4636

Klein, L.C. & Garvey, G.J. (1982). Silicon alkoxides in glass technology, *ACS Symposium Series*, Vol. 194, No. Soluble Silicon, pp. 293-304, ISSN 0097-6156

Klein, L.C. (1988). *Sol-Gel Technology for Thin Films, Fibers, Preforms, Electronics and Specialty Shapes*, Noyes, ISBN 0-8155-1154-X, Park Ridge, New Jersey, U.S.A.

Klukowska, A., Posset, U., Schottner, G., Wis, M.L., Salemi-Delvaux, C. & Malatesta, V. (2002). Photochromic hybrid sol-gel coatings: preparation, properties, and applications, *Materials Science*, Vol. 20, No. 1, pp. 95-104, ISSN 0137-1339

Li, P., de Groot, K. & Kokubo, T. (1996). Bioactive $Ca_{10}(PO_4)_6(OH)_2$–TiO_2 composite coating prepared by sol–gel process. *Journal of Sol–gel Science and Technology.* Vol. 7, No. 1, pp. 27–34, ISSN 0928-0707

Macey, H.H. (1942). Clay-water relationships and the internal mechanism of drying, *Transactions of the British Ceramic Society*, Vol. 41, pp. 73-120, ISSN 0371-5469

Mackenzie, J.D. (1988). Applications of the sol-gel process. *Journal of Non-Crystalline Solids*, Vol. 100, No. 1-3, pp. 162-168, ISSN 0022-3093

Mackenzie, J.D. & Bescher, E.P. (1998). Structures, properties and potential applications of Ormosils, *Journal of Sol-Gel Science and Technology*, Vol. 13, No. 1/2/3, pp. 371-377, ISSN 0928-070

Matsuura, Y., Matsukawa, K., Kawabata, R., Higashi, N., Niwa, M. & Inoue, H. (2001). Synthesis of polysilane-acrylamide copolymers by photopolymerization and their application to polysilane-silica hybrid thin films, *Polymer*, Vol. 43, No. 4, pp. 1549-1553, ISSN 0032-3861

Moore, F. (1961). Mechanism of moisture movement in clays with particular reference to drying-concise review, *Transactions of the British Ceramic Society*, Vol. 60, pp. 517-539, ISSN 0371-5469

Morrison, R.T. & Boyd R.N. (1966). *Organic Chemistry*, Allyn & Bsacon, ISBN 0136436773, Boston, U.S.A.

Novak, B.M. (1993). Hybrid nanocomposite materials - between inorganic glasses and organic polymers, *Advanced Materials*, Vol. 5, No. 6, pp. 422-433, ISSN 1521-4095

Ohtsuki, C., Kokubo, T. & Yamamuro, T. (1992). Mechanism of apatite formation on $CaOSiO_2P_2O_5$ glasses in a simulated body fluid. *Journal of Non-Crystalline Solids*, Vol. 143, No. 1, pp. 84-92, ISSN 0022-3093

Paul, A. (1990). *Chemistry of Glasses*, Chapman and Hall, ISBN 0412230208, New York, U.S.A.

Ragel, C.V. & Vallet-Regí, M. (2000). In vitro bioactivity and gentamicin release from glass-polymer-antibiotic composites, *Journal of Biomedical Materials Research*, Vol. 51, No. 3, pp. 424-429, ISSN 1552-4965

Sanchez, C. & Ribot, F. (1994). Design of hybrid organic-inorganic materials synthesized via sol-gel chemistry. *New Journal of Chemistry*, Vol. 18, No. 10, pp. 1007-1047, ISSN 1144-0546

Schmidt, H., Scholze, H. & Kaiser, A. (1984). Principles of hydrolysis and condensation reaction of alkoxysilanes, *Journal of Non-Crystalline Solids*, Vol. 63, No. 1-2, pp. 1-11, ISSN 0022-3093

Shih, W.Y., Alkay, I.A. & Kikuchi, R. (1987), Phase diagrams of charged colloidal particles, *Journal of Chemistry and Physics*, Vol. 86, No. 9, pp. 5127-5132, ISSN 0021-9606

Spanhel, L., Popall, G. & Muller, G. (1995). Spectroscopic properties of sol-gel derived nanoscaled hybrid materials, *Journal of chemical sciences*, Vol. 107, No. 6, pp. 637-644, ISSN 0974-3626

Stauffer D., Coniglio A. & Adam M. (1982). Gelation and critical phenomena. *Advances in Polymer Science*, Vol. 44, No. Polymer Networks, pp. 103-158, ISSN 0065-3195

Stockmayer W.H. (1945). *Advancing Fronts in Chemistry. Vol. 1. High Polymers,* Reinhold Publishing Company, New York, U.S.A.

Stoeber, W., Fink, A. & Bohn, E. Controlled growth of monodisperse silica spheres in the micron size range, *Journal of Colloid and Interface Science,* Vol. 26, No. 1, pp. 62-69, ISSN 0021-9797

Teoli, D., Parisi, L., Realdon, N., Guglielmi, M., Rosato, A. & Morpurgo, M. (2006). Wet sol-gel derived silica for controlled release of proteins, *Journal of Controlled Release*, Vol. 116, No. 3, pp. 295-303, ISSN 0168-3659

Uhlmann, D.R., Zelinski, B.J.J. & Wnek, G.E. (1984). The ceramist as chemist - opportunities for new materials, *Materials Research Society Symposium Proceedings*, Vol. 32, No. Better Ceramics Through Chemistry, pp. 59-70, ISSN 0272-9172

Ulrich, D.R. (1988). Prospects of sol-gel processes, *Journal of Non-Crystalline Solids*, Vol. 100, No. 1-3, pp. 174-193, ISSN 0022-3093

Vallet-Regì, M., Ramila, A., del Real, R.P. & Perez-Pariente, J. (2000). A new property of MCM-41: drug delivery system, *Chemistry of Materials*, Vol. 13, No. 2, pp. 308-311, ISSN 0897-4756

Vallet-Regí, M. (2001). Ceramics for medical applications, *Journal of the Chemical Society Dalton Transactions*, No. 2, pp. 97-108, ISSN 1472-7773

Vallet-Regí, M. & Arcos, D. (2006). Nanostructured hybrid materials for bone tissue regeneration, *Current Nanoscience*, Vol. 2, No. 3, pp. 179-189

Vallet-Regí, M. (2006). Bone repair and regeneration: possibilities, *Materialwissenschaft und Werkstofftechnik*, Vol. 37, No. 6, pp. 478-484, ISSN 1521-4052

Vallet-Regí, M. (2006). Ordered mesoporous materials in the context of drug delivery systems and bone tissue engineering, *Chemistry - A European Journal*, Vol. 12, No. 23, pp. 5934-5943, ISSN 1521-3765

Voronkov, M.G., Mileshkevich, V.P. & Yuzhelevskii, Y.A. (1978). *Studies in Soviet Science. The Siloxane Bond: Physical Properties and Chemical Transformations*, Consultants Bureau, ISBN 0306109409, New York, U.S.A.

Wang, D.S. & Pantano, C.G. (1992). Surface chemistry of multicomponent silicate gels, *Journal of Non-Crystalline Solids Advanced Materials from Gels*, Vol. 147-148, pp. 115-122, ISSN 0022-3093

Wei, Y., Xu, J., Dong, H., Dong, J.H., Qiu, K. & Jansen-Varnum, S.A. (1999). Preparation and physisorption characterization of d-glucose-templated mesoporous silica sol-gel materials, *Chemistry of Materials*, Vol. 11, No. 8, pp. 2023-2029, ISSN 0897-4756

Zallen R. (1983). *The Physics of Amorphous Solids,* Wiley, ISBN 9783527602797, New York, U.S.A.

Young, S.K., Gemeinhardt, G.C., Sherman, J.W., Storey, R.F., Mauritz, K.A., Schiraldi, D.A., Polyakova, A., Hiltner, A. & Baer, E. (2002). Covalent and non-covalently coupled polyester-inorganic composite materials, *Polymer,* Vol. 43, No. 23, pp. 6101-6114, ISSN . 0032-3861

Part 3

Biomedical Engineering

17

Diabetes Mechanisms, Detection and Complications Monitoring

Dhanjoo N. Ghista[1], U. Rajendra Acharya[2], Kamlakar D. Desai[3],
Sarma Dittakavi[4], Adejuwon A. Adeneye[5] and Loh Kah Meng[6]
[1]Department of Graduate and Continuing Education,
Framingham State University, Framingham, Massachusetts,
[2]School of Engineering, Division of ECE, Ngee Ann Polytechnic,
[3]Mukesh Patel School of Technology Management & Engineering,
[4]Biomedical Engineering department, Osmania University,
[5]Department of Pharmacology, Faculty of Basic Medical Sciences,
Lagos State University College of Medicine, Ikeja,
Department of Pharmaceutical Sciences,
College of Pharmacy,
University of Kentucky, Kentucky,
[6]VicWell Biomedical Private Limited,
[1,5]USA
[2,6]Singapore
[3,4]India
[5]Lagos State, Nigeria

1. Introduction

Historically, the word *diabetes* was coined from the Greek word meaning a *siphon* by the 2nd century Greek physician, Aretus the Cappadocian. He used the word to connote a condition of passing water (urine) like a *siphon*. Later the Latin description *mellitus* meaning *sweetened or honey-like* was added. Put together, the term *diabetes mellitus* was literarily used to denote a disease condition which was associated with *the persistent passage of sweetened urine* (Krall & Braser, 1999).

In 1999, the World Health Organization described diabetes mellitus as a metabolic disorder of multiple aetiology characterized by chronic hyperglycaemia (the fasting blood glucose level equal or above 200 mg/dl taken at least twice, on different occasions) with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action, or both. In other words, diabetes mellitus is a chronic disease with insidious onset in which the fasting blood glucose is persistently raised above the normal range values, the normal range being between 60 to 120 mg/dl of blood [Krall & Braser, 1999]. It occurs either because of a lack of insulin (the hormone responsible for glucose metabolism), or due to the presence of certain factors opposing the action of insulin on the body tissues that are involved in glucose metabolism, particularly, the liver and the skeletal muscles.

The consequence of insufficient insulin action is hyperglycaemia which may be associated with many associated metabolic abnormalities notably the development of hyperketonaemia resulting from disordered protein metabolism, and derangements in fatty acid and lipids metabolism. If the fasting blood glucose lies between 100 to 130 mg/dl, it is referred to as *Prediabetes* which is associated with an increased tendency or potential of developing *frank* diabetes. A fasting blood glucose of 140 mg/dl or higher is consistent with either type of diabetes mellitus, particularly, when accompanied by classic symptoms of diabetes [Diabetes Control and Complication Trial Research Group, 1997].

2. Diabetes mechanisms

Defects in glucose metabolizing machinery (such as defective insulin secretion, insulin action due to de-expression of insulin receptors or insensitivity of expressed insulin receptors and glucose transporters, decreased peripheral glucose utilization and defective glucose metabolizing enzymes, *etc.*) and consistent efforts of the physiological system to correct the imbalance in glucose metabolism or maintain glucose homeostasis (such as increased insulin secretion, lipolysis, gluconeogenesis, glycogenolysis, *etc.*) place an over exertion on the endocrine system, resulting in hyperglycaemia. The persistent chronic exposure of pancreatic β-cells to the supraphysiological glucose concentrations (hyperglycaemia) results in non-physiological and potentially irreversible β-cell damage, a term known as glucose toxicity which is a gradual, time-related onset of irreversible lesion to pancreatic β-cellular components of insulin content and secretion.

Multiple biochemical pathways and cellular mechanisms for glucose toxicity have been identified and these include glucose autoxidation (resulting from oxidative stress in the presence of chronic hyperglycaemia), protein kinase C (PKC) activation, increased flux through the hexosamine biosynthesis pathway (HBP), formation of advanced glycation end-products (AGEs), altered polyol pathway flux and altered gene expression. However, all these pathways share in common the formation of highly reactive oxygen intermediates (ROIs) or reactive oxygen species (ROS) which in excess amount and on prolonged exposure induce chronic oxidative stress on the pancreatic β-cell population, which in turn causes defective insulin gene expression and insulin secretion as well as increase pancreatic β-cell death.

Hyperglycaemia leads to the production of ROS which modulates various biological functions by stimulating transduction signals, some of which are involved in the pathogenesis of diabetes mellitus. Thus, redox-sensitive signalling pathways have been shown to play a pivotal role in the development, progression, and damaging effect on β-cells population within the pancreatic islet of Langerhans. In the pancreatic tissues, as hyperglycaemia worsens, the redox-sensitive signalling pathways mediating insulin synthesis, storage and release from the pancreatic β-cells becomes compromised progressively. In addition, the oxidative stress induced by chronic hyperglycaemia promotes pancreatic β-cells apoptosis which ultimately resulting in an overt reduction in the insulin secreting pancreatic β-cells population. The hallmarks of these molecular events are pancreatic β-cells failure and hypoinsulinaemia, which constitute the major pathogenic factors in type 1 diabetes mellitus.

Similarly, chronic hyperglycaemia-induced oxidative stress (the presence of an excess amount of reactive oxygen intermediates, due to an imbalance between their formation and degradation as a result of chronic hyperglycaemia) has been considered a proximate cause and common pathogenic factor for tissue/systemic complications of diabetes such as endothelial cells (micro- and macro-angiopathies), nerve cells (neuropathy), proximal renal epithelial cell (nephropathy), pancreatic β-cells (pancreatic β-cell failure) through lipid peroxidation and glycation mechanisms in these organs. Hyperglycaemia has been shown to result in glycation (a non-enzymatic conjugation of glucose to proteins leading to the formation of advanced glycation (glycosylation) end-products (AGEs) and tissue damage. Increased glycation and build-up of tissue AGEs have been implicated in the aetiology of diabetes mellitus, its complications and progression because they alter glucose metabolizing enzyme activity, decrease ligand binding, modify protein half-life and alter immunogenicity.

One mechanism by which the effects of glucose toxicity result in chronic hyperglycaemia are thought to be mediated is oxidative stress [Baynes, 1991; Evans *et al.*, 2002], and hyperglycaemia is known to be one of the main causes of oxidative stress in type 2 diabetes mellitus [Bonnefont-Rousselot, 2002; Robertson *et al.*, 2003]. Oxidative stress is a state of imbalance between free radical generation and mopping up.

Oxidative stress is known to play a pivotal role in the pathogenesis of insulin resistance which is itself is thought to be mediated via its contribution to glucose toxicity, particularly, in insulin target tissues including the pancreatic β-cells [Gleason *et al.*, 2000; Fantus, 2004]. Tissues such as the mesangial cells (in the kidneys), retinal cells and pancreatic islets are least endowed with intrinsic antioxidant enzyme expression, including *superoxidases-1* and *-2*, *catalase* and *glutathione peroxidase* [Hayden & Tyagi, 2002; Robertson, 2004]. Prolonged exposure of pancreatic β-cell to hyperglycaemia, as in diabetes, results in decreased expression of the antioxidant gene *γ-glutamylcysteine ligase* (*γ-GCL*) and down-regulation of the rate-limiting enzyme for glutathione synthesis [Robertson, 2004]. The *γ-GCL* catalyses the rate-limiting step in the synthesis of γ-glutamyl cysteine from cysteine, which forms the substrate for the second enzyme regulating glutathione synthesis [Yoshida *et al.*, 1995; Tanaka *et al.*, 2002]. Reduced gluthathione plasma and tissue concentrations, as marked by elevated levels of ceruloplasmin, promote free radical generation, production of advanced glycation products (AGEs) and acute flunctuations in glucose concentrations.

In addition, oxidative stress promotes the onset and development of diabetes mellitus by directly decreasing insulin sensitivity and causing direct cytotoxicity to the pancreatic insulin-producing β-cells [Maiese *et al.*, 2007]. The generated ROS penetrates through the cell membranes and reacts with the membrane phospholipids through the process of lipid peroxidation as well as reacts with the mitochondrial DNA to distrupt the mitochondrial respiratory machinery (mitochondrial electron transport) which is regulated by NADPH ubiquinone oxidoreductase and ubiquinone-cytochrome c reductase systems [Maiese *et al.*, 2007].

Oxidative stress is known to depress the mitochondrial oxidoreductase and citrate synthase activities resulting in significant reductions in mitochondrial oxidative and phosphorylation activities as well as reduces the levels of mitochondrial proteins and mitochondrial DNA in adipocytes, particularly in type 2 diabetes mellitus (Petersen et al., 2003). Oxidative stress has been shown to trigger the opening of the mitochondrial membrane permeability transition pore which results in a significant depletion of mitochondrial NAD+ stores and subsequently apoptotic cell injury (Maiese et al., 2007). In the pancreatic tissues, these cellular events result in depletion of the β-cells population, insulin deficiency while in the skeletal muscle, it manifests as insulin resistance.

Oxidative stress is also known to modify a number of cellular signalling pathways that can results in insulin resistance. For example, a significant increase in muscle protein carbonyl content (often used as a reliable biological marker of oxidative stress) and elevated levels of malondialdehyde and 4-hydrononenal (as reliable indicators of lipid peroxidation) have been implicated in the aetiology of insulin resistance diabetes mellitus [Haber *et al.*, 2003].

3. Glucose-insulin regulatory system modeling and simulation of OGTT blood glucose concentration dynamics to obtain indices for diabetes risk and detection

This section deals with the bioengineering modelling of the glucose-insulin regulatory system and the OGTT blood glucose dynamics data, for more reliable detection of diabetes as well as designation of risk to diabetes.

The conventional way of diagnosing diabetes is based on designation of specific values of fasting plasma glucose equal or greater than 126 mg/dl (7.0 mmol/l), and (ii) 2-hour plasma glucose concentration equal or greater than 200 mg/dl (11.1 mmol/l) during OGTT. Instead of this rigid approach, we are proposing that for more reliable monitoring and diagnosis of diabetes, it is more relevant to mathematically characterise the trend of blood glucose concentration rise and decline after an oral intake of 75 g glucose load in OGTT. Hence, we provide the bioengineering analysis of the Glucose-insulin regulatory system and glucose response data, leading to the formulation of a novel nondimensional diabetes index for diagnosis of diabetic patients as well as of those who are at risk of becoming diabetic.

So, in this section, we present the Glucose-Insulin Regulatory System (GIRS) modeling in the form of governing differential equations, and converge to the equation representing blood glucose response to glucose infusion rate. This equation forms the basis of modeling of the Oral Glucose Tolerance Test (OGTT). We then demonstrate how this OGTT model equation's solutions can simulate the OGTT data, to evaluate the model parameters distinguishing diabetes subjects from normal subjects. The climax to this section is the formulation of the Non-dimensional Diabetes Index (DBI), involving combination of the model parameters into just "one number" by which we can reliably detect diabetes. In fact, by determining the range of values of DBI for a big patient population, we can even detect "patients at risk of being diabetic".

3.1 Differential equation model of the glucose-insulin system

With reference to the Blood Glucose-Insulin Control System (depicted in Fig. 1), the corresponding first-order differential equations of the insulin and glucose regulatory sub-systems are given by equations (1) and (2) [Dittakavi et al., 2001].

$$x' = p - \alpha x - \beta y \tag{1}$$

$$y' = q - \gamma x - \delta y \tag{2}$$

where x′ and y′ denote the first time-derivatives of x and y, x: insulin output, y: glucose output, *p*: insulin input, *q*: glucose input, for unit blood-glucose compartment volume (*V*). In these equations, the glucose-insulin model system parameters (regulatory coefficients) are $\alpha, \beta, \gamma, \delta$.

These coefficients, when multiplied by the blood-glucose compartment volume *V* (which is proportional to the body mass) denote, respectively,

- the sensitivity of insulinase activity to elevated insulin concentration (αV),

- the sensitivity of pancreatic insulin output to elevated glucose concentration (βV),
- the combined sensitivity of liver glycogen storage and tissue glucose utilization to elevated insulin concentration (γV), and
- the combined sensitivity of liver glycogen storage and tissue glucose utilisation to elevated glucose concentration (δV).

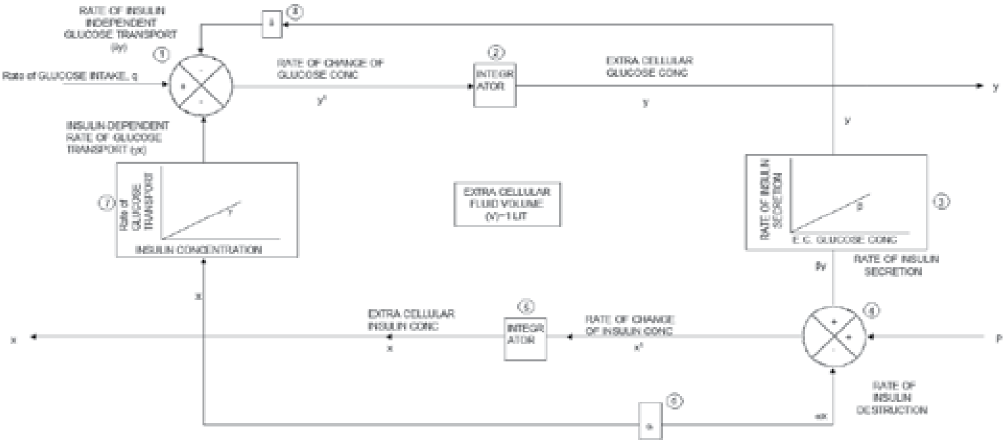

RATE OF INSULIN INDEPENDENT GLUCOSE TRANSPORT (δy)
①
Rate of GLUCOSE INTAKE, q
INSULIN DEPENDENT RATE OF GLUCOSE TRANSPORT (γx)
RATE OF CHANGE OF GLUCOSE CONC
y'
②
INTEGR ATOR
EXTRA CELLULAR GLUCOSE CONC
y
y
y
⑧
δ
⑦
Rate of GLUCOSE TRANSPORT
γ
INSULIN CONCENTRATION
EXTRA CELLULAR FLUID VOLUME (V)=1 LIT
RATE OF INSULIN SECRETION
β
E.C. GLUCOSE CONC
③
βy
RATE OF INSULIN SECRETION
④
x
EXTRA CELLULAR INSULIN CONC
x
⑤
INTEGR ATOR
RATE OF CHANGE OF INSULIN CONC
x'
p
x
⑥
α
αx
RATE OF INSULIN DESTRUCTION

Fig. 1. Physiological model of the Blood Glucose Control system
(represented by equations 1 and 2).

From equations (1) and (2), the differential equation model in glucose concentration (y) for insulin infusion rate (p = 0) and glucose in flow rate (q), is obtained as

$$y'' + y'(\alpha + \beta) + y(\alpha\delta + \beta\gamma) = q' + \alpha q \tag{3}$$

where y' and y'' denote first and second time derivatives of y.
The transfer-function corresponding to Eqn. (3) is obtained by taking the Laplace transforms on both sides (assuming the initial conditions to be zero). Thereby, we obtain (for glucose response)

$$Y(s)/Q(s) = \frac{(s+\alpha)}{s^2 + s(\alpha+\delta) + (\alpha\delta + \beta\gamma)} = G(s) \tag{4}$$

3.2 Model analysis to simulate Oral Glucose Tolerance Test (OGTT)

The OGTT model-simulation response curve is considered to be the result of giving an impulse glucose dose (of 4 gm of glucose/liter of blood-pool volume) to the combined system consisting of GI tract and blood glucose concentration (BGCS). Now, we can put down the transfer-function (TF) of the gastro-intestinal (GI) tract to be 1/(s + α), because the intestinal glucose-concentration variation is an exponential decay, and the exponential parameter value is close to that of the parameter α. When we multiply this GI tract TF [1/(s + α)] by the TF of the blood-pool glucose-metabolism given by Eqn. (4), and put Q(S) = 'G' gm of glucose per litre of blood-pool volume per hour, we get

$$\mathrm{Y}(\mathrm{s}) = \mathrm{G} / \left\{s^2 + s(\alpha+\delta) + (\alpha\delta + \beta\gamma)\right\} \tag{5}$$

The corresponding governing differential equation is now:

$$y'' + 2Ay' + \omega_n^2 y \;=\; G\delta(t)$$

or

$$y'' \;+ \lambda T_d y' + \lambda y = G\delta(t) \tag{6}$$

wherein ωn (= $\lambda^{1/2}$) is the natural frequency of the system, A is the attenuation or damping constant of the system, $\lambda = 2\,A\,/\mathrm{Td} = \omega_n^2$, and $\omega = (\omega_n^2 - A^2)^{1/2}$ is the angular frequency of damped oscillation of the system.

The solution of Eq. (6), for an under-damped response (corresponding to that of normal subjects, represented by the lower curve in Fig. 2) is given by

$$y(t) = \;(G\,/\omega)e^{-At}\,\sin\,\omega t, \tag{7}$$

where in ω (or ω_d) = $(\omega_n^2 - A^2)^{1/2}$.

The solution for over-damped response (corresponding to that of diabetic subjects, represented by the upper curve in Fig 2) is given by:

$$y(t) = (G/\omega)e^{-At}\,\sinh\,\omega t \tag{8}$$

where in ω (or ω_d) = $(A^2 - \omega_n^2)^{1/2}$

The solution for a critically-damped response (in which $A = \omega_n$), which applies to subjects at risk of becoming diabetic (whose blood glucose response curve would lie between the two curves of normal and diabetic subjects), is given by:

$$y(t) = G\,t\,e^{-At}; \tag{9}$$

for $\omega_n^2 = A^2 = \lambda$, and derivative-time period $T_d = \dfrac{2A}{\lambda} = \dfrac{2A}{\omega_n^2}$

These solutions are employed to simulate the clinical data, and to therefore evaluate the model-system parameters A and ω (or λ and T_d), to not only differentially-diagnose diabetes subjects as well as sbut also to characterize resistance-to-insulin.

Now, we can employ equations (7) and (8) to simulate the OGTT data shown in Fig. 2 to obtain the value of parameters: (i) $\lambda = 2.6\mathrm{hr}{-2}$, $Td = 1.08$ hr, for the normal subject, and (ii) $\lambda = 0.27\mathrm{hr}{-2}$ and $Td = 6.08$ hr, for the diabetic subject [Ghista, 2004].

We now formulate the Non-dimensional Diabetes Index (*DBI*), as

$$DBI = AT_d = \frac{2A^2}{\lambda} = \frac{2A^2}{\omega_n^2} \tag{10}$$

The value of *DBI* for the normal subject is 1.3, whereas for the diabetic subject it is 4.9. We have further found (in our initial clinical tests) that *DBI* for normal subjects is less than 1.6, while the *DBI* for diabetic patient is greater than 4.5. Hence a DBI value of 2-4 can suggest that the subject is at risk of becoming diabetic. This is a testimony of how well we have simulated the OGTT by our BME model and employed this DBI to diagnose diabetes.

$$\begin{array}{ll}
y(t) = \frac{G}{\omega} e^{-ck} \sinh \omega t & y(t) = \frac{G}{\omega} e^{-ck} \sinh \omega t \\
(AT_d = 4.9) & (AT_d = 1.3) \\
(A = 0.808 \ hr^{-1}, \lambda = 0.2657 \ hr^{-2} & (A = 1.4 \ hr^{-1}, \lambda = 2.6 \ hr^{-2} \\
T_d = 6.08 \ hr, & T_d = 1.08 \ hr, \\
G = 2.9464 \ gL^{-1}hr^{-1}) & G = 1.04 \ gL^{-1}hr^{-1})
\end{array}$$

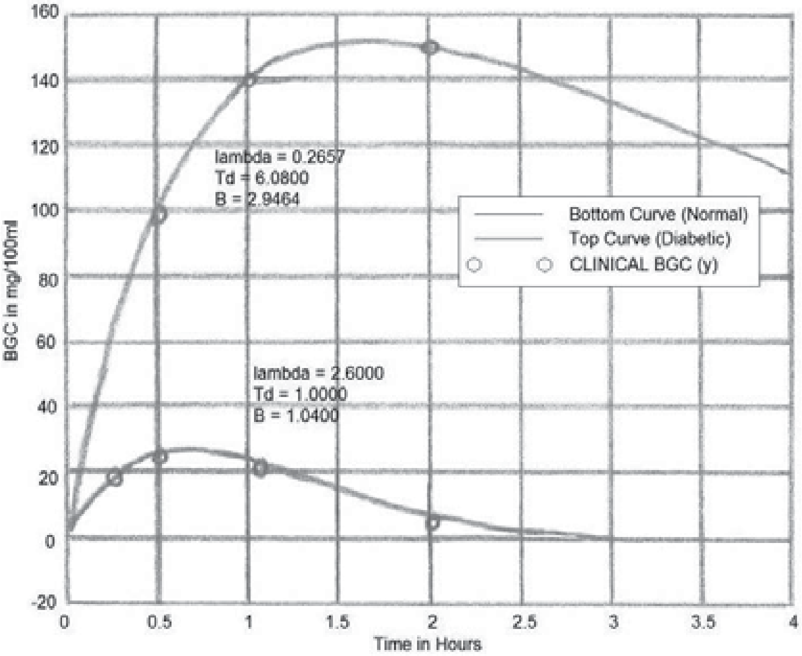


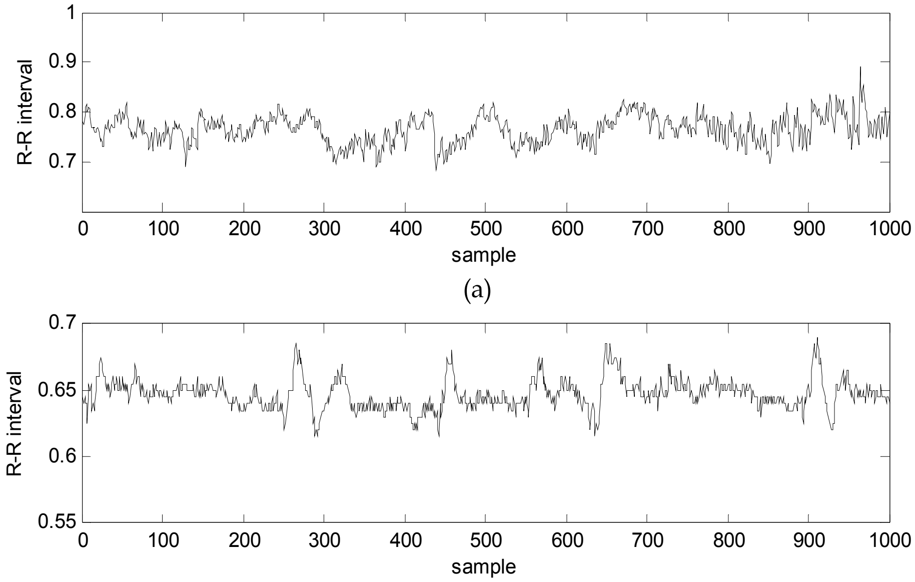
Fig. 2. OGTT Response Curve [Ghista, 2004], showing the glucose concentration responses of normal and diabetic subject.

4. Biomedical signal processing and image processing techniques for diabetes analysis

This section presents different signal and image processing methods that are used to evaluate the effect of diabetes on different organs.

4.1 Analysis of the heart rate variability signal

Heart rate variability (HRV) decreases in patients with diabetes [Acharya et al., 2006; Acharya et al., 2011b; Faust et al., 2011]. This variability can be analyzed in the time domain, frequency domain, and by using non-linear methods. Fig. 3 shows typical HRV signals of normal and diabetes subjects. Visually, it is difficult to notice the variability in these two signals. Hence, analysis in time domain and frequency domain with the use of non-linear methods is necessary. These methods are explained in this section.



(a)

(b)

Fig. 3. Typical heart rate signals; (a) normal (b) diabetes.

4.1.1 Time domain analysis

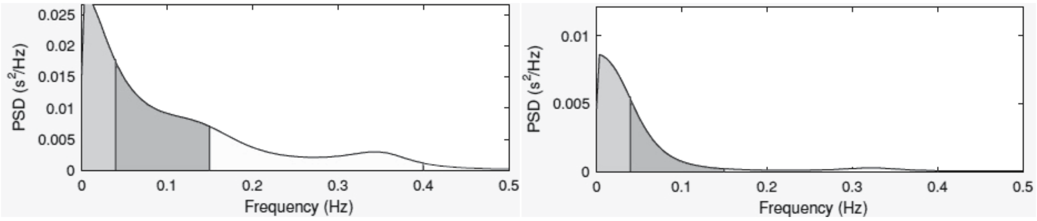
The time-and frequency-domain measures of HRV were analyzed by the Task Force of the European Society of Cardiology [Task Force, 1996]. Several time domain parameters are calculated from the original R-R interval: mean R-R interval, standard deviation of the NN intervals (SDNN), standard deviation of differences between adjacent RR (NN) intervals (SDSD), Standard Error, or Standard Error of the Mean (SENN), which is an estimate of the standard deviation of the sampling distribution of means based on the data, number of successive difference of intervals which differ by more than 50 msec expressed as a percentage of the total number of ECG cycles analyzed (pNN50%).

The HRV triangular index (TINN) is the integral of the density distribution (i.e. the number of all NN intervals) divided by the maximum of the density distribution. Thus, six standard measures namely Mean RR, SDNN, SENN, SDSD, pNN50% and TINN were studied.

4.1.2 Frequency domain analysis

Spectral analysis of HRV signal results in three main components: high frequency (HF) component, low frequency (LF) component, and very low frequency (VLF) component [Task Force, 1996]. The influence of the vagus nerve in modulating the sinoatrial node is indicated by the HF component (0.15Hz -40Hz) of the spectrum. The LF component (0.04Hz-.155 Hz) indicates the sympathetic effects on the heart. The VLF component (0.003Hz -.04 Hz) explains many details of the heart, chemoreceptors, thermareceptors, and renin-angiotensin system [Task Force, 1996; Kamath et al., 1987; Van der Akker et al., 1983].

Fig. 4 shows a typical power spectral density (PSD) distribution of the heart rate signals obtained from a normal subject (Fig. 4-a) and a diabetes patient (Fig. 4-b). The beat to beat variation is greater in the normal heart rate signal compared to the diabetes heart rate signal. Hence, the power spectral density is more predominant in HF in the normal subject[Faust et al., 2011].



(a) (b)

Fig. 4. Typical power spectral density of heart rate signal (a) normal (b) diabetes subject. The PSD of normal heart rate signal has LF, HF components. The diabetic heart rate signal, however, does not have HF components due to lower variability in the heart rate signal [Acharya et al., 2011b].

4.1.3 Non-linear parametric analysis of heart rate signals

Various non-linear parameters can be used to analyze the diabetes heart rate signals. They are Approximate Entropy (ApEn), Correlation Dimension (CD), Largest Lyapunov Exponent (LLE), The Hurst exponent (H), Recurrence plot (RP), and Fractal Dimension (FD).

The Approximate Entropy *ApEn* measures regularity of the time series. The method proposed by Pincus et al can be used to evaluate the ApEn [Pincus, 1991]. For the data points $x(1), x(2), ..., x(N)$, with an embedding dimension *m*, the ApEn or $APEN$ is given by:

$$APEN(m,r,N) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\log C_i^m(r) - \frac{1}{N-m}\sum_{i=1}^{N-m}\log C_i^{m+1}(r) \quad (11)$$

where $C_i^m(r) = \frac{1}{N-m+1}\sum_{j=1}^{N-m+1}\Theta\left(r - \left\|\mathbf{x}_i - \mathbf{x}_j\right\|\right)$ is the correlation integral. For this study, m is set to 2, and r is chosen as 0.15 times the standard deviation of the original data sequence, and N is the total number of data points.

The *Correlation dimension (CD)* is a quantitative measure of the informational complexity of the heart rate signal [Grassberger, 1983]. Some unique ranges of CD for different cardiac diseases have been proposed by Acharya et al. [2007]. The formula for CD involves the correlation function *C*(*r*), which is the probability that two arbitrary points on the orbit are closer together than *r*. This is done by calculating the separation between every pair of *N* data points and sorting them into bins of width d*r* proportionate to *r*. The correlation dimension can be calculated by using the distances between each pair of points in the set of N number of points, $a(i,j) = \left|X_i - X_j\right|$

$$C(r) = \frac{1}{N^2} \times \left(Number\ of\ pairs\ of\ (i,j)\ with\ a(i,j) < r\right) \quad (12)$$

Correlation dimension (CD) is given by:

$$CD = \lim_{r \to 0}\frac{\log\left(C(r)\right)}{\log(r)} \quad (13)$$

The Largest Lyapunov Exponent (LLE) measures the predictability of the system and determines sensitivity of the system to initial conditions [Rosenstien et al., 1993]. A positive LLE indicates chaos. The LLE is estimated by using a least squares fit to "average "line, and is given by:

$$y(n) = \frac{1}{\Delta t}\langle \ln\left(d_i\left(n\right)\right)\rangle \tag{14}$$

where $d_i(n)$ is the distance between i^{th} phase-space point and its nearest neighbor at n^{th} time step, and $\langle . \rangle$ denotes the average overall phase space points.

The Hurst Exponent (HE) indicates the self-similarity and correlation properties of heart rate signal. The *HE* has been defined and proposed by Dangel et al [Dangel et al., 1999]. Unique range of H values has been proposed by Acharya et al, for various cardiac states [Acharya et al., 2007].

$$\mathrm{H} = \log\left(\mathrm{R}/\mathrm{S}\right)/\log\left(\mathrm{T}\right) \tag{15}$$

where T is the duration of the sample of data and R/S is the corresponding value of rescaled range. An *HE* value of 0.5 indicates the presence of a random walk, *HE* < 0.5 depicts anti persistence, and *HE* > 0.5 indicates the persistence in the signal.

The *Recurrence plot (RP)* can be used to unearth the non-stationarity in the heart rate signals [Acharya et al., 2006], and was originally introduced by Eckmann et al. [Eckmann et al., 1987].

A Fractal is a set of points which, when looked at smaller scales, looks similar to the whole group [Madelbrot, 1983]. The Fractal Dimension (FD) determines the complexity of the time series. FD has been used in heart rate analysis to recognise and differentiate specific states of physiologic functions [Acharya et al., 2007].

The heart rate signal is a non-linear and non-stationary signal. The hidden intricacies of the signal can be easily extracted using non-linear analysis methods. The heart rate variation is more random in normal subjects as compared to the diabetes subjects. Hence, most of these non-linear parameters may show distinct values for normal and diabetes subjects. These clinically significant non-linear parameters can be fed into the classifiers as features for automatic classification. Moreover, these non-linear parameters can be combined in the form of an integrated index [Ghista, 2004; 2009a; 2009b]. Such an index may have unique range of values for normal and diabetes classes. Hence, one can diagnose normal and diabetes subjects by just using one index value without the need for automatic classifiers.

4.2 Image processing of digital fundus images in diabetic retinopathy

Diabetic retinopathy is an important complication of diabetes. As the diabetes retinopathy progresses, the number of blood vessels varies, and the exudates appear in the advanced DR stages [Yun et al., 2008; Acharya et al., 2011a]. Different image processing techniques have been used to extract blood vessels and exudates in DR subjects, and these techniques are explained in this section. Moreover, techniques for plantar pressure images analysis, which have proved to be useful in detecting diabetic neuropathy conditions, are also been presented in this section.

4.2.1 Retinal blood vessels detection

The detailed steps involved in the blood vessel detection are shown in Fig. 5 [Nayak et al., 2008; Acharya et al., 2011a; Acharya et al. 2009; Acharya et al., 2011b]. The green component of the RGB (Red, Green Blue) blood vessel image is considered for this study. The border of the image is obtained by applying an edge detection algorithm on the inverted green component of the image. Morphological operation is performed by using a disk shaped structuring element (SE) for blood vessels detection. Adaptive histogram equalization is then performed on these images to enhance the image, and subsequently, morphological opening operation is performed using a ball structuring element. Thresholding is carried out on the resulting image followed by the median filtering to obtain the boundary of the image. The small holes are then filled and the boundary is removed. Finally, the image with only blood vessels is obtained (Fig. 7) [Acharya et al., 2011b]. It can be seen from Fig. 7(a) that the number of blood vessels is different in the normal and the proliferative diabetes retinopathy (PDR) classes.

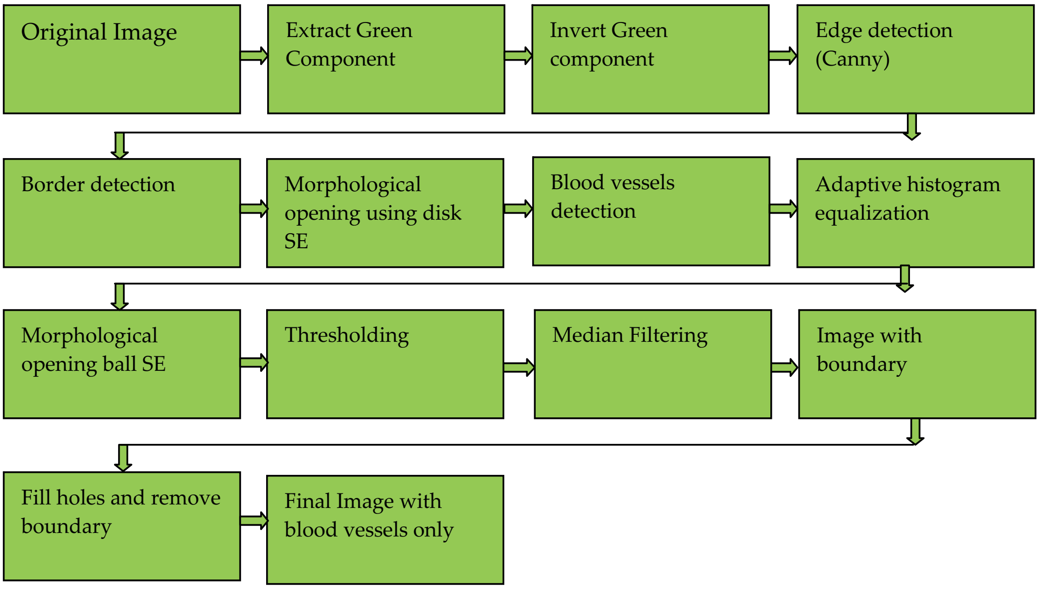

Original Image
Extract Green Component
Invert Green component
Edge detection (Canny)
Border detection
Morphological opening using disk SE
Blood vessels detection
Adaptive histogram equalization
Morphological opening ball SE
Thresholding
Median Filtering
Image with boundary
Fill holes and remove boundary
Final Image with blood vessels only

Fig. 5. The block diagram for detecting retinal blood vessels.

4.2.2 Exudates detection in digital fundus images

Fig. 6 shows the block diagram of the exudates extraction in digital fundus images [Acharya et al., 2008; Nayak et al., 2008; Acharya et al., 2011a; Acharya et al., 2011b]. The green component of the original image is extracted and subjected to the morphological closing operation by using octagonal shaped structuring element. Then, the resulting image is subjected to thresholding, and morphological closing operation is carried out by using disk shaped SE.

The edges are detected by using the Canny method. Subsequently, an 80x80 region of interest (ROI) is considered to remove the optic disc, and then the border of the image is also removed. Finally, by performing morphological erosion operation with disk shaped SE of size 3, the final image with only exudates is obtained (Fig. 7) [Acharya et al., 2011b]. It can be seen from the Fig. 7(b) that there are no exudates in the normal image, while the PDR image has exudates.

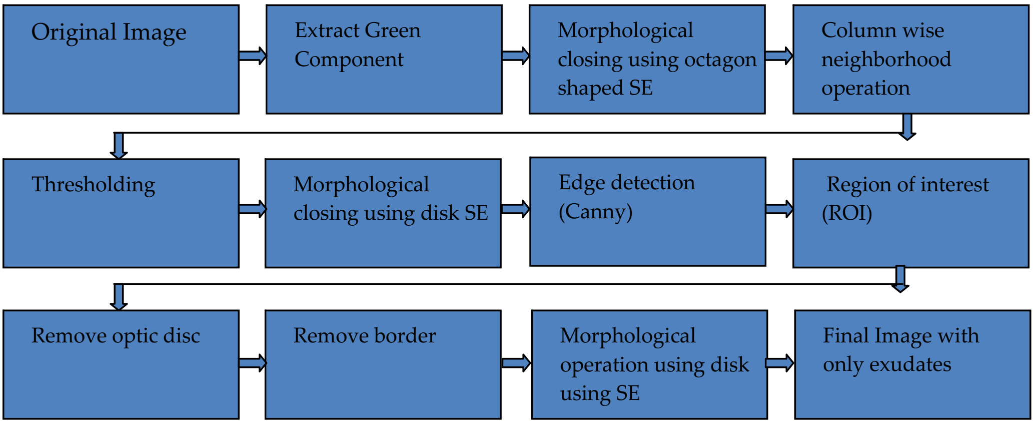

Original Image
Extract Green Component
Morphological closing using octagon shaped SE
Column wise neighborhood operation
Thresholding
Morphological closing using disk SE
Edge detection (Canny)
Region of interest (ROI)
Remove optic disc
Remove border
Morphological operation using disk using SE
Final Image with only exudates

Fig. 6. The block diagram for detecting exudates in digital fundus images.

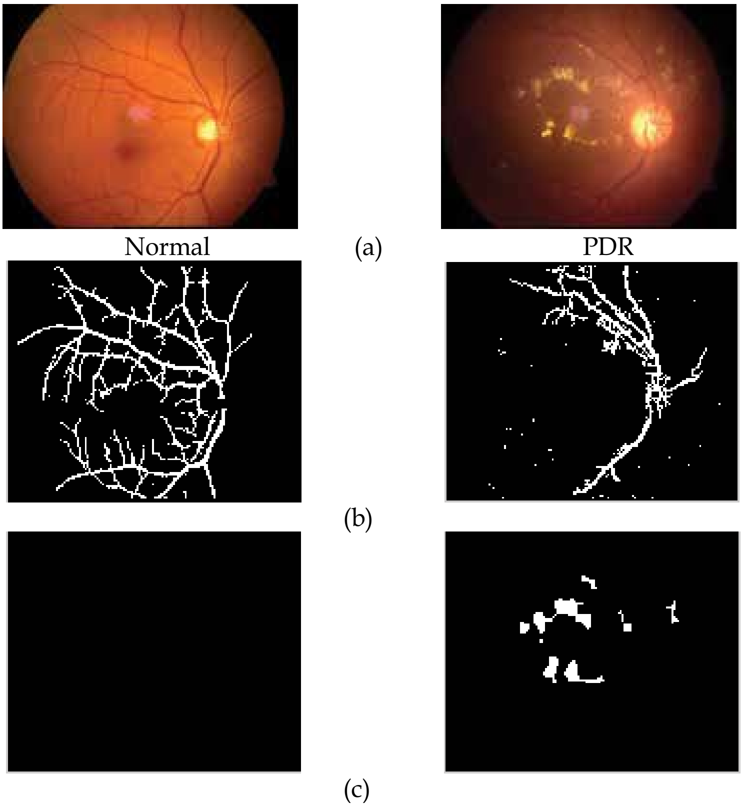
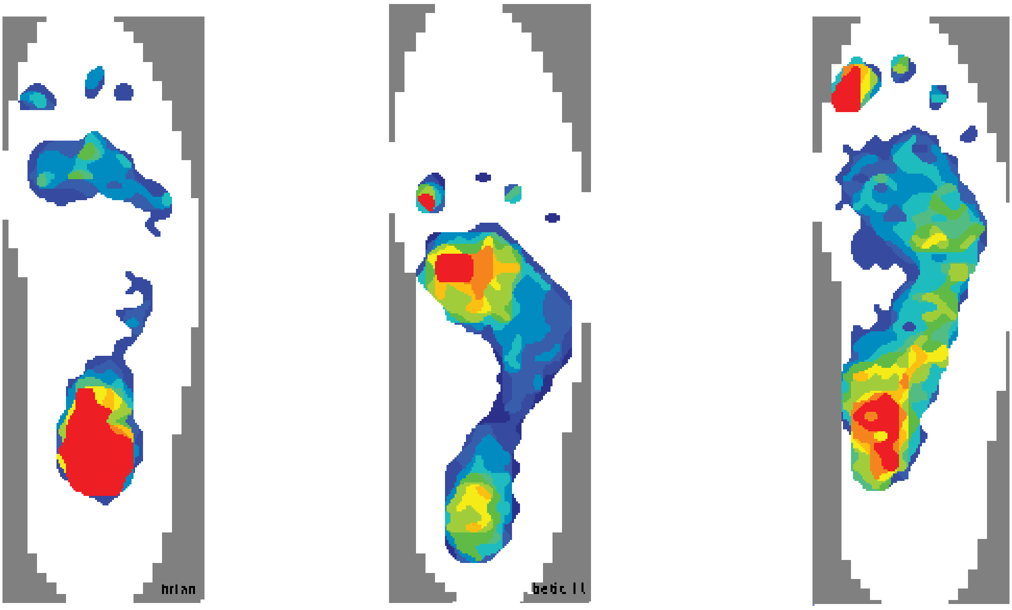

Normal
(a)
PDR
(b)
(c)

Fig. 7. Results of blood vessel detection and exudate detection from normal and PDR images. (a) Original normal and PDR images (b) Results of blood vessel detection (c) Results of exudate detection. The number of blood vessels are different for normal and PDR images, and exudates are absent in the normal fundus image.

4.3 Plantar pressure distribution image analysis

Fig. 8 shows the plantar pressure distribution images of normal subjects, and subjects with diabetes type II without and with neuropathy. It can be seen from the figure that the pressure distribution is different for normal, diabetes without and with neuropathy subjects [Acharya et al., 2008; Acharya et al., 2011b]. This difference can be further analyzed using Fourier transform and discrete wavelet transform (DWT).



(a) (b) (c)

Fig. 8. Static pedobarograph images of (a) the normal foot, (b) a diabetic foot with neuropathy, and (c) a diabetic foot without neuropathy.

The important feature used to diagnose the normal, diabetes type II with and without neuropathy classes is the power ratio (PR) that is obtained using the Fourier transform [Rahman et al., 2006]. This method is clearly explained below.

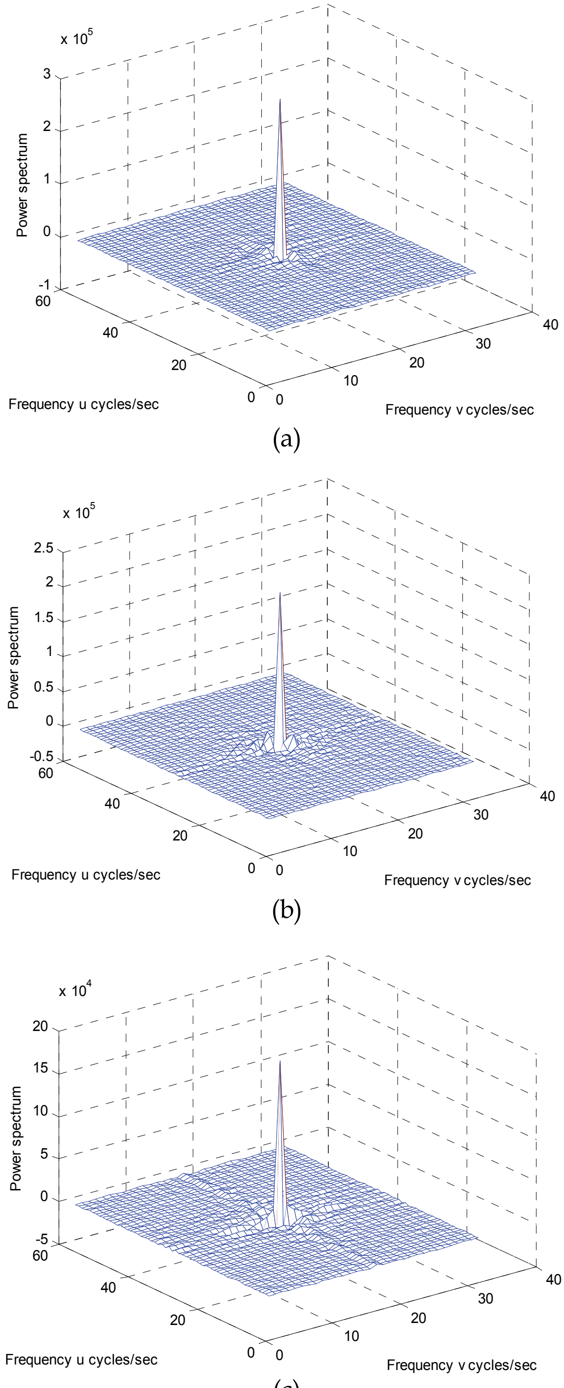
Fourier domain analysis: The Fourier spectrum *F(u,v)* of each region of the image can be obtained by using the below equation (16) [Cavanagh et al., 1991]. In this equation, *M* and *N* represent the numbers of rows and columns of the image. The power ratio (*PR*) is the ratio of the high frequency power (*HFP*) to the total power (*TP*). The Fourier spectrum is given by

$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y).e^{-j2\Pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} \tag{16}$$

where x, y, u and v are the variables.

$F(0,0)$ is the DC component of the image in the frequency domain and is the sum of all the pixels of an image in spatial domain [Cavanagh et al., 1991]. The *total power (TP)* of the image is given by

$$TP = \left\{\sum\sum |F(u,v)|^2\right\} - |F(0,0)|^2 \tag{17}$$



(a)

(b)

(c)

Fig. 9. Typical power spectra after deleting the DC component from region 6 of the left foot for (a) normal subject (b) diabetes subject without neuropathy (c) diabetes subject with neuropathy[Acharya et al., 2011b].

The low frequency and high frequency components are separated by S_o, which is given by

$$S_o = \begin{cases} \dfrac{M}{4} & if\ M \leq N \\ \dfrac{N}{4} & if\ M < N \end{cases} \tag{18}$$

$$LFP = \left\{ \sum_{S(u,v)=0}^{S_o} \left|F(u,v)\right|^2 \right\} - \left|F(0,0)\right|^2 \tag{19}$$

$$HFP = TP - LFP \tag{20}$$

$$PR = \left(\frac{HFP}{TP} \right) x100 \tag{21}$$

where *LFP, HFP*, and *PR,* denote the low frequency power, high frequency power, and the power ratio, respectively.

Fig. 9 shows the typical power spectra obtained for a normal subject, having diabetes without neuropathy, and subject having diabetes with neuropathy. It is a 3D figure, with *u* and *v* frequencies corresponding to row and column. The Y-axis indicates the power. The power spectrum of normal class has a peak in the centre and very small peaks around it. In the case of diabetes without neuropathy, the adjacent peaks are slightly larger; in the case of diabetes with neuropathy, there are dominating peaks on four sides. These plots are unique and depict variation of power spectrum. The PR values extracted from various regions of the plantar image are shown in Table 1[Acharya et al., 2011b].

Type	Control subjects (CS)	Diabetic control (DC)	Neuropathic (N)	p-value
Region 1	12.80 ± 3.49	9.562 ± 2.25	17.657 ± 3.27	<0.0001
Region 2	11.865 ± 2.13	9.678 ± 2.58	14.453 ± 2.31	<0.0001
Region 5	13.769 ± 3.31	9.512 ± 2.530	14.542 ± 2.69	<0.0001
Region 6	10.179 ± 2.09	9.697 ± 1.23	12.35 ± 2.19	<0.0001
Region 7	9.28 ± 6.03	8.67 ± 3.30	11.56 ± 1.45	<0.0001

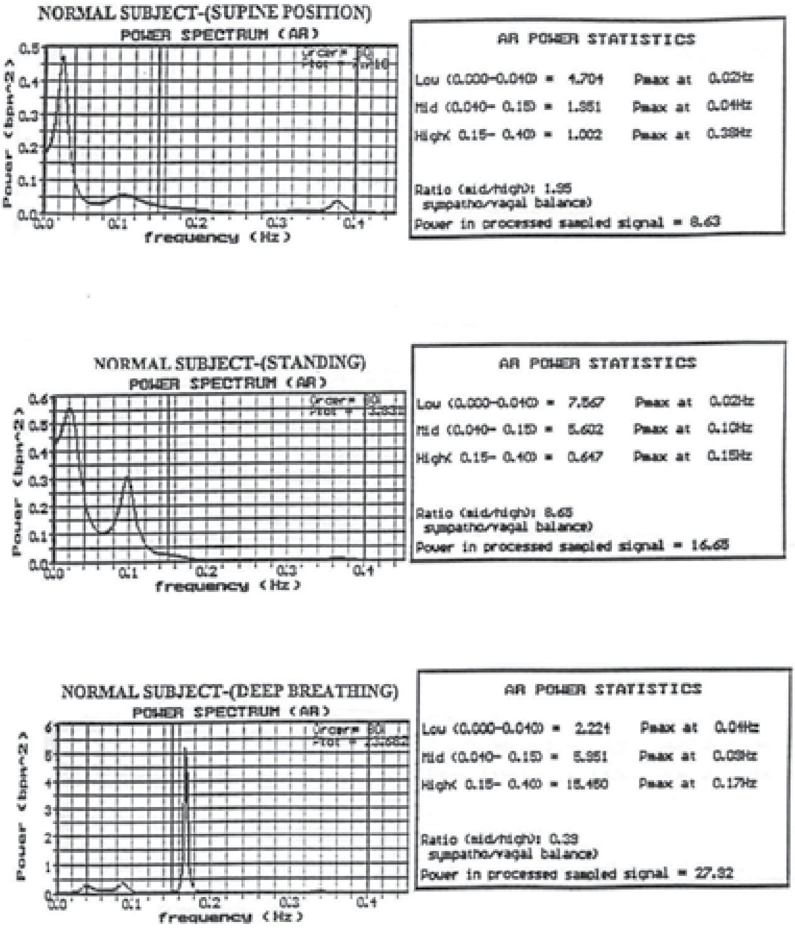
Table 1. Power ratio values for the various regions of the plantar pressure images obtained from the three classes.

The PR is the ratio of HF power to the total power. This value is higher for diabetes subjects with neuropathy when compared to the normal and diabetes without neuropathy subjects for regions 1, 2, 5, 6, and 7 (Table 1). These ranges are unique and clinically significant ($p<0.0001$). These PR features can be used to diagnose the three classes automatically using classifiers.

Likewise, DWT coefficients have also been used to identify the normal, diabetes type II with and without neuropathy classes [Acharya et al., 2008; Acharya et al., 2011b].

5. Diabetic autonomic neuropathy diagnosis from HRV power spectrum plots

The RR interval files are processed to get HRV and HRVPS [Desai, K.D et al., 2011]. The sampling frequency used to get HRV form RR file is 2Hz. The power spectrum plots depict power in $(BPM)^2$ versus Frequency (in Hertz). The auto regression statistics gives display of the following parameters:
Power under Low frequency range: frequency range from 000 to 0.04Hz
Power under Mid frequency range: frequency range from 0.04 to 0.15Hz
Power under High frequency range: frequency range from 0.15 to 0.40Hz
Sympatho/Vagal balance ratio: ratio of mid to high frequency powers
The Sympatho-Vagal ratio is found in the different frequency characteristics of the parasympathetic and sympathetic influences on heart rate. The HRVPS plots (for the supine, standing and deep breathing modes) are plotted with time-scale up to 150 seconds and heart rate scale in the range of 40 bpm to 140 bpm.



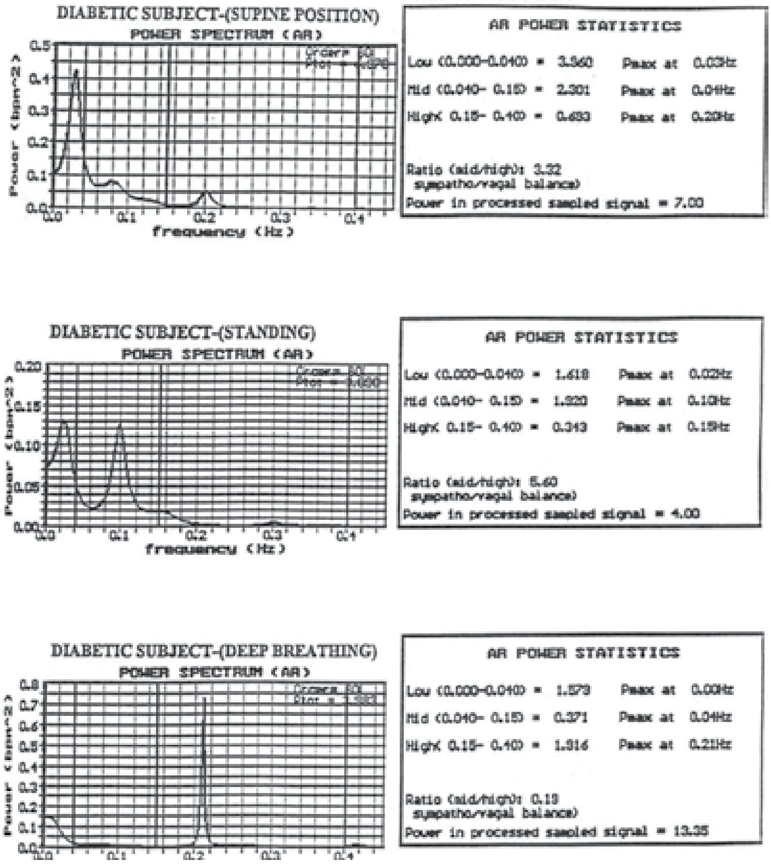
AR POWER STATISTICS
Low (0.000-0.040) = 4.704 Pmax at 0.02Hz
Mid (0.040- 0.15) = 1.951 Pmax at 0.04Hz
High(0.15- 0.40) = 1.002 Pmax at 0.38Hz
Ratio (mid/high): 1.95
sympatho/vagal balance)
Power in processed sampled signal = 8.63

AR POWER STATISTICS
Low (0.000-0.040) = 7.567 Pmax at 0.02Hz
Mid (0.040- 0.15) = 5.602 Pmax at 0.10Hz
High(0.15- 0.40) = 0.647 Pmax at 0.15Hz
Ratio (mid/high): 8.65
sympatho/vagal balance)
Power in processed sampled signal = 16.65

AR POWER STATISTICS
Low (0.000-0.040) = 2.221 Pmax at 0.04Hz
Mid (0.040- 0.15) = 5.951 Pmax at 0.09Hz
High(0.15- 0.40) = 15.460 Pmax at 0.17Hz
Ratio (mid/high): 0.39
sympatho/vagal balance)
Power in processed sampled signal = 27.32

Fig. 10. HRVPS plots of a normal subject in supine, standing and deep breathing modes. The power statistics on the right side show the power in low, medium and high frequency bands. There is an increase in the mid frequency power in standing position and in high frequency power in deep breathing mode.

Figure 10 displays the HRVPS of a typical normal subject in supine, standing and deep breathing modes. In this figure, the power statistics show the power in low, medium and high frequency bands. It can be noted that there is an increase in the mid-frequency power in standing position and in the high-frequency power in deep breathing mode. Figure 11 depicts the HRVPS plot of a typical diabetic subject in supine, standing and deep-breathing modes. Now, it can be seen that there is a decrease in mid-frequency power and in high-frequency power in deep-breathing mode compared to corresponding power levels of a normal subject (in Figure 10) [Desai, K.D et al., 2011].



AR POWER STATISTICS

Low (0.000-0.040) = 3.360 Pmax at 0.03Hz
Mid (0.040- 0.15) = 2.301 Pmax at 0.04Hz
High(0.15- 0.40) = 0.693 Pmax at 0.20Hz

Ratio (mid/high): 3.32
sympatho/vagal balance)
Power in processed sampled signal = 7.00

AR POWER STATISTICS

Low (0.000-0.040) = 1.618 Pmax at 0.02Hz
Mid (0.040- 0.15) = 1.920 Pmax at 0.10Hz
High(0.15- 0.40) = 0.343 Pmax at 0.15Hz

Ratio (mid/high): 5.60
sympatho/vagal balance)
Power in processed sampled signal = 4.00

AR POWER STATISTICS

Low (0.000-0.040) = 1.573 Pmax at 0.00Hz
Mid (0.040- 0.15) = 0.371 Pmax at 0.04Hz
High(0.15- 0.40) = 1.916 Pmax at 0.21Hz

Ratio (mid/high): 0.19
sympatho/vagal balance)
Power in processed sampled signal = 13.35

Fig. 11. HRVPS plots of a diabetic subject in supine, standing and deep breathing modes. The power statistics on the right side show the power in low, medium and high frequency bands. There is a decrease in the mid frequency power in standing position and in high frequency power in deep breathing mode compared to corresponding power levels of normal subject as shown in Fig 10a[Desai, K.D et al., 2011].

5.1 Diagnostic indices (based on HRVPS)

The analysis of HRV power spectra is commonly focused on the power in different frequency bands. In particular, the power in the high-frequency range reflects the fast parasympathetic never activity [Fallen et al., 1985], and the power in the mid-frequency range reflects both parasympathetic and sympathetic never activity [Akselrod, S., et al., 1981].

The ratio of the mid-frequency range power to the high-frequency range power is sometimes used as a relative index of the sympatho/vagal balance [Bianchi et al., 1990]. The high frequency range power ratio between supine and standing position is used as a parasympathetic index [Fallen et al., 1985]. The same ratio is used to study sympathetic function in standing position in the mid frequency range [Fallen et al., 1985]. Sympathetic vasomotor nerve function is quantified by the baro receptor oscillation frequency (i.e., the mid-peak frequency) in the HRVPS [Kamath et al., 1987].

In our study, autonomic function indices are defined in terms of spectral power indices and HRV period (or frequency) shift indices. The mid and high-frequency ranges are considered for defining indices. The diagnostic indices are based on parameters measured from the HRVPS. The values of diagnostic indices for the three groups of subjects have shown significant difference and can provide rational basis for selecting prognostic therapy before a diabetic patient develops cardiac arrhythmic complications.

The **diagnosis indices** are defined as follows:

$$I_1 = \text{Relative sympathetic} - \text{to} - \text{vagal balance index} = P_2 / P3\,, \tag{22}$$

where

(P_2)=area under HRVPS spectral plot between 0.04 Hz and 0.15 Hz

(P3)=area under HRVPS spectral plot between 0.15 and 0.4 Hz.

$$I_2 = \text{Orthostatic Stress Index} = (\text{P2sta} - \text{P2sup}) / \text{P2sup}\,, \tag{23}$$

where

(P2)=area under HRVPS spectral plot between 0.04 Hz and 0.15 Hz, and subscripts 'sta' and 'sup' refer to standing supine positions.

$$I_3 = \text{Sympatho} - \text{Vagal Integrity Index} = \sum(\text{Hrmax} - \text{Hrmin}) / n\,, \tag{24}$$

where

HRmax = Local maximum heart rate (beats per minute) during one breathing cycle.

HRmin = Local minimum heart rate (beats per minute) prior to local maximum in the same breathing cycle.

n=number of breathing cycles.

$$\begin{aligned} &I_{4std} = \text{Sympathetic HRV} - \text{Spectral Frequency Shift Index}(\text{standing}) = (F2_{std} - 0.1) / 0.1, \\ &\text{where } F_2 = \text{ Frequency of the Baroreceptor reflex peak} \end{aligned} \tag{25}$$

$$I_{5sup} = \text{Sympathetic HRV} - \text{Spectral Frequency Shift Index}(\text{supine}) = (F2_{supine} - 0.1) / 0.1 \tag{26}$$

$$I_6 = \text{Respiratory Stress Index} = \left(P_{3db} - P_{3supine} \right), \tag{27}$$

where

$$\begin{aligned} &\left(P_{2supine}\right) = \text{area under HRVPS spectral plot} \\ &\text{frequency 0.04 Hz and 0.15 Hz in supine position}\,. \end{aligned} \tag{28}$$

$$\left(P_{3db}\right) = \text{area under HRVPS spectral plot between frequency 0.15 Hz and 0.4 Hz in deep breathing}. \tag{29}$$

$$\left(P_{3supine}\right) = \text{area under HRVPS spectral plot between frequency 0.15 Hz and 0.4 Hz in supine position}. \tag{30}$$

The following Table 2 shows the calculated indices, for a sample normal subject, obtained from the HRVPS parameters.

Autonomic	Index Formula	Index Value
Relative Sympathetic-to-Vagal Balance Index	$I_1 = P_2/P_3$	$I_{1(sup)} = 2.06$ $I_{1(st)} = 7.82$ $I_{1(db)} = 0.32$
Orthostatic Stress Index	$I_2 = (P_{2sta} - P_{2sup})/P_{2sup}$	$I_2 = 1.64$
Sympatho-Vagal Integrity	$I_3 = \sum (HR_{max} - HR_{min}) / n$	$I_3 = 6.62$
HRV-Freq-Shift Index (standing)	$I_4 = (F_{2std} - 0.1) / 0.1$	$I_4 = -0.05$
HRV-Freq-Shift Index (supine)	$I_{5sup} = (F_{2supine} - 0.1) / 0.1$	$I_5 = -0.10$
Respirator Stress Index	$I_6 = (P_{3db} - P_{3supine}) / P_{3supine}$	$I_6 = 16.70$

Table 2. Computed Indices for a typical normal subject.

In this Table 2, $I_{1(sp)} = P_2/P_3$ (equation 22) in supine position; $I_{2(st)} = P_2/P_3$ (equation 22) in standing position; $I_{2(db)} = P_2/P_3$ (equation 22) in deep – breathing mode.
Now, diagnosis based on six indices makes it somewhat difficult to track in a patient as regards how much each index varies from its normal value, for making an appropriate diagnosis. So now we will adopt the novel approach, as in Ghista [Ghsta, 2004; 2009a], of formulating an index by combining the parameters in such a way that the index values are distinctly different for normal subjects, diabetics, and diabetics with ischemic heart disease. Hence, we are proposing that, from a diagnostic and classification viewpoint, it would be more convenient to formulate a DAN Integrated Index (DAN-IID) [Desai, K.D et al., 2011], as :

$$\text{DAN} - \text{IID} = \left[(\text{I1,st}) + (\text{I1,db}) + (\text{I2}) + (\text{I3}) + (\text{I6})\right] - -\left[(\text{I4}) + (\text{I5})\right] \tag{31}$$

5.2 Results and analysis: HRVPS of normal subjects, diabetic subjects, and diabetic subjects with ischemic heart disease

The instantaneous heart rate average (IHRav), average of difference between maximum and minimum heart rate over a cycle (ΔHrav), power and frequency measurements (P,F) measured from HRVPS are determined. There from the diagnostic indices are computed (as per equation 22-27).

Descriptive Statistics of Indices of the Three Groups

The computed indices for the three categories of subjects are displayed in the following Tables
Table 3 for normal subject group.
Table 4 for diabetic subject group.

Table 5 for IHF subject group.
Then, using the values in Tables (2), (3) and (4), the mean and standard deviation values of three groups are calculated and presented in the Table 6.

Name	I_{1sp}(N)	$I_{1,st}$(N)	$I_{1,db}$(N)	I_2(N)	I_3(N)	I_4(N)	I_5(N)	I_6(N)	DAN-IID
Ahamidm	2.08	14.57	0.51	3.78	3.26	-0.1	-0.07	16.26	38.55
Awmeah	2.07	7.83	0.33	1.64	6.62	-0.05	-0.1	16.7	33.27
Fahmia	1.88	2.67	0.36	2.04	8.56	0.13	-0.1	17.81	31.41
Fatimah	0.47	7.46	0.84	1.08	4.9	-0.57	0.07	0.24	15.02
Gitakr	1.93	3.93	2.49	2.66	9.51	0.03	-0.07	2.8	21.43
Indvai	1.63	8.25	0.26	2.01	3.86	-0.4	-0.07	16.14	30.85
Kploga	1.4	2.31	0.3	2.32	7.66	-0.12	0.5	10.67	22.88
Mattarh	1.21	9.17	2.02	2.2	3.59	-0.52	-0.07	1.53	19.1
Mohdsae	3.09	1.47	0.3	-0.65	4.43	-0.08	0.13	1.81	7.31
Mohsed	5.78	9.95	0.75	1.36	6.63	-0.02	0.1	21.87	40.48
Ramial	8.95	3.53	0.18	0.92	11.73	0.3	0.1	8.69	24.65
Sekarm	2.15	5.19	0.25	0.02	3.29	-0.05	0.3	9.88	18.38
Average	2.72 ±2.36	6.36 ±3.87	0.71 ±0.75	1.61 ±1.19	6.17 ±2.76	-0.12 ±0.25	6.0E-02 ±0.18	10.36 ±7.45	25.277 ±9.88

Table 3. Results of Indices for normal subject group.

Name	I_{1sp}(D)	$I_{1,st}$(D)	$I_{1,db}$(D)	I_2(D)	I_3(D)	I_4(D)	I_5(D)	I_6(D)	DAN-IID
Ahmedn	3.54	7.09	0.19	-0.06	2.83	-0.08	-0.57	1.88	12.58
Altmoh	2.07	4.95	0.31	-0.32	2.47	-0.57	-0.57	5.40	13.95
Aminaha	1.34	2.11	0.70	-0.28	2.94	-0.57	-0.57	1.08	7.69
Bakmh	4.25	2.45	0.52	-0.35	1.66	-0.57	-0.57	25.00	30.42
Elmamol	0.51	0.49	0.44	-0.41	2.25	-0.57	-0.57	-0.47	3.44
Fikria	3.57	12.72	0.16	0.78	1.51	-0.57	-0.57	16.89	33.2
Ghyarh	3.78	7.75	1.56	-0.36	2.81	-0.52	-0.57	0.20	13.05
Humoya	3.58	3.78	0.68	2.80	3.68	-0.35	-0.57	4.59	16.45
Kmilmo	2.84	5.30	0.59	0.11	2.38	-0.30	-0.57	1.15	10.4
Krshpr	0.85	0.59	0.13	-0.36	1.86	-0.57	-0.57	20.42	23.78
Kurubrl	1.55	1.74	3.08	0.03	1.44	-0.57	-0.57	3.78	11.21
Mahabs	1.54	5.78	1.06	0.51	6.91	-0.57	-0.57	1.28	16.68
Mohdosb	4.39	1.92	0.32	-0.76	2.01	0.03	0.01	1.32	4.77
Mohikat	2.41	0.55	0.29	-0.29	1.87	-0.57	-0.23	13.25	16.47
Muisdr	0.86	1.08	3.30	-0.30	2.98	-0.10	-0.57	0.34	8.07
Nasah	2.59	26.95	1.19	1.83	2.09	-0.32	0.57	0.44	32.25
Naya	0.75	3.51	3.92	-0.71	1.36	-0.57	-0.57	-0.58	8.64
Salmm	0.35	1.89	0.54	-0.30	0.78	-0.57	-0.57	-0.59	3.43
Average	2.26 ±1.36	5.03 ±6.31	1.05 ±1.17	8.66E- 02±0.9	2.43 ±1.32	0.43 ±0.2	-.455 ±0.29	5.29 ±7.94	14.804 ±9.43

Table 4. Results of Indices for diabetic subject group.

Name	$I_{1,sp}$(H)	$I_{1,st}$(H)	$I_{1,db}$(H)	I_2(H)	I_3(H)	I_4(H)	I_5(H)	I_6(H)	DAN-IID
Aminase	0.45	0.74	0.99	-0.81	1.45	-0.57	-0.57	1.74	5.25
Hamamak	6.22	1.73	0.52	-0.63	1.77	-0.57	-0.57	13.10	17.63
Mayara	2.03	3.80	1.00	0.52	2.90	-0.23	-0.08	5.53	14.06
Mdshr	1.33	1.05	0.11	0.21	3.41	-0.57	-0.05	9.62	15.02
Mohmust	2.14	2.45	1.34	0.10	2.22	-0.57	-0.28	0.69	7.63
Omarsh	2.95	2.60	0.22	-0.49	2.17	-0.20	-0.02	1.79	6.51
Shamsa	0.66	1.10	2.57	-0.16	1.78	-0.57	-0.57	0.25	6.68
Tamebr	2.79	7.10	2.09	2.50	1.58	-0.57	-0.15	9.60	23.59
Average	2.32 ±1.82	2.57 ±2.09	1.1 ±0.87	0.155 ±1.05	2.16 ±0.68	-0.48 ±0.16	-.28 ±0.24	5.29 ±4.93	12.046 ±10.85

Table 5. Results of Indices for ischemic heart disease subject group.

Index	Normal (N)		Diabetic (D)		Diabetic + IHD (H)	
	Mean	Sd	Mean	Sd	Mean	Sd
I_1 (supine)	2.719	2.357	2.266	1.356	2.32	1.821
I_1 (standing)	6.361	3.864	5.036	6.312	0.155	1.49
I_1 (deep breathing)	0.715	0.755	1.053	1.166	1.107	0.869
I_2 (orthostatic stress)	1.614	1.185	0.085	0.908	0.155	1.049
I_3 (sympatho-vagal integrity)	6.195	2.736	2.435	1.323	2.16	0.681
I_4 (sym HRVPS freq shift by standing)	-0.121	0.257	-0.439	0.203	-0.481	0.165
I_{5sup} (sym HRVPS freq shift in sup)	0.06	0.185	-0.519	0.152	-0.286	0.248
I_6 (resp stress index)	10.366	7.447	5.261	7.969	5.29	4.92
DAN-IID	25.277	9.88	14.804	9.43	12.046	6.57

Table 6. Descriptive Statistics of indices of the three groups.

Diagnostically significant indices

In order to demonstrate the effectiveness of the diagnostic indices (I_1 to I_6) to distinguish the three groups, the diagnostically significant indices are calculated using Mann Whitney Wilicoxon Rank test (Non-Parametric Tests), and the p values(<0.05)are tabulated in Table 7 below[Desai, K.D et al., 2011].

Index	Significance Between Two Groups	P-value(<0.05)
I_1 (standing)	N & H	0.0109
I_2	N & H	0.0253
I_3	N & H	0.0004
I_4	N & H	0.0025
I_5	N & H	0.0083
I_2	N & D	0.0020
I_3	N & D	0.0000
I_4	N & D	0.0015
I_5	N & D	0.0000
I_6	N & D	0.0422
I_5	H & D	0.0105

Table 7. Diagnostically Significant Indices.

5.3 Physiological relevance of the computed indices

The computed indices reflect the sympatho-vagal interactions that modulate cardiovascular function. The low-frequency component (in the 0.04Hz to 0.15Hz range) of the HRV power spectrum (F_2 peak) is an indicator of sympathetic modulation, and the high frequency component (in the 0.15Hz to 0.4Hz range) in the HRV power spectrum (F_3 peak) is a marker of vagal modulation.

The index I_1 (= P_2/P_3) represents relative sympathetic-to-vagal balance, I_1 is found to be reduced to a very low value, from 6.361 to 0.155 in standing position in the case of diabetics with ischemic heart disease. This indicates that diabetics with ischemic heart disease are not able to withstand orthostatic stress or load. Patients recovering from an acute myocardial infarction can be expected to have an increased I1 index during early convalescence, and a return to a normal value by 6 to 12 months

The orthostatic stress index I_2 shows significant reduction from a normal value of 1.614 to 0.085 in diabetics, and, to 0.155 in diabetics with ischemic heart disease. A similar trend is noted for **the sympatho-vagal integrity index I_3,** showing reduction in the index value from a normal value of 6.19 to 2.43 in the case of diabetics, and to 2.16 in diabetics with ischemic heart disease. This is indicative of damage to the sympathetic and parasympathetic systems controlling the SA node pacing activity

The sympathetic HRVRS frequency-shift Index in standing position (I_{4sd}) and Sympathetic HRVPS frequency-shift Index in supine position (I_{5sup}) are found to be decreased in diabetics as well as in diabetics with ischemic heart disease patients, compared to the normal subject group. This is indicative of the increased delay (of more than 10 seconds) in case of diabetics as well as diabetics with ischemic heart disease, due to demyelination of their nervous control system controlling the heart rate.

The Respiratory Stress Index I_6 denotes the effectiveness of vagal control on heart rate variation, and is found to be considerably reduced from a normal value of 10.36 to 5.26 in diabetics, and to 5.29 in diabetics with ischemic heart disease.

Thus the indices derived from the HRV power spectrum represent non-invasive signatures of the balance between sympathetic and parasympathetic components of the autonomic nervous system. These indices are shown to characterize diabetic autonomic neuropathy state, and to hence distinguish diabetics and diabetics with ischemic heart disease.

Integrated index composed of power-spectral indices

We have shown how well the HRVPS indices differentiate normal subjects from diabetics and diabetics with ischemic heart disease.

We now compute the values of this integrated Index (DAN-IID) for normal subjects (in Table 3), diabetic subjects (in Table 4), and diabetic patients with ischemic heart disease (in Table 5) . From these Index values, we compute its mean values and standard deviations, for normals, diabetics, and diabetics with ischemic heart disease (IHD). These values are tabulated in Table 6. It can be clearly seen, from this Table 6, that our integrated Index can be employed to effectively differentiate and diagnose diabetic subjects and diabetics with IHD. The Index can also be employed to assess the efficacy of diabetic medication and insulin administration.

We next make a distribution plot of this Integrated Index for normals, diabetics, and diabetics with IHD, in Figure 12. This plot graphically illustrates how well this integrated Index separates normal subjects, diabetic patients, and diabetic patients with ischemic heart disease[Desai, K.D et al., 2011].

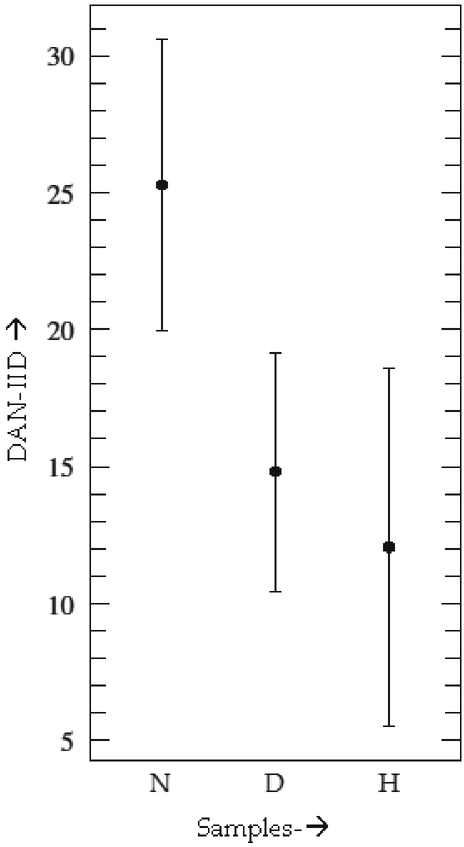


Fig. 12. Variation of DAN-IID for (N) normal subjects, (D) diabetic patients, and (H) diabetics with IHD. It can be noted that this DAN-IID clearly separates diabetics and diabetics with IHD from normal subjects.

6. Activity-based dynamic insulin infusion system

In section 3, we have introduced the glucose-insulin regulatory system and applied it to model the OGTT. We came up with our novel DBI, by means of which we can even detect supposedly normal subjects who are at risk of becoming diabetic. Now, we continue on the trail of this glucose-insulin regulatory system, by presenting its application to illustrate how for a diabetic patient the glucose level keeps going up after meal, and how it is regulated by automated infusion of insulin.

Herein, we demonstrate the operation of a Glucose activity-based Dynamic Insulin infusion (or release) system. The current insulin infusion systems are based on the diabetic patient's known activities history, in order to estimate the required insulin amount. These techniques do not allow the patients to deviate too much from their normal daily activities [Naylor et al., 1996]. Hence, our approach focuses on regular sampling of the diabetic patients' blood glucose concentration through a sensor, to compute the required amount of insulin to be released into the blood stream.

The amount of insulin infused to bring the blood glucose concentration down is regulated by a Closed-loop PD (Proportional-Derivative) Control system algorithm (Fig. 13). The closed loop system continuously monitors the blood glucose concentration at 0.5 h interval. Once the system detects that the blood glucose concentration exceeds a predetermined threshold e.g. 120mg/dl [International Diabetes Federation], the system is alarmed and 'calculates' the amount of insulin required [Loh, 2004] to bring the blood glucose concentration below the threshold.

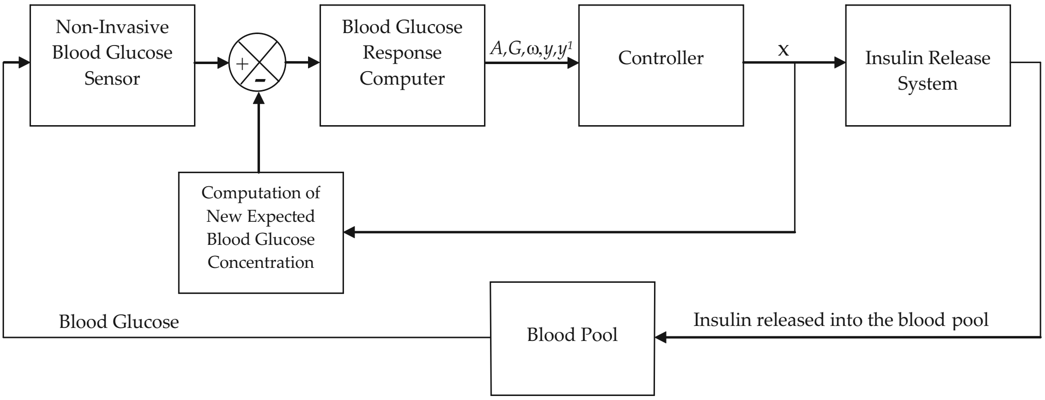

Non-Invasive Blood Glucose Sensor
+
−
Blood Glucose Response Computer
A, G, ω, y, y^1
Controller
x
Insulin Release System
Computation of New Expected Blood Glucose Concentration
Blood Glucose
Blood Pool
Insulin released into the blood pool

Fig. 13. Block diagram of the Glucose Regulating Insulin Release (GRIR) system: The glucose sensor monitors the existing blood glucose level. The error between the glucose sensor level and the computed expected blood glucose concentration is fed into the Closed-loop PD (Proportional-Derivative) Control system, and its algorithm computes the amount of insulin (x) to be released. Accordingly, the required amount of insulin is released into the blood. This now readjusts the blood glucose level, which is again monitored by the sensor.

Then, Fig. 14 shows the results of the application of the Insulin Infusion Release system of Fig 13. The diabetic subject D18's unaided glucose clinical data is fed into the system. On the Y axis, we have plotted blood-glucose concentration above the patient's glucose concentration of 120 mg/dl (or 1.2 g/l) at time 0 after meal. The insulin is released at 0.5 hour, 1 hour and 1.5 hours after meal. In figure 16, it is seen, how following insulin infusion, the blood glucose comes down. Once the blood glucose concentration drops below the threshold, the controller will stop releasing insulin into the blood stream.

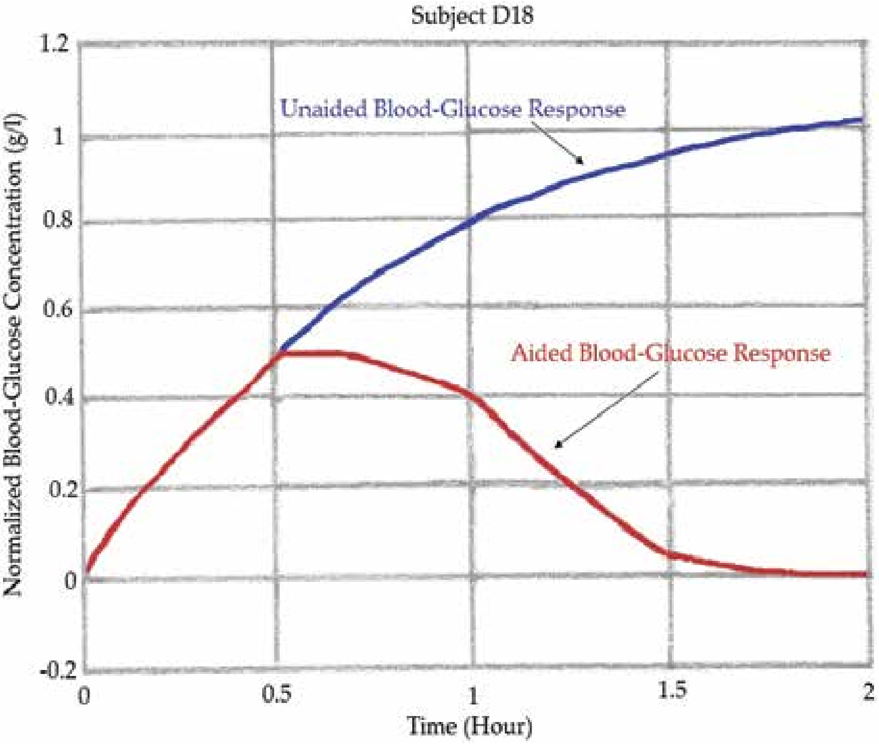


Fig. 14. The subject's unaided blood glucose concentration at time 0 is above 120mg/dl. The system is alarmed and samples the blood glucose concentration at 0.5h (170 mg/dl). The system sends a bolus of insulin 10mU/dl into the blood stream. The system keeps monitoring the resulting blood glucose concentration at 1.0h and 1.5 hour intervals, and infuses computed insulin bolus into the blood stream to bring the blood glucose concentration below the threshold value.

Thus, we have demonstrated the capability of the activity based adaptive dynamic real-time insulin release system. This system is able to protect the users from hyperglycemia.

7. Conclusion

This chapter is framed to provide useful insights into: (i) the mechanisms of diabetes; (ii) how the bioengineering analysis of the glucose regulatory system can be employed to diagnose diabetic patients and subjects at risk of becoming diabetic, based on an integrated index composed of parameters of the governing differential equation to simulate blood glucose concentration data of OGTT; (iii) parameters of time-and frequency-domain measures of HRV can be employed to differentiate diabetic subjects from normal subjects; (iii) processing of retinal digital fundus images to characterize retinopathy, and analysis of plantar pressure distribution images of normal subjects, and subjects with diabetes type II without and with neuropathy, (iv) diagnosis of diabetic autonomic neuropathy by means of a novel intregrated index composed of parameters of heartrate variability power-spectrum plots; (v) how we can apply the glucose-insulin regulatory system to illustrate how for a diabetic patient the glucose level keeps going up after meal, and how it is can be regulated by automated infusion of insulin.

8. Acknowledgments

Authors thank Journal of Mechanics in Biology and Medicine (published by World Scintific Publishing Co. Pte Ltd) for giving permission reproduce the portions of materials from their published papers, (i) U Rajendra Acharya, Jasper Tong, Vinitha Sree, Chua Kuang Chua, Tan Peck Ha, Dhanjoo N Ghista, Subhagata Chattopadhyay, Kwan-Hoong Ng, Jasjit S Suri, "Computer-Based Identification of Type 2 Diabetic Subjects With and Without Neuropathy using Dynamic Planter Pressure and Principal Component Analysis", Journal of Medical Systems, 2011 (In Press: DOI: 10.1007/s10916-011-9715-0); (ii) Desai, K.D, Dhanjoo N. Ghista, Issam Jaha El Mugamex, U Rajendra Acharya, Michael Towsey, Sultan Abdul Ali, Mohammed Saeed, M. Amin Fikri, *"Diabetic Autonomic Neuropathy Detection By Heart-Rate Variability Power-Spectral Analysis"*, Journal of Mechanics in Medicine and Biology, 2011 (In Press); (iii) Dhanjoo N Ghista, "Nondimensional Physiological Indices for Medical Assessment", Journal of Mechanics in Medicine and Biology, 9(4), 2009, 643-669.

9. References

Acharya, U.R.; Lim, C.M.; Ng, E.Y.K.; Chee, C. & Tamura, T. (2009). Computer-based detection of diabetes retinopathy stages using digital fundus images. *Proc Instn Mech Engrs. J Eng Med Part H,* Volume 223, No. 5, pp. 545-553.

Acharya, U.R.; Tan, P.H.; Subramaniam, T.; Tamura, T.; Chua, K.C.; Goh, S.C.E.; Lim, C.M.; Goh, S.Y.D.; Chung, K.R.C.& Law, C.(2008). Automated identification of diabetic type 2 subjects with and without neuropathy using wavelet transform on pedobarograph. *J Med Syst*, Volume 32, No. 1, pp.21-29.

Acharya, U.R.; Joseph, P.K.; Kannathal, N.; Lim, C.M. & Suri, J.S.(2006). Heart rate variability: A Review, *Med Biol Eng Comput*, Volume 44, No. 12, pp. 1031-1051.

Acharya, U.R.; Suri, J.S.; Spaan, J.A.E. & Krishnan, S.M.(2007). Advances in Cardiac Signal Processing. Springer Verlang GmbH. Berlin, Heidelberg.

Acharya, U.R.; Rahman, M.A.; Aziz, Z.; Tan, P.H.; Ng, E.Y.K. & Yu, W. (2008). Computer-based identification of plantar pressure in type 2 diabetes subjects with and without neuropathy. *Journal of Mechanics in Medicine and Biology,* Volume 8, No.3, pp. 363-375.

Acharya, U.R.; Faust, O.; Dua, S.; Hong, S.J.; Yang, T.S.; Lai, P.S. & Choo, K. (2011a). Computer-based diagnosis of diabetes retinopathy stages using digital fundus images. Dua S, Acharya UR, Ng EYK (Eds.), 'Computational image modeling of human eye with applications', *World Scientific Publishing Company.*

Acharya, U.R.; Ghista, D.N.; Nergui, M.; Chattopadhyay, S.; Ng, E.Y.K.; Sree, V.S.; Tong, J.W.K.; Tan, J.H.; Loh, K.M.& Suri, J.S. (2011b). Diabetes Mellitus: Enquiry into its Medical aspects and Bioengineering of its Monitoring and Regulation. *Journal of Mechanics in Medicine and Biology* (In Press).

Akselrod, S., et al (1981), Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control, Science, Vol. 213, 10, pp. 220-222 .

Baynes, J.W. (1991). Role of oxidase stress in development of complications in diabetes (Review). *Diabetes*, Vol. 40, No. 4, pp. 405-412.

Bianchi, A., et al (1990) Spectral analysis of heart rate variability signal and respiration in diabetic subjects, Medical & Biological Engineering & Computing, Vol. 28, pp. 205-211.

Bonnefont-Rousselot, D. (2002). Glucose and reactive oxygen species. *Current Opinion in Clinical Nutrition and Metabolic Care*, Vol. 5, No. 5, pp. 561-568.

Cavanagh, P.R.; Sims, D.S.& Sanders, Jr. L.J. (1991). Body mass is a poor predictor of peak plantar pressure in diabetic men. *Diabetes Care*, Volume 14, pp.750-755.

Dangel, S.; Meier, P.F.; Moser, H.R.; Plibersek, S. & Shen, Y.(1999). Time series analysis of sleep EEG. *Computer Assisted Physics,* pp. 93-95.

Desai, K. D.; Ghista, D.N.; Mugamex, I. J. El.; Acharya, U.R.; Towsey, M.; Ali, S.A.; Saeed, M., Fikri, M. A. (2011). Diabetic Autonomic Neuropathy Detection By Heart-Rate Variability Power-Spectral Analysis*, Journal of Mechanics in Medicine and Biology,* 2011 (In Press);

Dittakavi, S.S. & Ghista, D.N.(2001). Glucose tolerance tests modeling & patient simulation for diagnosis. *J Mech Med Biol,* Volume 1, No.2, pp.193-223.

Eckmann, J.P.; Kamphorst, S.O. & Ruelle, D.(1987). Recurrence plots of dynamical systems. *Europhys Lett,*Volume 4, pp. 973-977.

Evans, J.L.; Golfine, I.D.; Maddux, B.A. & Grodsky, G.M. (2002). Oxidative stress and stress-activated signalling pathways: a unifying hypothesis of type 2 diabetes. *Endocrine Reviews*, Vol. 23, No. 5, pp. 599-622.

Fallen, E.L., Nandogopal, D., Connonlly, S., and Ghista, D.N. (1987) "How reproducible is the power spectrum of heart rate variability in health subjects?", Proceedings of International Symposium on Neural and Cardiovascular Mechanisms, Bologna, Italy May, 1985.

Fantus, G. (2011). Diabetes, glucose toxicity. http://www.endotext.org/diabetes/diabetes12new/diabetes12.htm

Faust, O.; Acharya, U.R.; Molinari, F.; Chattopadhyay, S. & Tamura, T. (2011). Linear and Non-Linear Analysis of Cardiac Health in Diabetic Subjects. *Biomedical Signal Processing and Control*, 2011

Ghista, D.N.(2004). Physiological systems' numbers in medical diagnosis and hospital cost effective operation. *J Mech Med Biol.*, Volume 4, No. 4,pp. 401–418.

Ghista, D.N.(2009a). Nondimensional physiological indices for medical assessment. *J Mech Med Biol.,* Volume 9, No. 4, pp. 643-669.

Ghista, D.N. (2009b). Applied Biomedical Engineering Mechanics. *CRC Press.*

Gleason, C.E.; Gonzalez, M.; Harmon, J.S. & Robertson, R.P. (2000). Determinants of glucose toxicity and its reversibility in the pancreatic islet beta-cell line, HIT-T15. *American Journal of Physiology, Endocrinology and Metabolism*, Vol. 279, No. 5, pp. E997-E1002.

Grassberger, P. & Procassia, I.(1983). Measuring the strangeness of strange attractors. *Physica,* Volume D9, pp.189-208.

Haber, C.A.; Lam, T.K.; Yu, Z.; Gupta, N.; Goh, T.; Bogdanovic, E.; Giacca, A. & Fantus, I.G. (2003). N-acetylcysteine and taurine prevent hyperglycaemia-induced insulin resistance *in vivo*: possible role of oxidative stress. *American Journal of Physiology: Endocrinology and Metabolism*, Vol. 285, No. 4, pp. E744-E753.

Hayden, M.R. & Tyagi, S.C. (2002). Islet redox stress: the manifold toxicities of of insulin resistance, metabolic syndrome and amylin derived islet amyloid in type 2 diabetes mellitus. *Journal of Pharmacology*, Vol. 3, No. 4, pp. 86-108.

International Diabetes Federation. Website: http://www.idf.org/2000, last accessed in Dec 2010.

Kamath, M.V.; Ghista, D.N.; Fallen, E.L.; Fitchett, D.; Miller, D. & McKelvie, R.(1987). Heart rate variability power spectrogram as potential noninvasive signature of cardiac regulatory system response, mechanisms, and disorders. *Heart Vessels*, Volume 3, pp.33-41.

Krall, L.P. & Beaser, R.S. (1989). *Joslin Diabetes Manual* (12th edition), Lippincott Williams and Wilkins, ISBN 978012111200, London (United Kingdom).

Loh, K.C. (2004). Pharmacology of Oral Anti-hyperglycaemic Agents & Insulin (Invited Article). *Singapore Family Physician*, Volume 30, pp.16-20.

Maiese, K.; Chong, Z.Z. & Shang, Y.C. (2007). Mechanistic insights into diabetes mellitus and oxidative stress. *Current Medicinal Chemistry*, Vol. 16, No. 16, pp. 1729-1738.

Mandelbrot, B.B.(1983). Geometry of Nature. Freeman San Francisco.

Nayak, J.; Bhat, P.S.; Acharya, U.R.; Lim, C.M. & Gupta, M.(2008). Automated identification of different stages of diabetic retinopathy using digital fundus images. *J Med Syst*, Volume 32, No. 2, pp. 107-115.

Naylor, C.D.; Sermer, M.; Chen, E. & Sykora, K.(1996). Cesarean delivery in relation to birth weight and gestational glucose tolerance: pathophysiology or practice style? Toronto Tri-Hospital Gestational Diabetes Investigators. *JAMA,* Volume 275, pp.1165.

Petersen, K.F.; Befroy, D.; Dufour, S.; Dziura, J.; Ariyan, C.; Rothman, D.L.; DiPietro, L.; Cline, G.W. & Shulman, G.I. (2003). Mitochondrial dysfunction in the elderly: possible role in insulin resistance. *Science*, Vol. 300, No. 5622, pp. 1140-1142.

Pincus, S.M.(1991). Approximate entropy as a measure of system complexity. *Proc National Academic Science,* Volume 88, pp.2297-2301.

Rahman, M.A.; Aziz, Z.; Acharya, U.R.; Tan, P.H.; Natarajan, K.; Ng, E.Y.K.; Law, C.; Subramaniam, T. & Shuen, W.Y.(2006). Analysis of plantar pressure in diabetic Type 2 subjects with and without neuropathy. *Innov Technol Biol Med*, Volume 27, No. 2, pp.46-55.

Robertson, R.P. (2004). Chronic oxidative stress as central mechanism for glucose toxicity in pancreatic islet beta cells in diabetes. *Journal of Biological Chemistry*, Vol. 279, No. 41, pp. 42351-42354.

Robertson, R.P.; Harmon, J.; Tran, P.O.; Tanaka, Y. & Takahashi, H. (2003). Glucose toxicity in β-cells: Type 2 diabetes, good radicals gone bad, and glutathione connection. *Diabetes*, Vol. 52, No. 3, pp. 581-587.

Rosenstien, M.; Colins, J.J. & De Luca, C.J.(1993). A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D*, Volume 65, pp. 117-134.

Tanaka, Y.; Tran, P.O.; Harmon, J. & Robertson, R.P. (2002). A role for glutathione peroxidase in protecting pancreatic beta cells against oxidative stress in a model of glucose toxicity. *Proceedings of National Academy of Science (USA)*, Vol. 99, No.19, pp. 12363-12368.

Task Force of the European Society of Cardiology and North American Society of Pacing and electrophysiology.(1996). Heart Rate Variability: Standards of measurement, physiological interpretation and clinical use. *Eur Heart J,* Volume 17, pp.354-381.

Van der Akker, T.J.; Koeleman, A.S.M.; Hogenhuis, L.A. & Rompelman, G.(1983). Heart-rate variability and blood pressure oscillations in diabetics with autonomic neuropathy. *Automedica*, Volume 4, pp.201-208.

Yoshida, K.; Hirokawa, J.; Tagami, S.; Kawakami, Y.; Urata, Y. & Kondo, T. (1995). Weakened cellular scavenging activity against oxidative stress in diabetes mellitus: regulation of glutathione synthesis and efflux. *Diabetologia*, Vol. 38, No. 2, pp. 201-210.

Yun, W.L.; Acharya, U.R.; Venkatesh, Y.V.; Chee, C., Lim C.M. & Ng, E.Y.K.(2008). Identification of different stages of diabetic retinopathy using retinal optical images. *Information Sciences*, Volume 178, No. 1, pp.106-121.

18

Domain-Specific Software Engineering Design for Diabetes Mellitus Study Through Gene and Retinopathy Analysis

Hua Cao, Deyin Lu and Bahram Khoobehi
Louisiana State University, University of Mississippi Medical Center, LSU Eye Center, USA

1. Introduction

Software engineering designs and practices differ widely among various application domains. This chapter is concentrating on high performance software engineering design for bioinformatics and more specifically for diabetes mellitus study through gene and retinopathy analysis. Complex gene interaction study offers an effective control of blood glucose, blood pressure and lipids. Early detection of retinopathy is effective in minimizing the risk of irreversible vision loss and other long-term consequence associated with diabetes mellitus.

Type 2 diabetes mellitus is a disorder of glucose homeostasis involving complex gene and environmental interactions that are incompletely understood. Mammalian homologs of nematode sex determination genes have recently been implicated in glucose homeostasis and type 2 diabetes mellitus. The Fem1b knockout (Fem1b-KO) mice have been developed, with targeted inactivation of Fem1b, a homolog of the nematode fem-1 sex determination gene. It shows that the Fem1b-KO mice display abnormal glucose tolerance and that this is due predominantly to defective glucose-stimulated insulin secretion. Arginine-stimulated insulin secretion is also affected. These data implicate Fem1b in pancreatic islet function and insulin secretion, strengthening evidence that a genetic pathway homologous to nematode sex determination may be involved in glucose homeostasis and suggesting novel genes and processes as potential candidates in the pathogenesis of diabetes mellitus. In addition, this chapter is going to introduce basic gene analysis approaches that can be applied on diabetes mellitus study. These approaches include searching Genbank online database using BLAST, mapping DNA, locating genes, aligning different DNA or protein sequences, determining genotypes, and comparing nucleotide or amino acid sequences using global and local alignment algorithms. Fem1b gene, as an example, is going to be discussed with these basic gene analysis approaches.

Diabetic retinopathy is the leading cause of new cases of blindness among Americans aged 20 to 64 in both predominantly white and black populations [1]. Despite the recommendation for yearly eye examinations and efforts to achieve this, of the approximately 17 million Americans with diabetes, about 6 million nationwide remain undiagnosed and untreated, or not receiving annual eye examinations, which can lead to diabetic retinopathy [2].

Early indications of retinal blood vessel abnormalities and complications provide important indicators for clinical timely diagnosis and treatment of diabetes mellitus and eye disorders. The software engineering design tool facilitates increasing the number of annual diabetic screening eye examinations, thereby reducing the long time wait for diabetic patients to receive eye examinations. Common activities in software engineering approach for retinopathy include single or multi-modality retinal image registration, fusion, vessel pattern recognition, arteries & veins identification, and vessel diameter measurement. These methods play a major role in the development of better methods of diagnosing and treating diabetic retinopathy. Fusing the multi-modality retinal images, which usually requires intensive computational effort, is a very challenging problem because of the possible vast content change and non-uniform distributed intensities of the involved images.

This chapter is going to present a novel approach of retinal image fusion. Control points are detected at the vessel bifurcations using adaptive exploratory algorithm. Mutual-Pixel-Count (MPC) maximization based heuristic optimization adjusts the control points at the sub-pixel level. The iteration stops either when MPC reaches the maximum value, or when the maximum allowable loop count is reached. A refinement of the parameter set is obtained at the end of each loop, and finally an optimal fused image is generated at the end of the iteration. By locking the multi-modality retinal images into one single volume, the algorithm allows ophthalmologists to match the same eye over time to get a sense of disease progress and pinpoint surgical tools to increase accuracy and speed of the surgery. The new algorithm can be easily expanded to human or animals' 3D eye, brain, or body image registration and fusion.

2. Diabetes mellitus type 2 and fem1b gene

2.1 Diabetes mellitus type 2 occurrence and its diagnosis

Diabetes mellitus type 2 is a metabolic disorder that is characterized by high blood glucose in the context of insulin resistance and relative insulin deficiency [3]. The pathophysiology of Type 2 diabetes mellitus involves impaired insulin secretion, and impaired insulin action in regulating glucose and fatty acid metabolism in the liver, skeletal muscle, and adipose tissue. Many individuals with Type 2 diabetes mellitus have hypertension and perturbations of lipoprotein metabolism, as well as other manifestations of the insulin resistant syndrome. In addition to the risk for development of diabetes - specific complications of retinopathy, Type 2 diabetes mellitus is recognized as a substantial risk factor for cardiovascular disease [4].

It is recommended by the National Diabetes Data Group that diagnosis of diabetes mellitus be based on [5]

1. Two fasting plasma glucose levels of 126 mg/dL (7.0 mmol/L) or higher;
2. Two two-hour postprandial plasma glucose (2hrPPG) readings of 200 mg/dL (11.1 mmol/L) or higher after a glucose load of 75 g;
3. Two casual glucose readings of 200 mg per dL (11.1 mmol per L) or higher.

Fasting plasma glucose was selected as the primary diagnostic test because it predicts adverse outcomes (e.g., retinopathy) and is easy to perform in a clinical setting.

A mammalian Fem1 gene family, encoding homologs of fem-1, has been characterized and consists of at least three members in the mouse, designated Fem1a, Fem1b, and Fem1c; these have highly conserved homologs in humans, designated FEM1A, FEM1B, and FEM1C, respectively. Mammalian homologs of two other nematode sex determination genes, tra-2 and tra-3, have recently been implicated in glucose homeostasis and type 2 diabetes mellitus. In producing susceptibility to type 2 diabetes mellitus, NIDDM1 is known to interact with a gene, whose identity is unknown, on human chromosome 15 near the CYP19 locus at 15q21.3 [6]. This is near 15q22, where FEM1B, the human homolog of mouse Fem1b, localizes [7].

2.2 Glucose and insulin measurements

Blood glucose from tail vein was measured using an OneTouch FastTake Glucometer. Insulin was measured from plasma, tissue extracts, or cell supernatants using the Rat (Mouse) Sensitive Insulin radioimmunoassay (RIA) kit and the manufacturer's instructions. For the intraperitoneal glucose tolerance test (iPGTT), intraperitoneal insulin tolerance test (iP-ITT), and acute-phase glucosestimulated insulin secretion (A-GSIS) test, there were 12 animals in each group (12 male homozygous Fem1b-KO, 12 male wild-type controls, 12 female homozygous Fem1b-KO, and 12 female wild-type controls), aged 3 to 4 months. The arginine-stimulated insulin secretion test compared eight Fem1b-KO homozygous males with eight wild-type males, aged 6 months. D-Glucose (200 mg/ml) was administered at 2 mg/g body weight by intraperitoneal injection [8]. Tail vein blood was sampled for blood glucose determination from nonsedated animals before and at 15, 30, 60, and 120 min after glucose administration.

2.3 Fem1b-KO mice development

In this study, the gene targeting by homologous recombination has been used to generate Fem1b knockout (Fem1b-KO) mice with inactivation of the Fem1b gene. It was performed with a deletion of Fem1b coding exon 1, which contains the translation initiation codon and the first two ankyrin repeats [9]. The results show that these mice display abnormal glucose homeostasis, with abnormal glucose tolerance tests and defective glucose-stimulated insulin secretion. These findings indicate that Fem1b is involved in pancreatic islet β-cell function and provide further evidence for involvement of a pathway resembling nematode sex determination in mammalian glucose homeostasis. This approach utilized standard methodology (Figure 1) and the basic elements of the targeting vector and screening strategy by Southern blot and PCR genotyping. Figure 2 was generated by Zeiss AvioVision with the immunohistochemical analysis demonstrates a loss of specific Fem1b staining in islets of Fem1b-KO homozygotes.

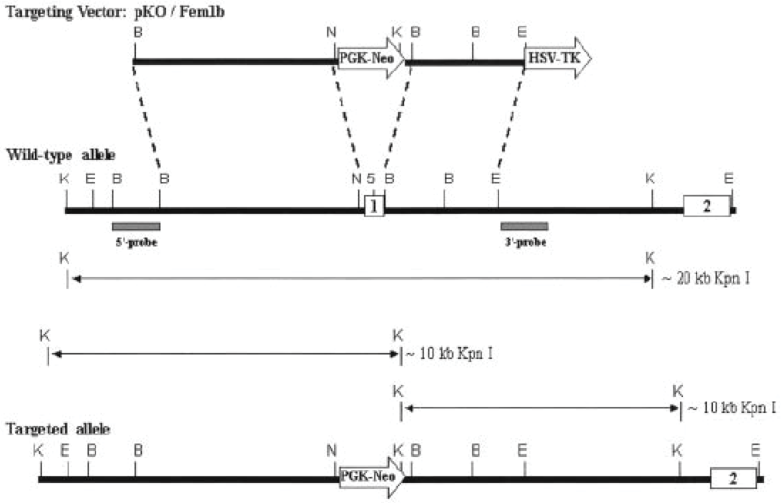


Fig. 1. General strategy, with expected KpnI digestion products along with 5_ and 3_ probes for Southern blot. The boxes labeled 1 and 2 represent exons 1 and 2, respectively. Exon 1 is replaced by the PGK-Neo gene (labeled arrow) in the targeted allele.

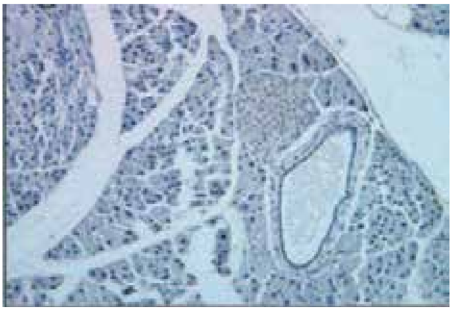


Fig. 2. Immunostaining of homozygous Fem1b-KO pancreas with anti-Fem1b C-terminus antibody Li-51, demonstrating the absence of specific staining for Fem1b.

2.4 Glucose homeostasis in Fem1b-KO mice

As noted above, mammalian homologues of nematode sex determination genes have recently been shown to be involved in glucose homeostasis and type 2 diabetes mellitus. Based on this logic, glucose homeostasis was evaluated in the Fem1b-KO mice by using established experimental methods. As a first-line screen, these mice were between 3 and 4 months of age. The iP-ITT showed minimally abnormal results (Figure 3), suggesting that insulin resistance is not the primary defect in homozygotes, although it could be contributing.

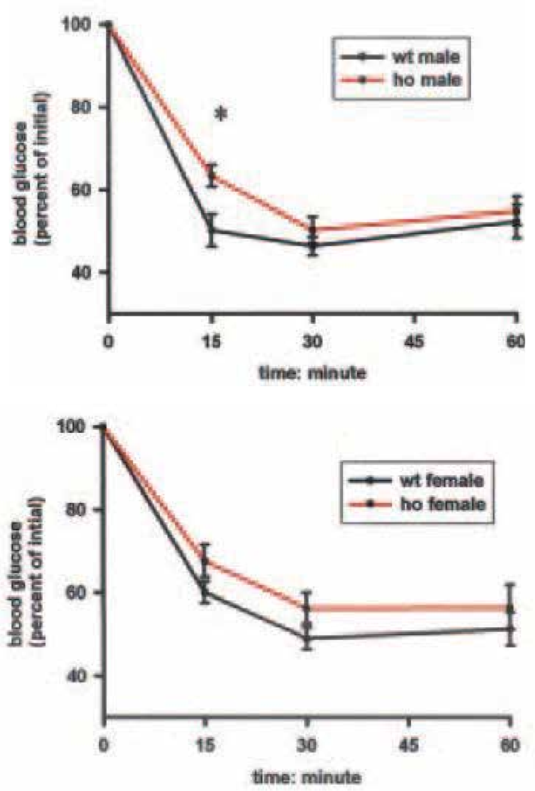


Fig. 3. Intraperitoneal insulin tolerance test (males and females).

To evaluate whether the defective acute-phase insulin secretion is related to a defect in secretion per se as opposed to a defect in insulin production, the insulin content was measured o in these mice (Figure 4), which demonstrates that Fem1b-KO homozygotes have increased insulin content compared to that of wild-type controls.

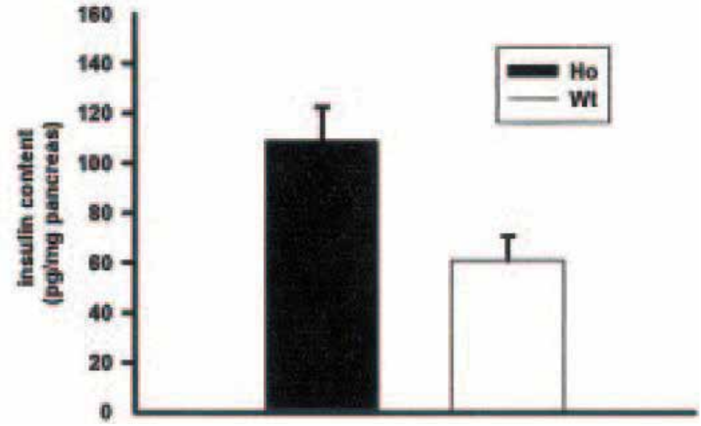


Fig. 4. Pancreatic insulin content in fasted mice (four Fem1b-KO homozygous mice and four wild-type mice).

2.5 Immunostaining of Fem1b in pancreatic islets

In humans, FEM1B has been shown to be expressed within whole pancreas [10], but cell type distribution within this organ was unknown. Immunostaining of wild-type pancreas with immunoaffinity-purified antibody shows that Fem1b protein is expressed in pancreatic islets (Figure 5).

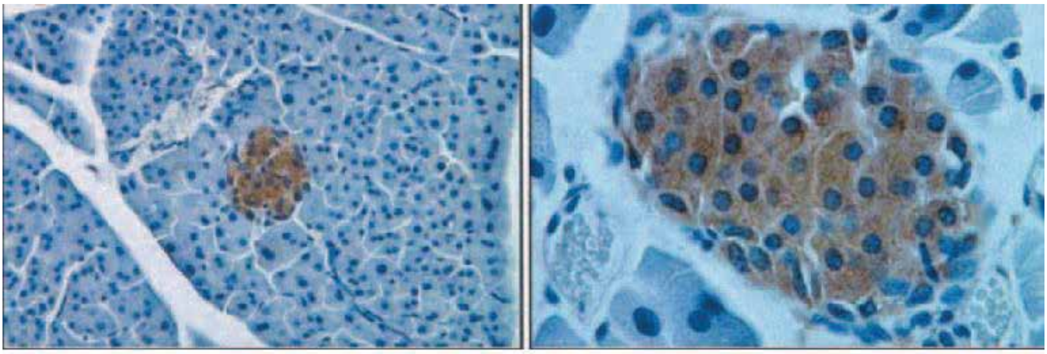


Fig. 5. Lower and higher magnification, respectively, of immunoperoxidase staining with antibody Li-51 (against the C terminus of Fem1b) counterstained with hematoxylin.

Immunostaining of the pancreas with a commercially available goat polyclonal antibody against Fem1b demonstrates the same islet staining pattern, with an absence of specific staining in the Fem1b-KO homozygotes. Coimmunostaining with antibodies to insulin, a β-cell marker, demonstrates that Fem1b is expressed in virtually all β cells (Figure 6).

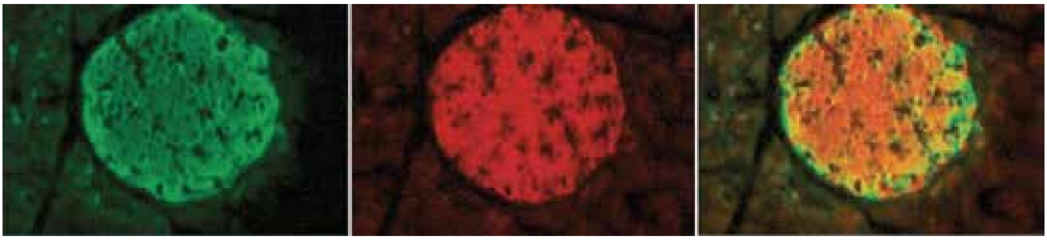


Fig. 6. Immunofluorescence staining of insulin (green), Fem1b (red), and merged image demonstrating that Fem1b is expressed not only in insulin-positive β cells but also in insulin-negative non-β cells.

The coimmunostaining with antibodies to glucagon and somatostatin, markers for α cells and δ cells, respectively, demonstrates that the Fem1b protein is also expressed in these non-β cells (Figure 7).

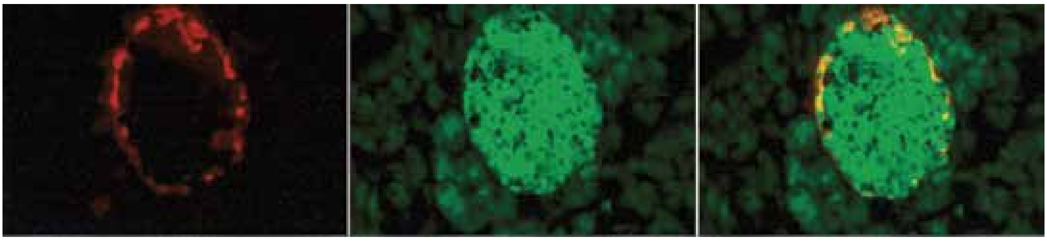


Fig. 7. Immunofluorescence staining with a combination of antibodies to glucagon and somatostatin (red) and Fem1b (green) and a merged image verifying expression of Fem1b within non- β cells.

3. Fem1b gene search and alignment

3.1 BLAST search

Basic Local Alignment Search Tool (BLAST) is an algorithm for comparing amino-acid sequences or the nucleotides. By performing a BLAST search, one is able to compare an unknown sequence with a library or database of known sequences, and identify library sequences that resemble the unknown sequence above a certain score percentage [11] (usually 40%). This chapter is going to give an example that follows the discovery of a previously unknown fem1b gene in the mouse and performs a BLAST search of the human genome to see if humans carry a similar fem1b gene. BLAST identifies sequences in the human fem1b genome that resemble the mouse fem1b gene based on similarity of sequence. Given a sequence of one fragment of mouse gene (Figure 8), BLAST software is going to search all human gene banks and find similar genes. http://blast.ncbi.nlm.nih.gov/Blast.cgi is the official BLAST website.

```
CAGAAGTCGG ATTGAAGCCT TGGAGCTCTT GGGTGCCTCC TTTGCAAATG ATCGTGAGAA
CTATGACATC ATGAAGACAT ACCACTATTT ATATTTAGCT ATGTTGGAGA GATTTCAGGA
TGGTGACAAC ATTCTTGAAA AAGAGGTTCT CCCACCCATC CATGCTTATG GCAACAGAAC
TGAGTGTAGG AACCCACAGG AATTGGAGGC TATTCGGCAA GACAGAGATG CTCTTCACAT
GGAGGGCCTT ATAGTTCGGG AACGGATTTT AGGTGCTGAC AACATTGATG TTTCCCACCC
CATCATTTAC AGAGGGGCTG TCTATGCTGA TAACATGGAG TTCGAGCAGT GCATCAAATT
```

Fig. 8. Unknown mouse gene.

When the results page appears, click the identifier with the highest score and you will see the following information. Here the highest score is 481. The score was calculated on the match quality and the length of the most-similar segments that occur between the unknown mouse gene and the target human fem1b gene.

Human - Sore: 481 Query Coverage 98%

```
LOCUS       NM_015322               2583 bp    mRNA    linear   PRI 26-DEC-2010
DEFINITION  Homo sapiens fem-1 homolog b (C. elegans) (FEM1B), mRNA.
ACCESSION   NM_015322
VERSION     NM_015322.3  GI:52851431
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2583)
  AUTHORS   Ewens,K.G., Stewart,D.R., Ankener,W., Urbanek,M., McAllister,J.M.,
            Chen,C., Baig,K.M., Parker,S.C., Margulies,E.H., Legro,R.S.,
            Dunaif,A., Strauss,J.F. III and Spielman,R.S.
  TITLE     Family-based analysis of candidate genes for polycystic ovary
            syndrome
  JOURNAL   J. Clin. Endocrinol. Metab. 95 (5), 2306-2315 (2010)
   PUBMED   20200332
  REMARK    GeneRIF: Observational study of gene-disease association. (HuGE
            Navigator)
REFERENCE   2  (bases 1 to 2583)
  AUTHORS   Subauste,M.C., Sansom,O.J., Porecha,N., Raich,N., Du,L. and
            Maher,J.F.
  TITLE     Fem1b, a proapoptotic protein, mediates proteasome
            inhibitor-induced apoptosis of human colon cancer cells
  JOURNAL   Mol. Carcinog. 49 (2), 105-113 (2010)
   PUBMED   19908242
```

When you scroll down the page, you see reach a long list of the human fem1b nucleotide sequence starting with

```
ATGGAAGGACTTGCGGGGTATGTTTACAAAGCAGCGAGCGAGGGTAAAGTGCTGACCCTGG
CTGCCCTATTATTAAATCGCTCGGAATCCGATATACGATACCTTCTTGGGTACGTTAGTCA
GCAAGGAGGCCAGCGGAGTACCCCCTTGATTATAGCCGCACGTAACGGCCACGCGAAGGTC
GTGAGACTCCTACTCGAACATTATCGGGTACAAACGCAGCAAACAGGAACCGTACGGTTCG
ACGGATACGTTATAGATGGCGCGACAGCTTTATGGTGTGCCGCAGGCGCCGGTCACTTCGA
AGTAGTC
```

3.2 Sequence statistics analysis

Sections of a nucleotide sequence with a certain percentage of A+T or C+G usually indicates intergenic parts of the sequence. Figure 9 is a plot of monomer densities and combined monomer densities. One can use such statistic plot to determine if the sequence has the characteristics of a protein-coding region.

Figure 10 is the visualization of the nucleotide distribution. Figure 11 is the codon distribution showing a high amount of GAA, GAT and AAC. The amino acids for GAA, GAT and AAC are Glutamate, Aspartate, and Asparagine respectively. The corresponding bar chart distribution is displayed at figure 12. It is noticeable that it contains high volume of leucine, alanine, and valine; low volume of tryptophan, methionine, and proline.

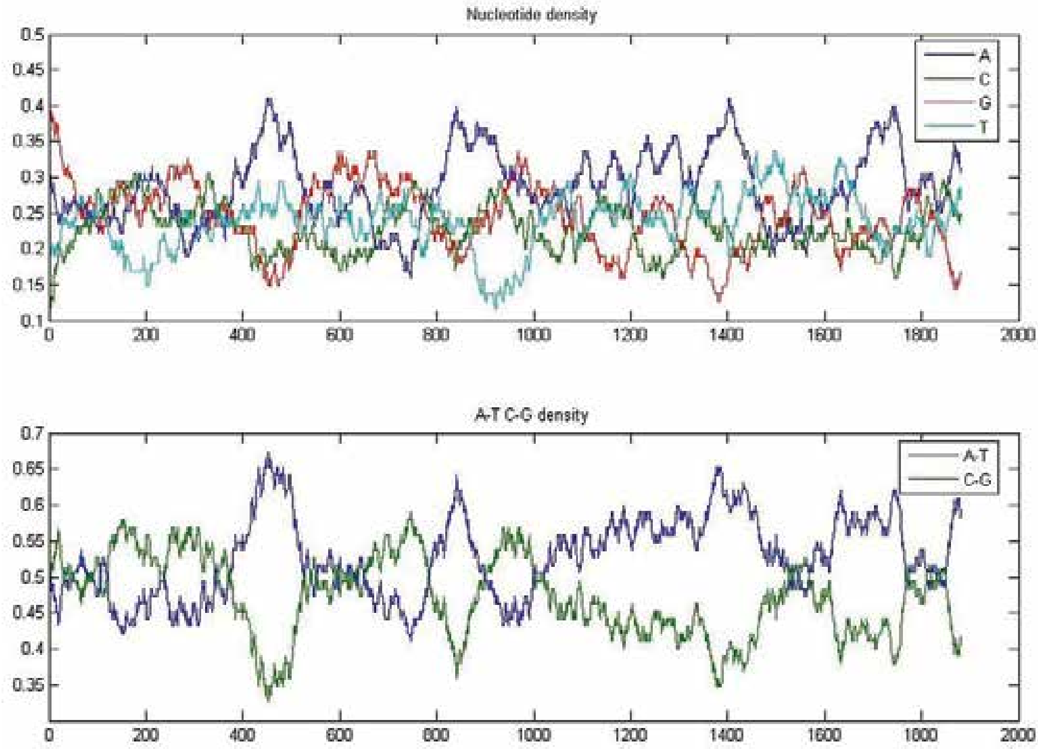


Fig. 9. Human fem1b gene's monomer densities and A-T &C-G combined monomer densities.

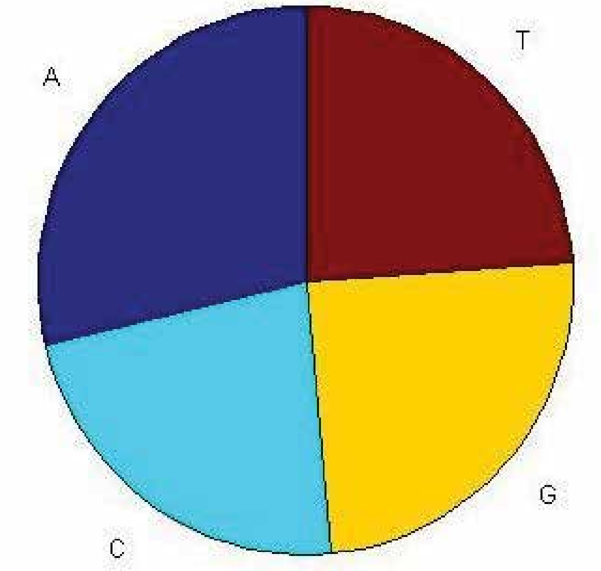
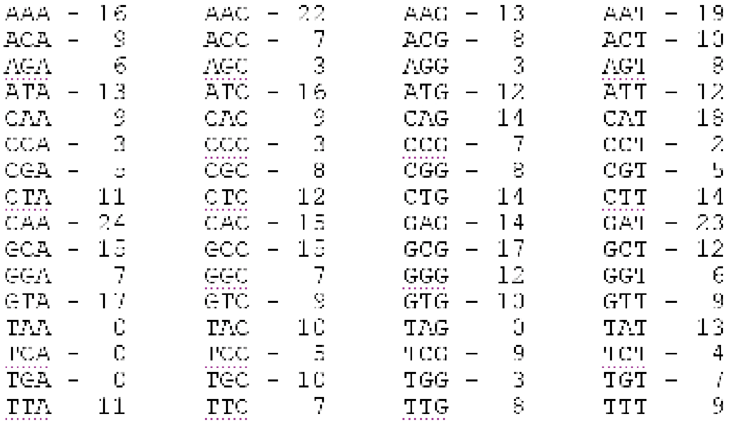


Fig. 10. Human fem1b gene's nucleotide distribution (A: 542, C: 426, G: 462, T: 451).



```
AAA - 16      AAC - 22      AAG - 13      AAT - 19
ACA -  9      ACC -  7      ACG -  8      ACT - 10
AGA    6      AGC    3      AGG    3      AGT    8
ATA - 13      ATC - 16      ATG - 12      ATT - 12
CAA    9      CAC    9      CAG   14      CAT   18
CCA -  3      CCC -  3      CCG -  7      CCT -  2
CGA -  5      CGC -  8      CGG -  8      CGT -  5
CTA   11      CTC   12      CTG   14      CTT   14
GAA - 24      GAC - 15      GAG - 14      GAT - 23
GCA - 15      GCC - 15      GCG - 17      GCT - 12
GGA    7      GGC    7      GGG   12      GGT    6
GTA - 17      GTC -  9      GTG - 10      GTT -  9
TAA    0      TAC   10      TAG    0      TAT   13
TCA -  0      TCC -  5      TCG -  9      TCT -  4
TGA -  0      TGC - 10      TGG -  3      TGT -  7
TTA   11      TTC    7      TTG    8      TTT    9
```

Fig. 11. Human fem1b gene codon distribution.

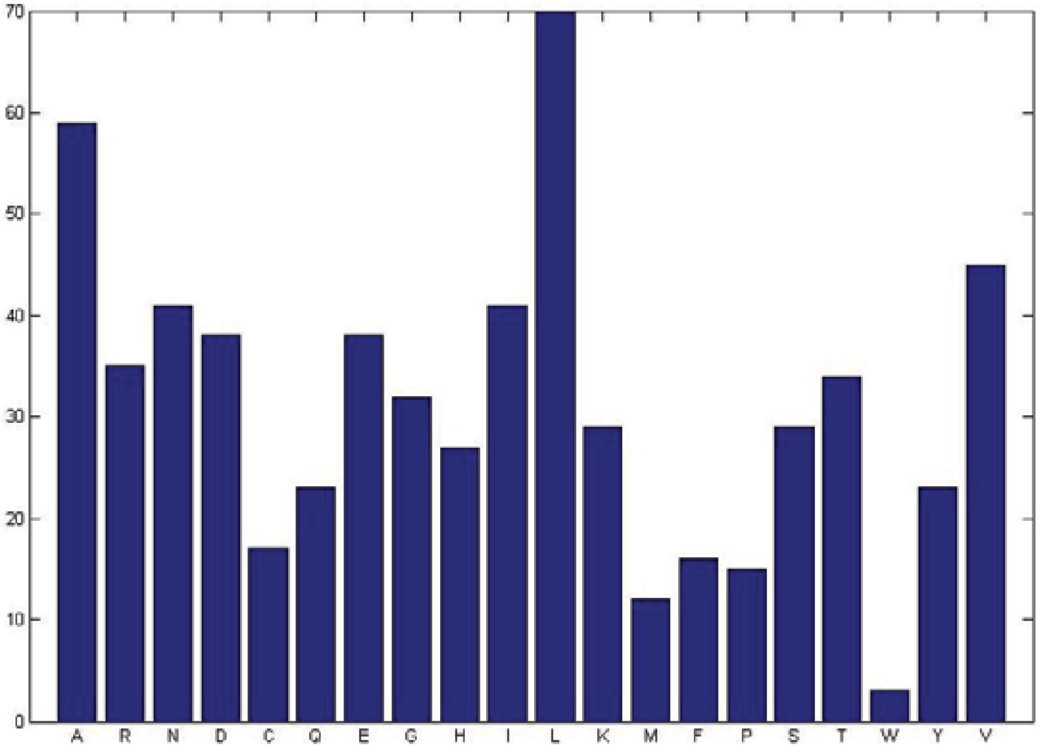


Fig. 12. Amino acids distribution of human fem1b gene.

A: 59 R: 35 N: 41 D: 38 C: 17 Q: 23 E: 38 G: 32 H: 27 I: 41 L: 70 K: 29 M: 12 F: 16 P: 15 S: 29 T: 34 W: 3 Y: 23 V: 45

3.3 Open reading frame of Fem1b gene from both human and mouse

An open reading frame (ORF) is a nucleotide sequence without having a stop codon in a given reading frame. ORFs can be identified by examining each of the three possible reading frames on each strand. A DNA sequence must contain a translation start codon and it is usually "AGT". Possible stop codons are "TAA", "TAG and "TGA" [11]. Identifying the start and stop codons for translation determines the ORF in a given nucleotide sequence. Once an ORF is located for a gene or mRNA, a nucleotide sequence can be translated into its corresponding amino acid sequence. Figure 13 – 15 display three reading frames for human's and mouse's fem1b gene sequences. Both genes show the longest ORF on the first reading frame.

Dot plots are one of the easiest ways to look for similarity between two sequences. The diagonal line shown in figure 16 indicates a good alignment between the human's and mouse's fem1b gene.

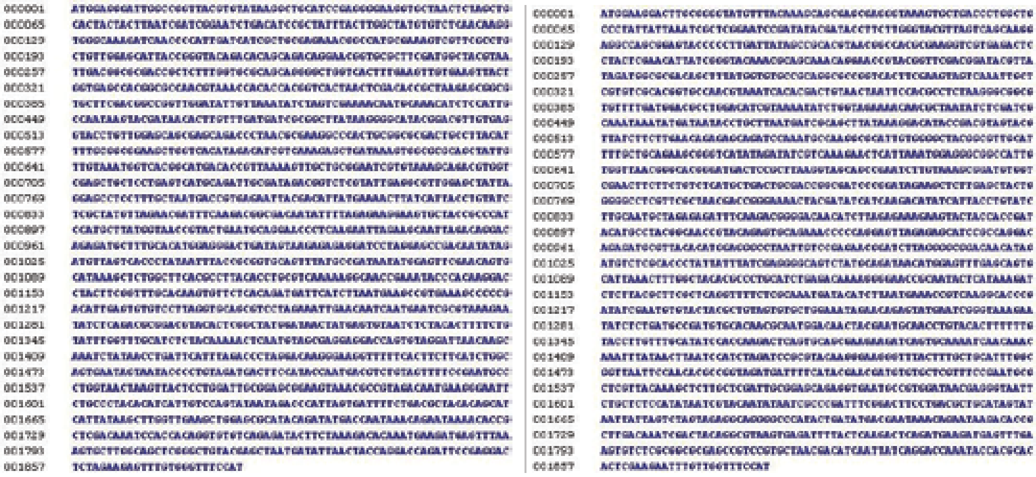


Fig. 13. First ORF of fem1b gene (left – mouse; right – human).

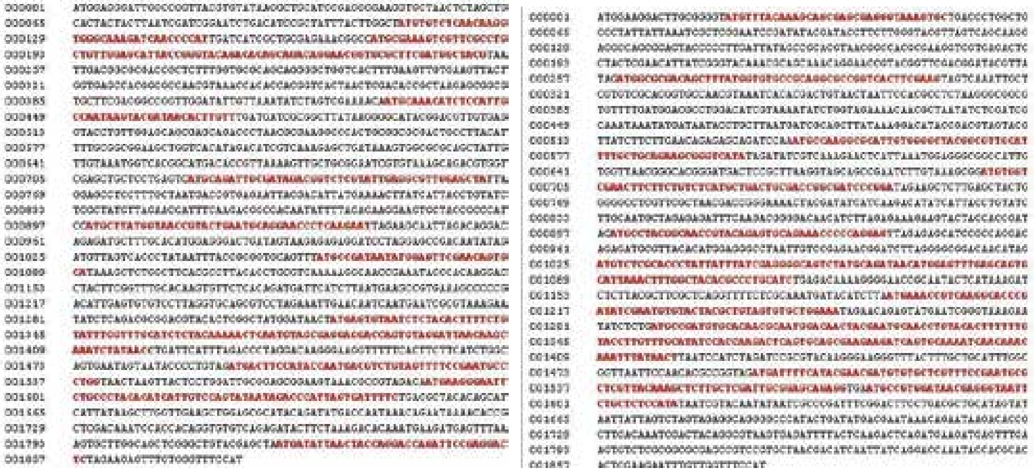


Fig. 14. Second ORF of fem1b gene (left – mouse; right – human).

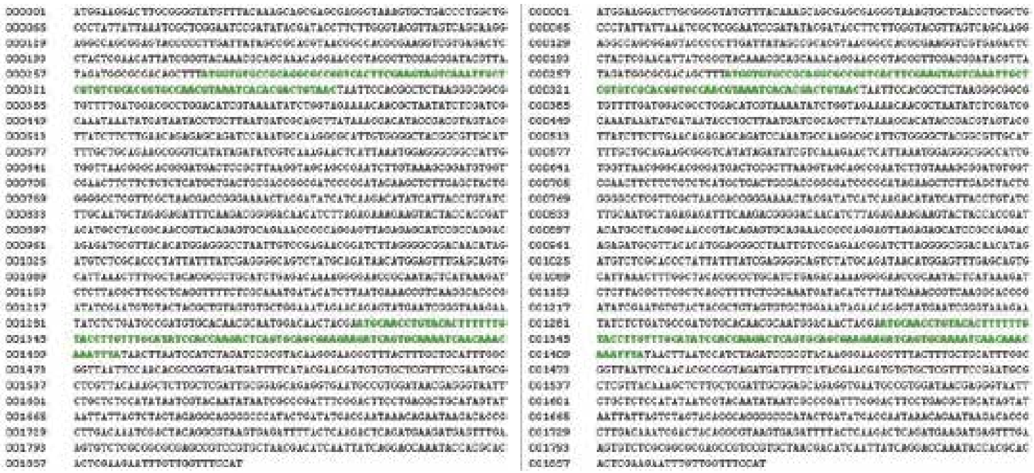


Fig. 15. Third ORF of fem1b gene (left – mouse; right – human).

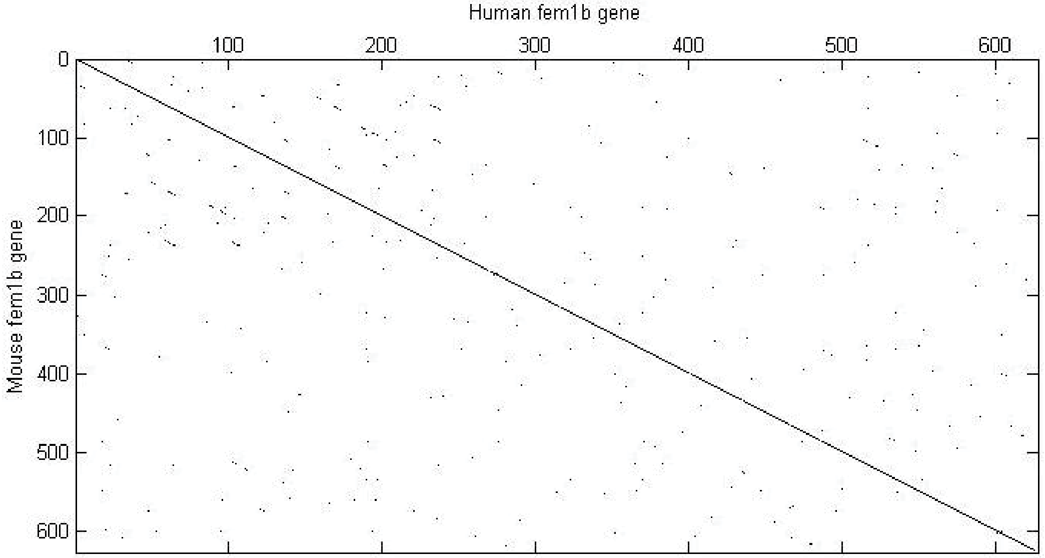


Fig. 16. Dot plot comparing the human and mouse amino acid sequences.

3.4 Sequence alignment

3.4.1 Global alignment

The Needleman-Wunsch algorithm, which was first published by Saul Needleman and Christian Wunsch in 1970 [12], performs a global alignment on two amino acid or nucleotide sequences. Such algorithm was the first application of dynamic programming to molecular sequence comparison. The following output was performed on two nucleotide sequences of mouse's and human's by the Needleman-Wunsch algorithm

```
Identities = 1430/1890 (76%), Positives = 1591/1890 (84%)

0001  ATGGAGGGATTGGCCGGTTACGTGTATAAGGCTGCATCCGAGGGGAAGGTGCTAACTCTAGCTG
      |||||:||| | || || || || || ||:||:||:: |||||| ||:|||||:|| ||:||||
0001  ATGGAAGGACTTGCGGGGTATGTTTACAAAGCAGCGAGCGAGGGTAAAGTGCTGACCCTGGCTG
0065  CACTACTACTTAATCGATCGGAATCTGACATCCGCTATTTACTTGGCTATGTGTCTCAACAAGG
      | ||| || |:||||| |||||||| || || || ||  |:||||| || || : |||:|||||
0065  CCCTATTATTAAATCGCTCGGAATCCGATATACGATACCTTCTTGGGTACGTTAGTCAGCAAGG

0129  TGGGCAAAGATCAACCCCATTGATCATCGCTGCGAGAAACGGCCATGCGAAAGTCGTTCGCCTG
      :|| ||: |:: :||||| ||||| || || ||: |:|||||||| |||||:|||||  | ||
0129  AGGCCAGCGGAGTACCCCCTTGATTATAGCCGCACGTAACGGCCACGCGAAGGTCGTGAGACTC
```

3.4.2 Local alignment

The Smith-Waterman algorithm was first published by Temple Smith and Michael Waterman in 1981 [13]. It is a well-known dynamic programming algorithm for local amino acid or nucleotide sequence alignment. Unlike the global alignment, the Smith-Waterman algorithm performs comparison among segments of all lengths and optimizes the similarity. It is guaranteed to find the optimal local alignment with respect to the scoring method. However, the Smith-Waterman algorithm requires *O(mn)* (m and n are the length of two input sequences) . In practical use, it has been replaced by the heuristic BLAST algorithm, which is much more efficient although not guaranteed to find the optimal alignments. The following output was from local alignment of the amino acid sequences of mouse's and human's using the Smith-Waterman algorithm.

```
Identities = 621/627 (99%), Positives = 627/627 (100%)

001  MEGLAGYVYKAASEGKVLTLAALLLNRSESDIRYLLGYVSQQGGQRSTPLIIAARNGHAKVVRL
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
001  MEGLAGYVYKAASEGKVLTLAALLLNRSESDIRYLLGYVSQQGGQRSTPLIIAARNGHAKVVRL

065  LLEHYRVQTQQTGTVRFDGYVIDGATALWCAAGAGHFEVVKLLVSHGANVNHTTVTNSTPLRAA
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
065  LLEHYRVQTQQTGTVRFDGYVIDGATALWCAAGAGHFEVVKLLVSHGANVNHTTVTNSTPLRAA
```

4. Diabetic retinopathy study through retinal image fusion

Hypoxia of the retina is believed to be a factor in the development of diabetic retinopathy, the leading cause of blindness worldwide. Retina image fusion provides a practical way for determination of the oxygenation status of the ocular fundus. Such method would be a valuable medical diagnostic tool for diabetic retinopathy [14], age-related macular degeneration, glaucoma [15], retinopathy of prematurity, and central retinal vein occlusion [16].

4.1 Acquisition of retinal images

Retinal images presented in this chapter were taken by a modified Topcon TRC-50EX fundus camera, with a lens and a c-mount through the vertical path of the camera. Hyperspectral images were taken through the vertical viewing port by an imaging spectrograph and digital camera (model VNIR 100; Photon Industries Inc., Mississippi Stennis Space Center, USA) across the fundus image (Figure 17).

The subjects of the retinal images were Cynomolgus monkeys of 4 to 4.5 years of age and 2.5 to 3 kg body weight with normal eyes [17]. The use of animals for taking retinal images was approved by Louisiana State University Health Sciences Center Institutional Animal Care and Use Committee [18]. This animal usage is also conformed to the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research. The monkeys were housed in an air conditioned room with normal temperature and humidity with a 12 hour light-dark diurnal cycle.

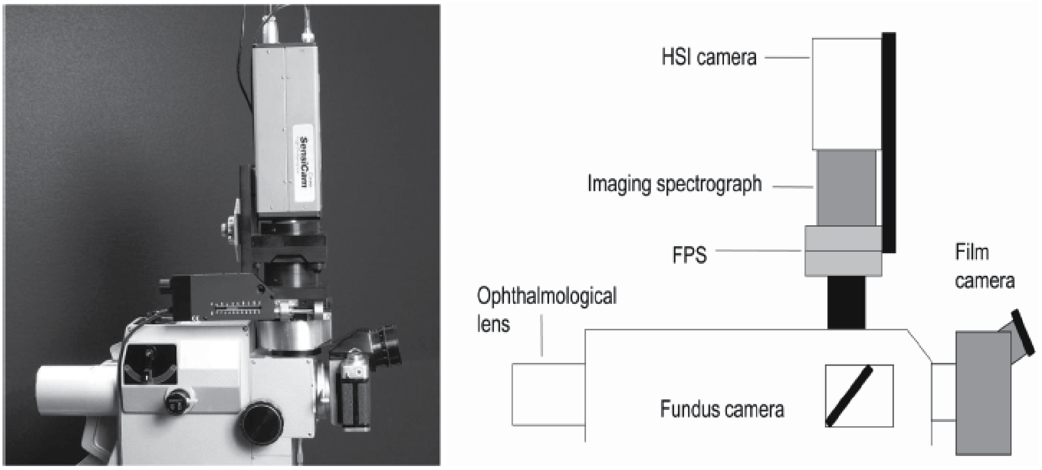


Fig. 17. Hyperspectral imaging system in relation to the fundus camera. The image is redirected upward by a mirror. The imaging system is translated over the camera port by a linear actuator mounted below the imaging spectrograph and CCD camera [17].

4.2 Retinal image fusion

There are five major steps involved in image fusion (Diagram 1):

1. The first step is the image segmentation. The segmentation subdivides an input image into its constituent regions or objects and extract/detect salient features/structure for the automated procedure.
2. The second step feature extraction is going to detect the salient structures on the target images for the feature-based approach.
3. The third step is the feature matching. The purpose of feature matching is to bringing together the information that represents same features detected at different images. The first three steps will provide the initial guess of the features for the fusion algorithm.
4. The fourth step is the optimization of the initial guess. The previously detected features will be adjusted in this step through a certain objective function.
5. The final step will transforms the images from single or different modalities into spatial alignment [19] through a certain mathematical model and then display combined view of the involved images.

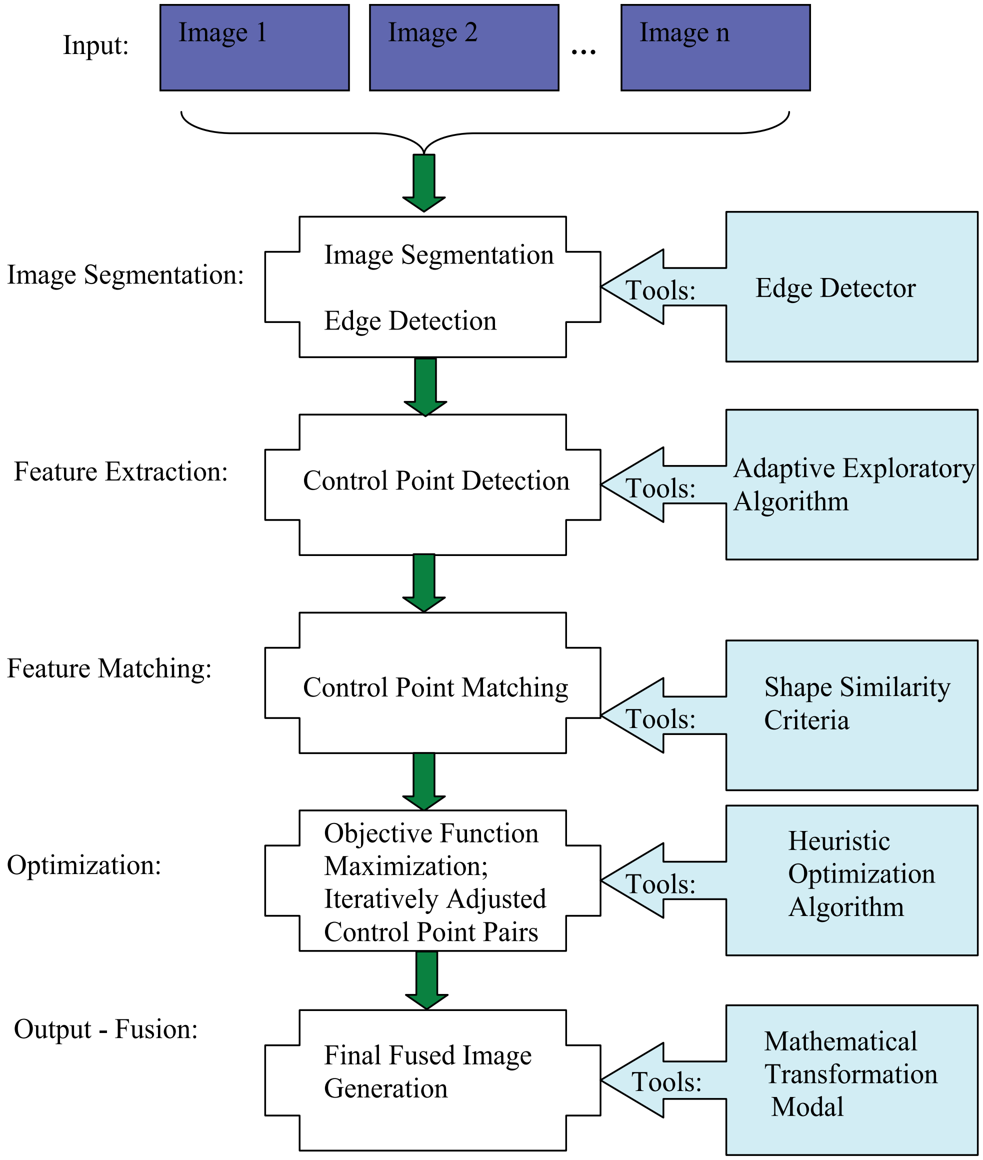

Input:
Image 1
Image 2
...
Image n
Image Segmentation:
Image Segmentation
Edge Detection
Tools:
Edge Detector
Feature Extraction:
Control Point Detection
Tools:
Adaptive Exploratory Algorithm
Feature Matching:
Control Point Matching
Tools:
Shape Similarity Criteria
Optimization:
Objective Function Maximization; Iteratively Adjusted Control Point Pairs
Tools:
Heuristic Optimization Algorithm
Output - Fusion:
Final Fused Image Generation
Tools:
Mathematical Transformation Modal

Diagram 1. Image fusion procedure.

4.2.1 Retinal vasculature extraction using canny edge detector

The Canny operator [20] is less likely than the others to be "fooled" by noise, and more likely to detect true weak edges. The Canny operator preserves most edges among all other edge detectors. Therefore, the Canny edge detector is employed in this research to extract the retinal vasculature edges. There are two criteria used in the Canny Operator to locate the rapidly changed intensity pixels. They are:

1. Pixels where the 1st derivative of the intensity is larger in magnitude than a certain threshold;

2. Pixels where the 2nd derivative of the intensity has a zero crossing.
3. Canny's method detects edges at the zero-crossings of the second directional derivative of the image. It performs zero-crossings of

$$\frac{d^2(G\times I)}{dn^2}=\frac{d\left(\left(\frac{dG}{dn}\right)\times I\right)}{dn} \tag{1}$$

where, n is the direction of the gradient of the image; G is the edge signal; I is the image intensity. The zero-crossings of Canny's method correspond to the first directional-derivative's maxima and minima in the direction of the gradient. Edges will be identified as the maxima in magnitude. Each pixel's edge gradient is computed and compared with the gradients of its neighbors along the gradient direction. If the central pixel is smaller, mark the current edge's intensity as 0; if largest among all neighbors, keep the original intensity. Based on the nine-pixel neighborhood, the normal to the edge direction has two u_x and u_y. In order to estimate the gradient on the discrete sampling, two pixels closest to u are selected. A plane can be identified by the gradient magnitudes of three pixels. By using this plane, the gradient magnitude and the intensity at each pixel on the line can be locally estimated. The gradient magnitude at $P_{x+1,\ y+1}$ and $P_{x-1,\ y-1}$ (Figure 18) can be calculated as:

$P_{x+1,\ y+1}$:
$$G(P_{x+1,y+1})=\frac{u}{u_y}G(x+1,y+1)+\frac{u_y-u_x}{u_y}G(x,y+1) \tag{2}$$

$P_{x-1,\ y-1}$:
$$G(P_{x-1,y-1})=\frac{u}{u_y}G(x-1,y-1)+\frac{u_y-u_x}{u_y}G(x,y-1) \tag{3}$$

If the gradient at $P_{x,y}$ is greater than both of $G(P_{x+1,y+1})$ and $G(P_{x-1,y-1})$, $P_{x,y}$ will be identified as a maximum.

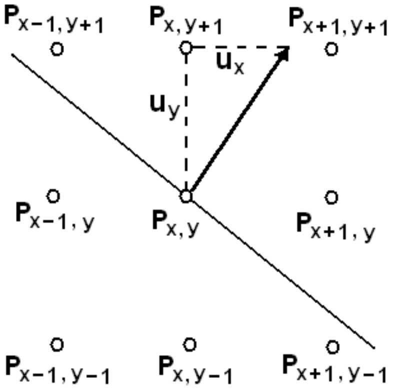


Fig. 18. Canny Edge Detection – Localization of Maxima [34].

In order to make the localization of magnitude maxima accurate, Canny defined a filter by optimizing a performance index which enhances real positive and real negative. The filter is used to minimize the probability of non-detected edge points and false detection.

$$SNR = \frac{\left| \int_{-w}^{w} G(-x) f(x) dx \right|}{n_0 \sqrt{\int_{-w}^{w} f^2(x) dx}} \tag{4}$$

where, SNR stands for Signal-to-Noise Ratio; *f* is the filter; denominator is the RMSE response to noise *n(x)*. The identification of the real edge localization is defined as:

$$Localization = \frac{1}{\sqrt{E[x_0^2]}} = \frac{\left| \int_{-w}^{w} G'(-x) f'(x) dx \right|}{n_0 \sqrt{\int_{-w}^{w} f'^2(x) dx}} \tag{5}$$

Two adaptive thresholds are used in Canny's method. They are high threshold and low threshold. The high threshold is used to find the start point of strong edges. Any points that meet the high threshold will be selected as the edge point. These start points are growing into different directions as long as there is no edge strength falling below the low threshold.

Figure 19 is the 3D shaded surface plots of the original retinal angiogram image. The *X*-*Y* axis corresponds to the original image size. The height *Z* axis is a single-valued function defined over a geometrically rectangular grid. *Z* specifies the color data as well as surface height, so color is proportional to surface height with range of [0, 1] of each pixel on the image. All the retinal salient features are preserved in the Canny edges. Those salient features are the retinal vessel bifurcations, from which the control points will be selected using the Adaptive Exploratory Algorithm. Figure 20 and 21 show the retinal vessel edges detected by the Canny operator.

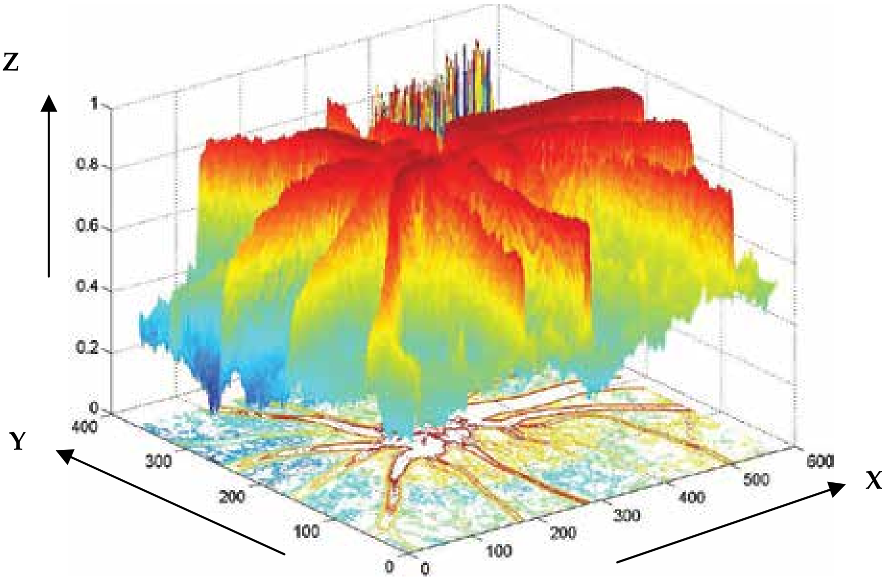


Fig. 19. 3D shaded surface plot of the angiogram image.

4.2.2 Control point detection

A good-guess of the initial control point selection ensures fused image generated at an efficient computational time. Bad control point selection will significantly increase the computation cost, or even cause the image fusion fail. Vessels or some particular abnormalities make images not necessarily matching the retina structures. Even when structure and function correspond, the abnormality still happens sometimes if inconsistence exists between structural and functional changes. Further more, angiogram images usually have higher resolution and are rich in information, whereas fundus images have lower resolution and are indeed abstract with some details or even missing some small vessels. Practically, those situations are unavoidable and will create difficulties in extracting the control points because the delineation of the vein boundaries may not be precise. In this study, control points are detected using the adaptive exploratory algorithm (Figure 20 and 21) [21].

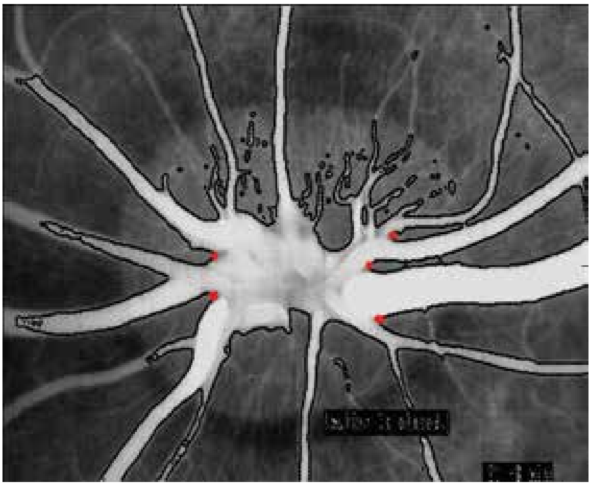


Fig. 20. Angiogram grayscale reference image's control point selection.

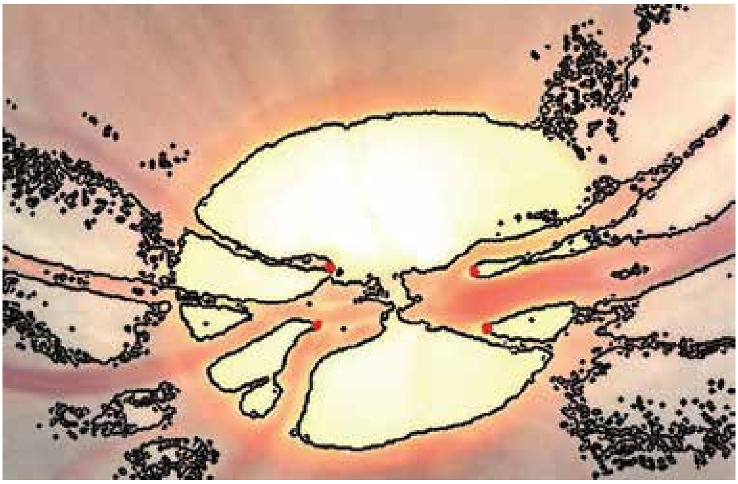


Fig. 21. Fundus true color input image's control point selection.

4.2.3 Heuristic optimization algorithm

An optimization procedure is required to adjust the initial good-guess control points in order to achieve the optimal result. The process can be formulated as a heuristic problem of optimizing an objective function that maximizes the Mutual-Pixel-Count between the reference and input images. The algorithm finds the optimal solution by refining the transformation parameters in an ordered way. By maximizing the objective function, one image's vessels are supposed to be well overlaid onto those of the other image (Figure 23). Mutual-Pixel-Count measures the optic nerve head vasculature overlapping for corresponding pixels in both images. It is assumed that the retinal vessels are represented by 0 (black pixel) and background is represented by 1 (white pixels) in the binary 2D map. When the vasculature pixel's transformed *(u, v)* coordinates on the input image correspond to the vasculature pixel's coordinates on the reference image, the MPC is incremented by 1 (Figure 22). MPC is assumed be maximized when the image pair is perfectly geometrically aligned by the transformation. After pre-processing, the binary images of the reference and input images are obtained, i.e. I_{ref} and I_{input}. Only black pixels from both images contribute to MPC. The ideal case is that all zero pixels of the input image are mapped onto zero pixels of the reference image. The problem can be mathematically formulated as the maximization of the following objective function:

$$f_{mpc}(x,y,u,v) = \sum_{I_{ref}(x,y)=1 \ \text{and} \ I_{input}(u,v)=1}^{u,v \in ROI} I_{input}(T_x(u,v),T_y(u,v)) \qquad (6)$$

where f_{mpc} denotes the value of the Mutual-Pixel-Count. T_x and T_y are the transformations for u and v coordinates of the input image. The ROI (Region-of-Interest) is the vasculature region where the MPC is calculated on.

Coordinates' adjustment is iteratively implemented until one of the following convergence criteria is reached: (1). Predefined maximum number of loops is reached; or (2). the updated f_{MPC} is smaller than ε , i.e.

$$\left| f_{MPC^{n+1}}(x,y,u,v) - f_{MPC^{n}}(x,y,u,v) \right| < \varepsilon \qquad (7)$$

where ε is a very small non-negative threshold.

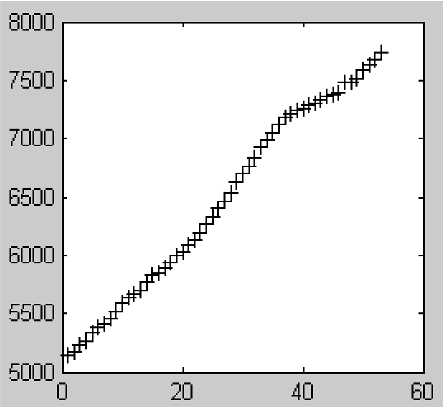
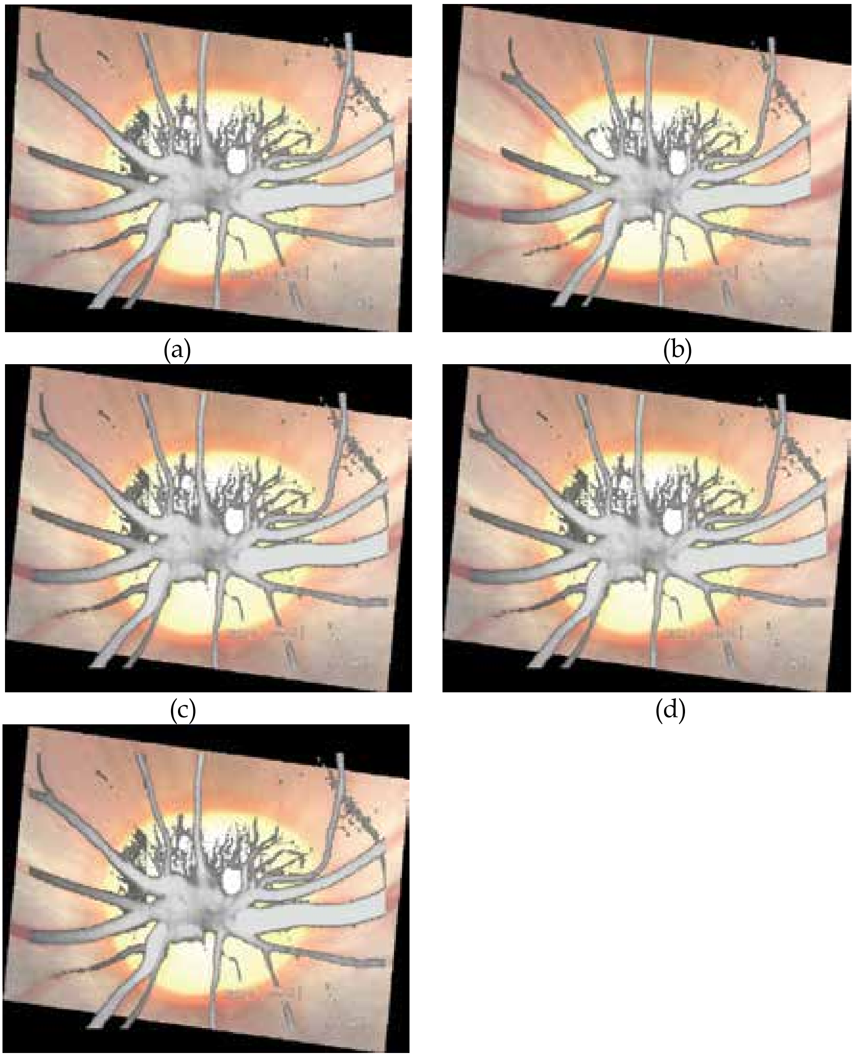


Fig. 22. f_{MPC} increasing during the iteration (Y-axis is f_{MPC}; X-axis is the loop count).



(a) (b)

(c) (d)

(e)

Fig. 23. Fused image improvement during the iteration. From (a)-(e) fMPC = 5144, 7396, 7484, 7681, 7732.

4.2.4 Affine transformation model

A mathematical model is the tool for transforming the target images and fusing them into one single volume. Affine model is a basic geometrical transformation in image processing and is defined as Eq. 8 and 9. The DOF of the affine model is 6 because it has six parameters, i.e. a_1, a_2, b_1, b_2, a_3, and a_4.

$$\begin{pmatrix} U \\ V \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & b_1 \\ a_3 & a_4 & b_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \begin{aligned} U(x, y) &= a_1x + a_2y + b_1 \quad (8) \\ V(x, y) &= a_3x + a_4y + b_2 \quad (9) \end{aligned}$$

Affine model's advantage lies in that it can measure lost information such as skew, translation, rotation, shearing and scaling that maps finite points to finite points and parallel lines to parallel lines (Figure 24 and Table 1). Its drawback lies in the strict requirement that at least six pairs of control points are needed [19].

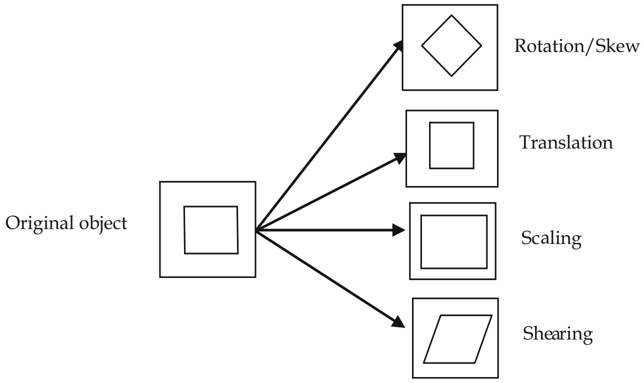


Fig. 24. Transformation Graphs.

Transformation	Description
Rotation/Skew	Points are rotated by an angle θ.
Translation	A linear shift in the position of the vertical and horizontal coordinates of the image in one plane to another set in the same spatial domain.
Scaling	A transformation of the horizontal and vertical coordinate points characterized by a certain scale factor.
Shearing	A transformation in which all points along a line remain fixed while other points are shifted parallel to the line by a certain distance proportional to their perpendicular distance from the line [22].

Table 1. Transformation Descriptions.

5. Conclusion and future directions

Fem1b functions in vivo to regulate insulin secretion and plasma glucose levels. Fem1b-KO mice do not have fasting hyperglycemia but rather have defective acutephase GSIS. Such defective acute-phase GSIS is the earliest detectable defect in humans destined to develop diabetes and may represent the primary genetic risk factor predisposing to diabetes [23]. With aging and superimposed insulin resistance, fasting hyperglycemia and overt diabetes later develop. Therefore, the Fem1b-KO mouse model is a key component of the complex pathogenesis of type 2 diabetes mellitus. Both male and female homozygotes display abnormal glucose tolerance. The role of Fem1b in pancreatic islet insulin secretion strengthens evidence that a genetic pathway homologous to nematode sex determination may be involved in mammalian glucose homeostasis. This novel pathway could be involved in the β-cell dysfunction seen in type 2 diabetes mellitus. Since Calpain-10/NIDDM1 is known to interact with a gene that is near where human FEM1B localizes, to increase susceptibility to type 2 diabetes[24], whether FEM1B could be the responsible interacting gene becomes a pertinent question Although the mechanism of this regulation by Fem1b remains to be established, this finding strengthens evidence that a genetic pathway homologous to nematode sex determination may be involved in mammalian glucose homeostasis and promises to offer insight into novel genes and processes as potential candidates in the pathogenesis of diabetes mellitus.

Multi-modality analysis has been emerging as a major trend in the remote sensing, computer visualization, and biomedical image fusion. Fusing biomedical images is a very challenging problem because of the possible vast content change and non-uniform distributed intensities of the involved images. The new algorithm presented in this chapter, which consists of the Adaptive Exploratory Algorithm for the control point detection and heuristic optimization fusion, is reliable and time efficient. The new approach has achieved an excellent result by giving the visualization of fundus image with a complete angiogram overlay. By locking the multi-sensor images in one place, the algorithm allows ophthalmologists to match the same eye over time to get a sense of disease progress and pinpoint surgical tools to increase accuracy and speed of the surgery. The new algorithm can be easily expanded to human or animals' 3D eye, brain, or body image feature extraction, registration, and fusion. Many biomedical registration and fusion methods are still primarily used for research activity [19]. Very few of them have been developed into the integrated user-friendly computer software. The eventual aim is to developing and distributing the advanced and easy-to-use software which is suitable for various clinical environments. This plan requires intensive user interface developing work, which allows users adjusting a few threshold parameters if necessary. The user interface must be stable, simple and informative for every day clinical routine, because most of the end-users do not have much knowledge about algorithms and computer programs. Working closely with clinicians is extremely important when applying the new methods to practical clinic applications. As computation speed has been dramatically increased, real-time live ophthalmic image processing [25] will be used to handle larger and larger volumes of data in short periods. The data transmission rate, image size, higher resolution pixels, and many other issues will inevitably stress such live imaging fusion systems. The algorithms presented in this book have potential ability to handle those challenges. The presented method is a promising step towards useful clinical tools for retinopathy diagnosis, and thus forms a good foundation for further development.

6. References

[1] Klein R, Klein BE, Moss SE, Davis MD, DeMets DL. "The Wisconsin Epidemiologic Study of Diabetic Retiopathy, II: prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 years." *Arch Ophthalmol;* 102: 520-526. 1984

[2] Ryan, M; "Diabetes The Leading Cause of New Cases of Blindness"; Ohio Department of Health, Public Affairs; November 2003;

[3] Brissova, M., M. Shiota, W. E. Nicholson, M. Gannon, S. M. Knobel, D. W. Piston, C. V. Wright, and A. C. Powers. "Reduction in pancreatic transcription factor PDX-1 impairs glucose-stimulated insulin secretion"; *J. Biol. Chem.* 277:P11225–11232. 2002

[4] KELLEY, D. E., and B. H. GOODPASTER. Effects of exercise on glucose homeostasis in Type 2 diabetes mellitus. *Med. Sci. Sports Exerc.,* Vol. 33, No. 6, Suppl., p S495–S501.

[5] J. MAYFIELD, "Diagnosis and Classification of Diabetes Mellitus: New Criteria"; Bowen Research Center, Indiana University, Indianapolis, Indiana

[6] Cox, N. J., M. Frigge, D. L. Nicolae, P. Concannon, C. L. Hanis, G. I. Bell, and A. Kong. "Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans"; *Nat. Genet.* 21: 213–215.1999

[7] Ventura-Holman, T., N. B. Hander, and J. F. Maher. “The human FEM1B gene maps to chromosome 15q22 and is excluded as the gene for Bardet-Biedl syndrome, type 4”: *Am. J. Med. Sci.* 319:268–270. 2000

[8] D. Lu, Tereza Ventura-Holman, J. Li, Robert W. McMurray, Jose S. Subauste, and Joseph F. Maher. “Abnormal Glucose Homeostasis and Pancreatic Islet Function in Mice with Inactivation of the Fem1b Gene”. *Molecular and Cellular Biology.* 25(15): 6570–6577;2005

[9] Ventura-Holman, T., M. F. Seldin, W. Li, and J. F. Maher. “The murine Fem1 gene family: homologs of the Caenorhabditis elegans sex-determination protein FEM-1”; *Genomics* 54:221–230. 1998

[10] Chan, S. L., K. O. Tan, L. Zhang, K. S. Yee, F. Ronca, M. Y. Chan, and V. C. Yu. “F1Aα, a death receptor-binding protein homologous to the Caenorhabditis elegans sex-determining protein, FEM-1, is a caspase substrate that mediates apoptosis”. *J. Biol. Chem.* 274:32461–32468. 1999

[11] J. Claverie; C. notredame; “Bioinformatics For Dummies, 2nd Edition”; ISBN: 978-0-470-08985-9; Wiley Publishing, Inc; 2007

[12] Needleman, Saul B.; and Wunsch, Christian D. "A general method applicable to the search for similarities in the amino acid sequence of two proteins"; *Journal of Molecular Biology*; V48, I3: P443–53; 1970.

[13] Smith, Temple; and Waterman, Michael. "Identification of Common Molecular Subsequences"; *Journal of Molecular Biology;* V147: P195–197. 1981

[14] Denninghoff KR, Smith MH, Hillman L. “Retinal imaging techniques in diabetes”; *Diabetes Technol Ther*. V2:111-113. 2002

[15] Schweitzer D, Hammer M, Kraft J, Thamm E, Konigsdorffer E, Strobel J. In vivo measurement of the oxygen saturation of retinal vessels in healthy volunteers. Transactions on Biomdeical Engineering. 46: 1454-1465. 1999

[16] Yoneya A, Saito T, Nishiyama Y, et al. “Retinal oxygen saturation levels in patients with central retinal vein occlusion”. *Ophthalmology*;109:1521-1526. 2002

[17] J. Beach, J. Ning, B. Khoobehi; “Oxygen Saturation in Optic Neve Head Structures by Hyperspectral Image Analysis”; Current Eye Research; V32, P161-170, 2007.

[18] B. Khoobehi, J. Beach, H. Kawano; “Hyperspectral imaging for measurement of oxygen saturation in the optic nerve head”; Invest Ophthalmol Vis Sci. V45, P1464–1472; 2004

[19] H. Cao; “Automated Fusion of Multi-Modality Biomedical Images”; ISBN: 363906622;VDM, Verlag Dr. Muller Publisher, August 2008.

[20] J.F. Canny; "A computational approach to edge detection"; IEEE Transactions on Pattern Analysis and Machine Intelligence; V8, P 679-698; 1986.

[21] H. Cao, B. Khoobehi, S. S. Iyengar; "Automated Optic Nerve Head Image Fusion of Nonhuman Primate Eyes Using Heuristic Optimization Algorithm"; 5th IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB 2008); Sun Valley, Idaho, USA; September 15-17, 2008; P 228-232

[22] Z. Millwala; “A dual-state approach to dental image registration”; Masters Thesis; *Lane Department of Computer Science and Electrical Engineering; West Virginia University*; 2004

[23] Denninghoff KR, Smith MH, Hillman L. “Retinal imaging techniques in diabetes”; *Diabetes Technol Ther*. V2:111-113. 2002

[24] Schweitzer D, Hammer M, Kraft J, Thamm E, Konigsdorffer E, Strobel J. In vivo measurement of the oxygen saturation of retinal vessels in healthy volunteers. Transactions on Biomdeical Engineering. 46: 1454-1465. 1999

[25] Yoneya A, Saito T, Nishiyama Y, et al. “Retinal oxygen saturation levels in patients with central retinal vein occlusion”. *Ophthalmology*;V109: P1521-1526. 2002

19

A Shape-Factor Method for Modeling Parallel and Axially-Varying Flow in Tubes and Channels of Complex Cross-Section Shapes

Mario F. Letelier and Juan S. Stockle
University of Santiago of Chile,
Chile

1. Introduction

In the study of some industrial, biological and natural fluidic systems it is often necessary to model fluid flow through tubes, channels or passages of complex geometries. The complexity may arise from the cross-sectional shape, or from longitudinal cross-section variation, or from both. Typical cases include flow of molten metals or plastics through dies and moulds, blood flow, microfluidic applications, and flow in porous media, among many others. Characteristics of these flows are laminar state, incompressibility, small rates of flow and varied time patterns. One field where pertinent applications are being developed at a fast rate is Microfluidics (Cetin and Li., 2008; Chen et al., 2008; Forte et al., 2008; Gebauer and Bocek, 2002 ; Mathies and Huang, 1992; Sommer et al., 2008; Srivastava et al., 2005; Woolley and Mathies, 1994; Yeger et al., 2006.) . In this specific field, present microchannel manufacturing techniques produce typically non-circular capillaries (Sommer et al., 2008). Also the introduction of electrical or magnetic field induce plastic behavior in the working fluid.

In particular, it is well known that blood is a biological fluid that behaves as a Newtonian fluid in arteries, veins and large capillaries, but becomes non-Newtonian in the smaller vessels, where the size of suspended particles is big as compared to the vessel´s diameter size (Pedley ,2008). A relevant problem in this field as to the method presented in the next sections is the analysis of diseased arteries and veins for quirurgical interventions. Specifically, stenosed arteries are blood conduits of irregular geometry in which cross-section geometry usually varies along the vessel length.

The above context implies that it is desirable, particularly for modeling and design purposes, to count with analytical techniques that can integrate variables such as the non-circular cross-section of conduits, axial variation of conduit geometry, and plastic flow in some cases.

In this chapter it is presented a method of analysis that allows to address in a general way the problem here outlined.

The standard analytical technique for tube flow problems is usually the search of specific solutions to the momentum equations with associated boundary and initial conditions (Batchelor, 2000). Otherwise numerical solutions are developed for some purposes (Xue et al.,1995).

The main aim of this chapter is, thus, to introduce and explore the potential use of a general analytical approach to irregular conduit flow, which makes it possible to determine velocity field, rate of flow, shear stress, recirculation regions and plug zones, this last when fluid plasticity is operant.

The method already referred to has been developed by the authors through specific applications mainly during the past decade. In this chapter some previous results are organized within a common analytical pattern, together with novel material.

This chapter includes sections for the *general model*, considering one velocity component and more than one velocity component versions, *applications* related to flow in straight tubes and to axially-varying flows, and a closing *conclusion* section.

2. The general model

The concept of "shape factor" herein used is applied to a function G of spatial coordinates, such as when $G = 0$, a series of closed curves are determined for a range of some parameters contained in G. One typical example is

$$G = 1 - r^2 + \varepsilon r^n \sin n\,\theta \tag{1}$$

In this (r, θ) are polar coorfinates, n is an integer number and ε is a parameter such as that for $\varepsilon = 0$ the curve described by (1) is a circle, and as ε increases, the shape evolves to some limiting shape, controlled by n. In all cases here considered, the maximum allowable value of ε is less than unity, and beyond that value, the curve is no longer a closed one. If $\varepsilon = \varepsilon_c$ is the critical, or maximum, allowable value of ε, then for the shape factor described by (1), ε_c is found to be

$$\varepsilon_c = \frac{2}{n}\left(\frac{n-2}{n}\right)^{(n-2)/2} \tag{2}$$

n	ε_c
3	0,385
4	0,250
5	0,186
6	0,148

A more general shape factor is

$$G = 1 - r^2 + \varepsilon_1\, r^{n_1} \sin n_1\,\theta + \varepsilon_2\, r^{n_2} \sin n_2 \theta + \cdots \tag{3}$$

which leads to more complex shapes. Some instances of these shapes are shown in Fig. 1.

For the purposes of this presentation, a general shape factor in polar coordinates can be defined as

$$G = h_0(r) + \varepsilon_1 h_1(r, \theta) + \varepsilon_2 h_2\,(r, \theta) + \cdots \tag{4}$$

in which h_1, h_2 … are boundary perturbation functions. For the case of channel flow, the structure of (4) may be the same, in which polar coordinated may be substituted by Cartesian coordinates. The specific characteristics of functions h_i are determined by the nature of the equations of motion and associated boundary conditions.

Two relevant cases can be highlighted, namely, flow with one velocity component, and flow with more than one velocity component.

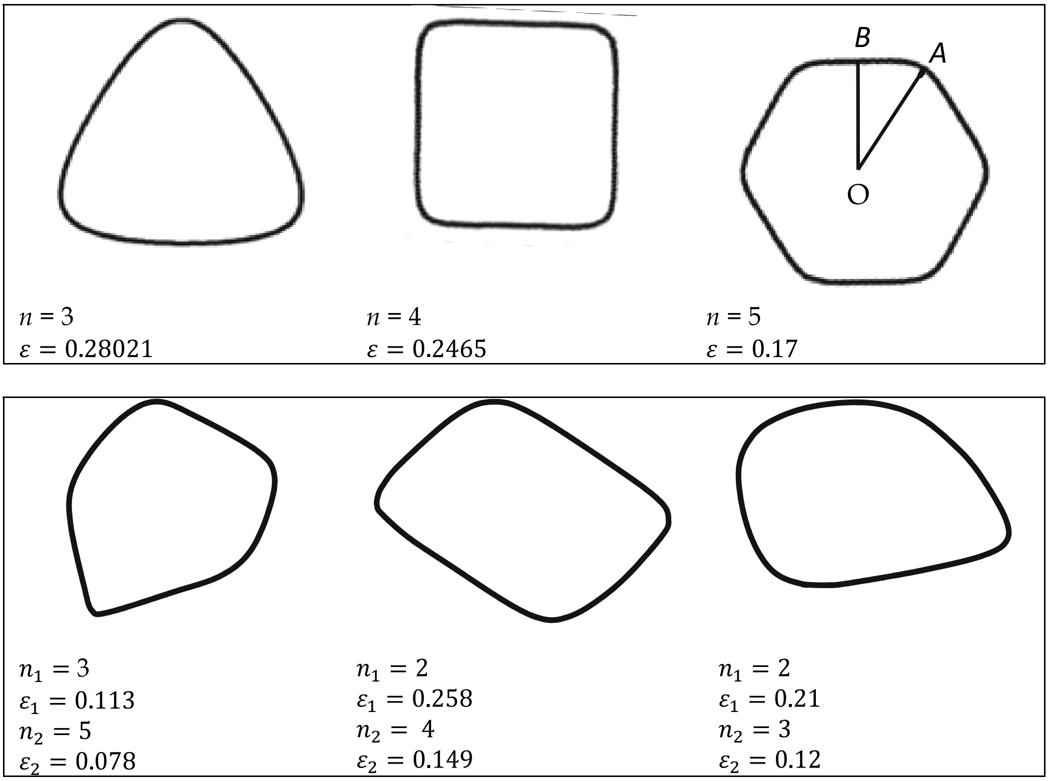

B
A
O
$n = 3$
$\varepsilon = 0.28021$
$n = 4$
$\varepsilon = 0.2465$
$n = 5$
$\varepsilon = 0.17$
$n_1 = 3$
$\varepsilon_1 = 0.113$
$n_2 = 5$
$\varepsilon_2 = 0.078$
$n_1 = 2$
$\varepsilon_1 = 0.258$
$n_2 = 4$
$\varepsilon_2 = 0.149$
$n_1 = 2$
$\varepsilon_1 = 0.21$
$n_2 = 3$
$\varepsilon_2 = 0.12$

Fig. 1. Samples of tubes contours from shape factors (1) and (3).

2.1 Flow with one velocity component

These are flows in straight tubes of constant cross-section. In these cases the axial velocity w can be modeled as

$$w(r,\theta,t) = G(r,\theta)[\,f_0(r,t) + \varepsilon\, f_1(r,\theta,t) + \varepsilon^2 f_2(r,\theta,t) + \cdots] \tag{5}$$

where, for the sake of simplicity, only one boundary perturbation function has been considered. Functions f_i are to be determined from the equation of motion in terms of a standard regular perturbation scheme around the small parameter ε.

2.2 Flows with more than one velocity component

These are mainly flows with axial variation of tube or channel geometry. In these cases the solution procedure will usually involve the use of a stream function Ψ. In such problems both Ψ and the velocity components should be zero at the boundary, a condition that can be met by defining

$$\Psi = G^2[g_0(r,t) + \varepsilon\, g_1(r,\theta,t) + \varepsilon^2 g_2(r,\theta,t) + \cdots] \tag{6}$$

where again functions g_1 have to be determined from the equations of motion. The definition of Ψ is given for every specific application in the corresponding section.

In the following some specific applications of this method of analysis are presented.

3. Applications

3.1 Flow in straight tubes of constant non-circular cross-section

3.1.1 Newtonian unsteady flow

For incompressible, developed and isothermal flow, the equation of motion are the standard Navier-Stokes and continuity equations. In dimensionless variables they are

$$\Omega\frac{\partial w}{\partial t}-\left(\frac{1}{r}\frac{\partial w}{\partial r}+\frac{\partial^2 w}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 w}{\partial \theta^2}\right)=\phi(t)=-\frac{\partial P}{\partial z}(t) \tag{7}$$

$$\frac{\partial w}{\partial z}=0 \tag{8}$$

In this

$$\Omega=\frac{\rho a^2}{\mu \mathrm{T}_0} \tag{9}$$

is the so-called unsteadiness number of the flow, which measures the relative importance of a temporal inertia force against a steady viscous force, and where ρ =density, α =reference tube radius, μ =dynamic viscosity, and T_0 =reference time.

A convenient solution of (7) (Letelier et al., 1995) for round tubes (ie for $\partial/\partial\theta=0$) can be worked out by postulating

$$w=A_2(1-r^2)+A_4(1-r^4)+A_6(1-r^6)+\cdots \tag{10}$$

Where $A_{2n}=A_{2n}(t)$ for $n=1{,}2{,}3\ldots\infty$. Equation (10) meets the no-slip boundary condition $w(1,t)=0$. After substituting (10) in (7) it is found that all functions A_{2n} can be expressed in terms of $A_2=A$ so that the axial velocity takes the form

$$w_0=(1-r^2)\left\{A+\frac{\Omega}{4^2}\frac{dA}{dt}(1+r^2)+\frac{\Omega^2}{4^2 6^2}\frac{d^2A}{dt^2}(1+r^2+r^4)+\cdots\right\} \tag{11}$$

where A is related to the forcing function $\phi(t)$ as follows

$$\frac{\phi}{4}=A+\frac{\Omega}{2^2}\frac{dA}{dt}+\frac{\Omega^2}{2^2 4^2}\frac{d^2A}{dt^2}+\cdots \tag{12}$$

In these expressions Ω can have any finite positive value.

According to (5), it is found

$$f_0=A+\frac{\Omega}{4^2}\frac{dA}{dt}(1+r^2)+\frac{\Omega^2}{4^2 6^2}\frac{d^2A}{dt^2}(1+r^2+r^4)\cdots \tag{13}$$

If

$$\mathrm{L}=\Omega\frac{\partial}{\partial \mathrm{t}}-\left(\frac{1}{\mathrm{r}}\frac{\partial}{\partial \mathrm{r}}+\frac{\partial^2}{\partial \mathrm{r}^2}+\frac{1}{\mathrm{r}^2}\frac{\partial^2}{\partial \theta^2}\right) \tag{14}$$

then, by collecting terms of order ε, it follows

$$L\{f_1(1-r^2)\}=L\{f_0 r^n \sin(n\theta)\} \tag{15}$$

wherefrom

$$f_1=\sin(n\theta)\sum_{i=1}^{\propto}\left\{\Omega^{\mathrm{i}}\frac{\mathrm{d^iA}}{\mathrm{dt^i}}\right\}\sum_{s=0}^{2i-2}\left\{C_{is}r^{i+s}\right\} \tag{16}$$

The constants C_{is} in these equations are obtained by putting the coefficients of all powers of r, for any $d^iA/dt^i (i = 1,2, \dots)$,equal to zero in (13). The result is

$$C_{10} = \frac{n-3}{16(n+1)} \tag{17}$$

$$C_{22} = \frac{72C_{10}+n-2.5}{576(n+2)} \tag{18}$$

$$C_{20} = \frac{(n+1)(576C_{22}+1)-144C_{10}-9}{576(n+1)} \tag{19}$$

and so on. Higher order terms in ε can be obtained in like fashion.
An example of velocity profiles is shown in figure 1 for $n = 6$, $\varepsilon = 0.148$ at two semi-axes (cf Fig. 1).
In this case the tube contour is an approximate hexagon and $\phi = cost$, ie a purely oscillatory flow is described.

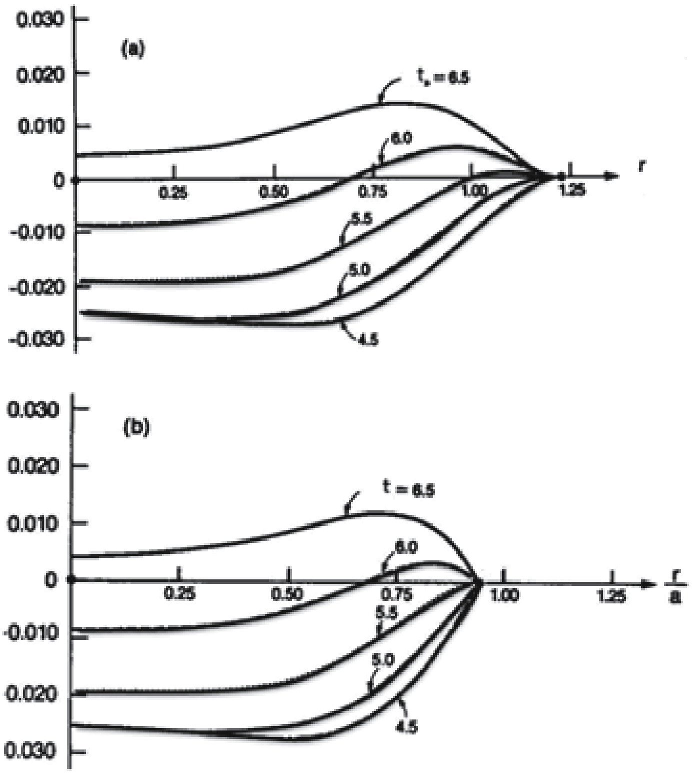


Fig. 2. Instantaneous velocity profiles for Ω = 40. (a) along maximum semi-axis OA; (b) along minimum semi-axis OB.

The structure of (4) makes it possible to apply a regular perturbation method of solution around the dimensionless parameter ε. Since ε is bounded for a given value of n, and is always less than unity, the solution becomes actually an exact one when enough terms are obtained.

3.1.2 Steady plastic flow

In this application it is considered steady flow of a Bingham plastic. Here (5) is also applicable, and the equation of motion, in terms of shear stress, as defined below, is

$$\frac{\tau_{rz}}{r} + \frac{\partial \tau_{rz}}{\partial r} + \frac{1}{r}\frac{\partial \tau_{\theta z}}{\partial \theta} = \phi \tag{20}$$

$$\tau_{rz} = -\left(1 - \frac{N}{I}\right)\frac{\partial w}{\partial r} \tag{21}$$

$$\tau_{\theta z} = -\left(1 - \frac{N}{I}\right)\frac{1}{r}\frac{\partial w}{\partial \theta} \tag{22}$$

and

$$I = \sqrt{\left(\frac{\partial w}{\partial r}\right)^2 + \left(\frac{1}{r}\frac{\partial w}{\partial \theta}\right)^2} \tag{23}$$

is the second invariant of the rate of deformation tensor. The dimensionless yield stress is

$$N = \frac{\tau_0 a}{w_0 \eta_0} \tag{24}$$

The momentum equation (20) is the standard one for parallel steady flow. Its structure has been made consistent with (21-22) and with the standard mathematical ordering of terms. The constitutive expressions (21-22) come from the applicable form of the Bingham fluid model. Defining

$$w = w_0 + \varepsilon\, w_1 + \cdots \tag{25}$$

then, from (5) it follows

$$w_0 = (1-r^2) f_0 \tag{26}$$

$$w_1 = (1 - r^2)\, f_1 + r^n\, f_0 \, sin\, n\theta \tag{27}$$

and the following equations are found

$$\tau_{rz} = N - \frac{\partial w_0(r)}{\partial r} - \varepsilon \frac{\partial w_1(r,\theta)}{\partial r} \tag{28}$$

$$\tau_{\theta z} = \left(\frac{N}{\frac{\partial w_0(r)}{\partial r}} - 1\right)\frac{\varepsilon}{r}\frac{\partial w_1(r,\theta)}{\partial \theta} \tag{29}$$

$$-\frac{1}{r}\frac{\partial w_0(r)}{\partial r} - \frac{\partial^2 w_0(r)}{\partial r^2} = \phi - \frac{N}{r} \tag{30}$$

$$\frac{1}{r}\frac{\partial w_1(r,\theta)}{\partial r} + \frac{\partial^2 w_1(r,\theta)}{\partial r^2} + \left(1 - \frac{N}{\frac{\partial w_0(r)}{\partial r}}\right)\frac{1}{r^2}\frac{\partial^2 w_1(r,\theta)}{\partial \theta^2} = 0 \tag{31}$$

Equations (28-29) are the result of substituting (25) in (23) and of ordering terms in powers of ε through a linearization procedure. From (30-31) it is found

$$w_0(r) = N(r - 1) + \frac{\phi}{4}(1 - r^2) \tag{32}$$

$$w_1(r,\theta) = A_0\left(1 + n^2 \sum_{i=1}^{n} \frac{(-1)^i \phi^i (n^2-(i-1)^2)!}{2^i N^i (i^2)!} r^i\right) cos(n\theta) \tag{33}$$

where

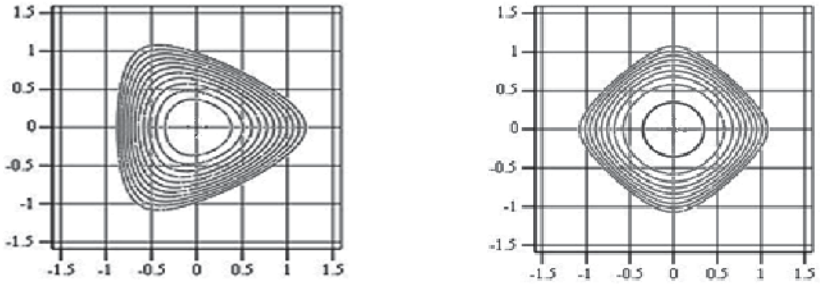
$$A_0 = \frac{2^n[n^2!]N^n}{(-1)^n \phi^n\left[n^2\left(n^2-1\ \right)(n^2-2^2)\ldots(n^2-(n-1)^2)\right]} \tag{34}$$

Functions $f_0, f,$ and following can be found equating terms in orders of ε in (5), ie

$$f_0 = \frac{w_0}{1-r^2} \tag{35}$$

$$f_1 = \frac{w_1 - r^n\, f_0\, \sin n\,\theta}{1-r^2} \tag{36}$$

In this both functions are continuous for $r = 1,$ and so can be built higher order functions. Isovel plots and plug zones for selected instances of flow are shown in figure 3. A plug zone is such that inside its limiting boundary the shear stress is less than the yield stress.



(a) (b)

Fig. 3. Isovels from (25) for (a) $n = 3,\ \varepsilon = 0.3,\ N = 0.2$ and $\phi = 4$; (b) $n = 4,\ \varepsilon = 0.24;$ $N = 0.7$ and $\phi = 4.$

According to fig.3, in both cases therein depicted, the plug zone appears at the center and is essentially circular.

3.2 Axially-varying flows in conduits

3.2.1 Newtonian flow in round tubes of arbitrarily axially-varying cross-section

A definition diagram is shown in fig.4

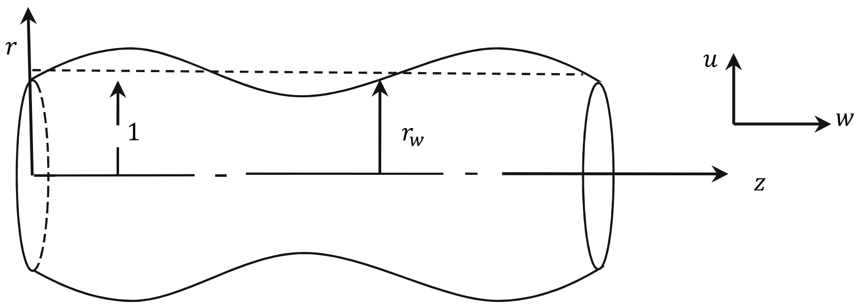


Fig. 4. Definitions diagram for flow in round tubes of axially-varying cross-section.

For this flow, the continuity and Navier-Stokes equations are

$$\frac{1}{r}\frac{\partial(ru)}{\partial r}+\frac{\partial w}{\partial z}=0 \tag{37}$$

$$u\frac{\partial u}{\partial r}+w\frac{\partial u}{\partial z}=-\frac{2}{Re}\frac{\partial P}{\partial r}+\frac{2}{Re}\left(\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial(ru)}{\partial r}\right)+\frac{\partial^2 u}{\partial z^2}\right) \tag{38}$$

$$u\frac{\partial w}{\partial r}+w\frac{\partial w}{\partial z}=-\frac{2}{Re}\frac{\partial P}{\partial z}+\frac{2}{Re}\left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial w}{\partial r}\right)+\frac{\partial^2 w}{\partial z^2}\right) \tag{39}$$

where Re is the Reynolds number. The velocity field is found after defining

$$u=\varepsilon u_1(r,z)\ldots \tag{40}$$

$$w=w_0(r)+\varepsilon w_1(r,z)\ldots \tag{41}$$

$$P=P_0(z)+\varepsilon P_1(r,z)\ldots \tag{42}$$

$$\Psi=\Psi_0(r)+\varepsilon\Psi_1(r,z)\ldots \tag{43}$$

The velocity is next expressed in terms of de stream function as follows

$$w=\frac{1}{r}\frac{\partial w}{\partial r} \tag{44}$$

$$u=-\frac{1}{r}\frac{\partial\Psi}{\partial z} \tag{45}$$

The wall radius is a function of the axial coordinate, that is here defined as

$$r_w=1+\varepsilon h(r)F(z) \tag{46}$$

in which ε is a small parameter. This algebraic structure allows to introduce a very large range of axial variation since h and F are arbitrary functions. Three cases will be considered. In the first case

$$h=r; \qquad F(z)=\sin\omega z \tag{47}$$

where is ω is an arbitrary frequency.
The stream function is modelled from (6) and (46), ie

$$\Psi=(r_w-r)^2[g_0(r)+\varepsilon\{g_{11}(r)\sin(\omega z)+g_{12}(r)\cos(\omega z)\}+0(\varepsilon^2)+\cdots] \tag{48}$$

wherefrom

$$\Psi_0=(1-r)^2 g_0(r)=\frac{\phi}{16}(r^2-1)^2 \tag{49}$$

and thus

$$g_0(r)=\frac{\phi}{16} \tag{50}$$

The first order stream funtion is

$$\Psi_1=(1-r)^2\big(g_{11}(r)sin(\omega z)+g_{22}(r)cos(\omega z)\big)+2(1-r)\,r\,sin(\omega z)\frac{\phi}{16} \tag{51}$$

which is next written as

$$\Psi_1 = \mathrm{H}_1(\mathrm{r}) * \sin(\omega \mathrm{z}) + \mathrm{H}_2(\mathrm{r}) * \cos(\omega \mathrm{z}) \tag{52}$$

and where H_1 and H_2 are unknowns that are modelled as finite polynominals of even order terms, ie

$$H_1(r) = a_0 + a_2r^2 + a_4r^4 + a_6r^6 + a_8r^8 + a_{10}r^{10} + a_{12}r^{12} + a_{14}r^{14} + a_{16}r^{16} \tag{53}$$

$$H_2(r) = b_0 + b_2r^2 + b_4r^4 + b_6r^6 + b_8r^8 + b_{10}r^{10} + b_{12}r^{12} + b_{14}r^{14} + b_{16}r^{16} \tag{54}$$

The coefficients a_i and b_j are determined by substituting (53) and (54) in the equation for Ψ found from (38-39) once (44-45) are substituted in there. Examples of typical streamline and isovelocity patterns are shown in figures 5 and 6. Streamlines are plotted from (48) and isovelocity are curves where $u^2 + v^2$ is constant.

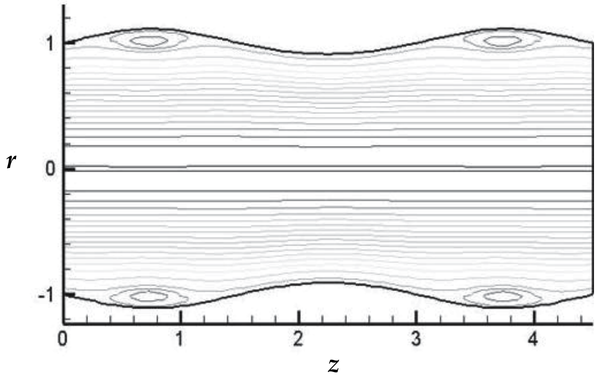


Fig. 5. Streamlines for Re=100 and ε =0.1.

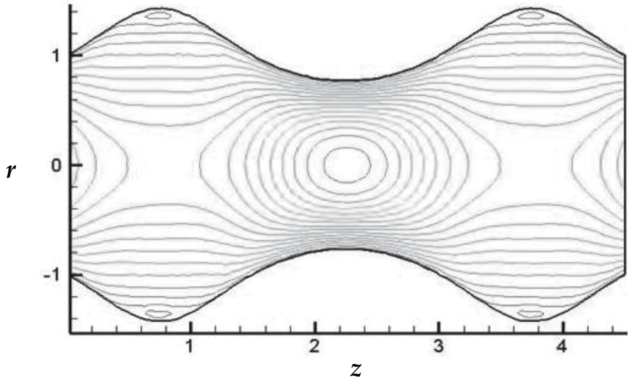


Fig. 6. Isovelocity lines for Re=1 and ε =0.3 from (48) and (44-45).

A second kind of contour is defined through the expression

$$F(z) = -0.028z^4 + 0.434z^3 - 2.156z^2 + 3.43z \tag{55}$$

which was transformed in a Fourier series in the range $0 \leq z \leq 3$ in terms of sine and cosine functions that allow a modeling similar, but more complex, to that already described. Examples of typical streamline and isovelocity patterns are shown in figure 7 and 8.

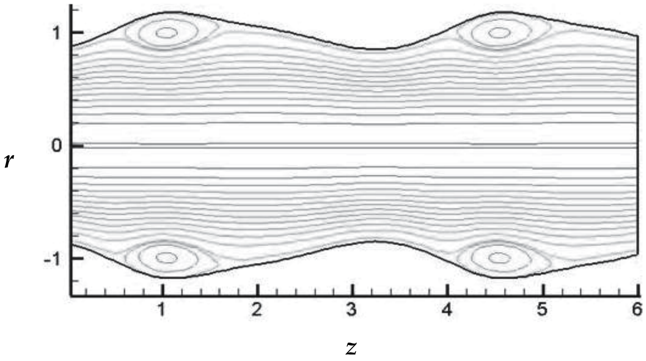


Fig. 7. Streamlines for Re=100 and ε =0.2.

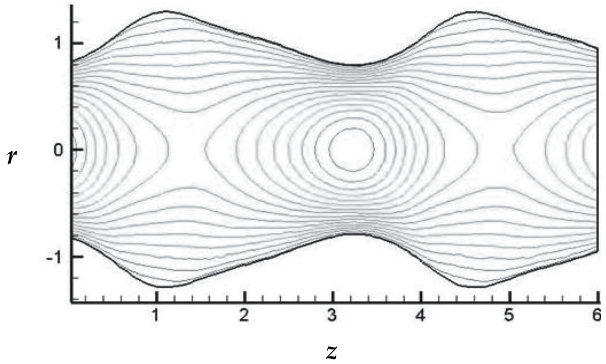


Fig. 8. Isovelocity lines for Re=1 and ε =0.3.

In similar fashion, the a third contour presented is defined by

$$F(z) = -0.03z^4 - 0.045z^3 + 0.405z^2 + 0.42z \tag{56}$$

For this case, typical isovelocity and isobaric curves are shown in figures 9 and 10.

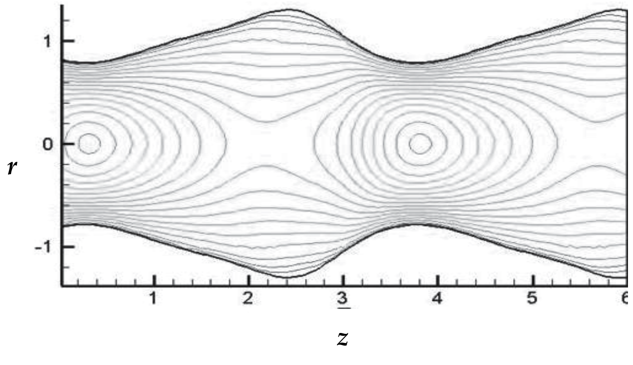


Fig. 9. Isovelocity lines for Re=1 and ε =0.3.

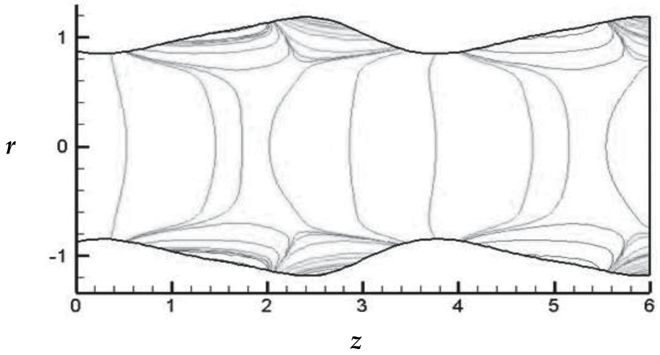


Fig. 10. Isobaric lines for Re=100 and ε =0.2 from (42).

3.2.2 Steady plastic flow in undulating channels

A definition diagram for this flow is shown in fig. 11.

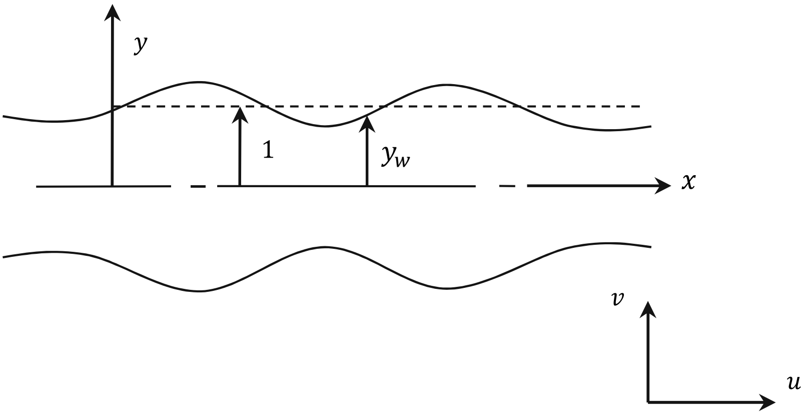


Fig. 11. Definition diagram for flow in undulating channels.

The Bingham constitutive equations are for this case are

$$\tau_{xx} = -\left(1 - \frac{N}{I}\right) 2 \left(\frac{\partial u}{\partial x}\right) \tag{57}$$

$$\tau_{yy} = -\left(1 - \frac{N}{I}\right) 2 \left(\frac{\partial v}{\partial y}\right) \tag{58}$$

$$\tau_{yx} = \tau_{xy} = -\left(1 - \frac{N}{I}\right) \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) \tag{59}$$

and the corresponding continuity and momentum equations are

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \tag{60}$$

$$u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=-\frac{1}{Re}\left(\frac{\partial \tau_{xx}}{\partial x}+\frac{\partial \tau_{xy}}{\partial y}\right)-\frac{1}{Re}\frac{\partial P}{\partial x} \tag{61}$$

$$u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}=-\frac{1}{Re}\left(\frac{\partial \tau_{xy}}{\partial x}+\frac{\partial \tau_{yy}}{\partial y}\right)-\frac{1}{Re}\frac{\partial P}{\partial y} \tag{62}$$

Next, the velocity, pressure and stream function are expanded as

$$u(x,y)=u_0(y)+\varepsilon u_1(x,y)\ldots \tag{63}$$

$$v(x,y)=\varepsilon v_1(x,y)\ldots \tag{64}$$

$$P(x,y)=P_0(x)+\varepsilon P_1(x,y)\ldots \tag{65}$$

$$\psi(x,y)=\psi_0(y)+\varepsilon\psi_1(x,y)\ldots \tag{66}$$

$$u(x,y)=\frac{\partial}{\partial y}\psi(x,y) \tag{67}$$

$$v(x,y)=-\frac{\partial}{\partial x}\psi(x,y) \tag{68}$$

$$\psi=(y-y_w)(g_0(y)+\varepsilon g_1(x,y)+\varepsilon^2 g_2(x,y)+\cdots) \tag{69}$$

where the wall is described by

$$y_w=1+\varepsilon F(y)sin(\omega x) \tag{70}$$

Following a procedure similar to the one presented in section 3.2.1, it is found

$$\psi_0=(y-1)^2 g_0(y) \tag{71}$$

$$g_0(y)=-\frac{1}{6}(2y+4-3N) \tag{72}$$

$$\psi_1=(1-y)^2\big(g_1(y)sin(\omega x)+g_2(y)cos(\omega x)\big)+2(1-y)F(y)sin(\omega z)f_0(y) \tag{73}$$

$F(y)$ is defined as $F=y$ and the first order stream function is modeled as

$$\psi_1=A(y)sin(\omega x)+B(y)cos(\omega x) \tag{74}$$

The unknown functions $A(y)$ and $B(y)$ are modeled as finite polynomial, ie

$$A(y)=a_0+a_1y+a_2y^2+a_3y^3+a_4y^4+a_5y^5+a_6y^6+a_7y^7+a_8y^8+a_9y^9 \tag{75}$$

$$B(y)=b_0+b_1y+b_2y^2+b_3y^3+b_4y^4+b_5y^5+b_6y^6+b_7y^7+b_8y^8+b_9y^9 \tag{76}$$

The coefficients a_i and b_j are determined by substituting (75) and (76) in the equations for ψ found from the momentum equations. In the following figures are presented plots of streamlines (equation (69)), isovelocity lines ($u^2+v^2=const.$), plug zones and axial velocity profiles

In figure 13 the plug zones are shown as shaded areas, which were determined by putting the condition that the shear stress should be equal or less the yield stress. The quasi-plug zones are zones where only $\tau_{rz}\leq 0$.

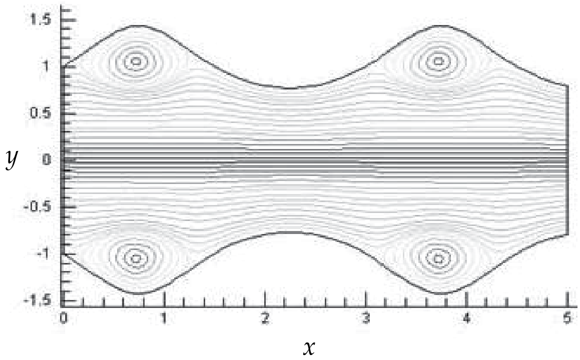


Fig. 12. Streamlines for Re=100, N=0.3 and ε = 0.3.

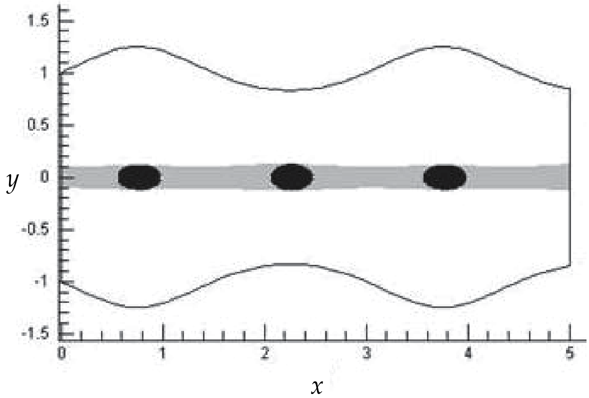


Fig. 13.Plug and quasi-plug zones for *Re*=1, *N*=02 and *ε*=0.2.

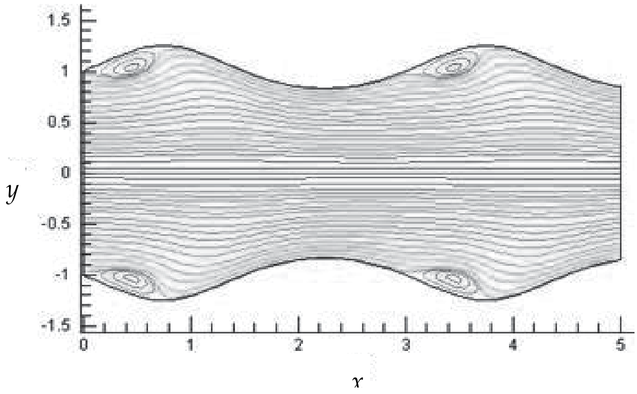


Fig. 14. Streamlines for *Re*=20, *N*=0.1 and ε=0.2.

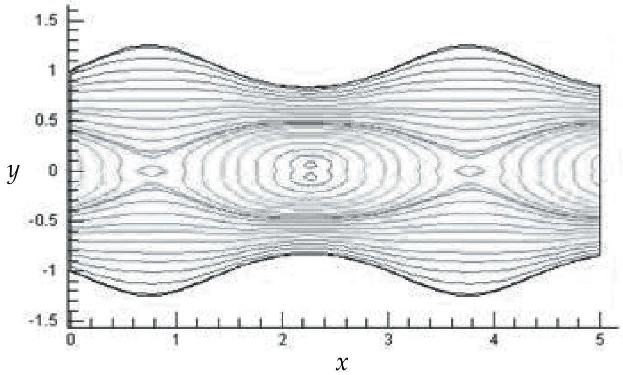


Fig. 15. Isovelocity lines for *Re*=1, *N*=0.1 and ε=0.2.

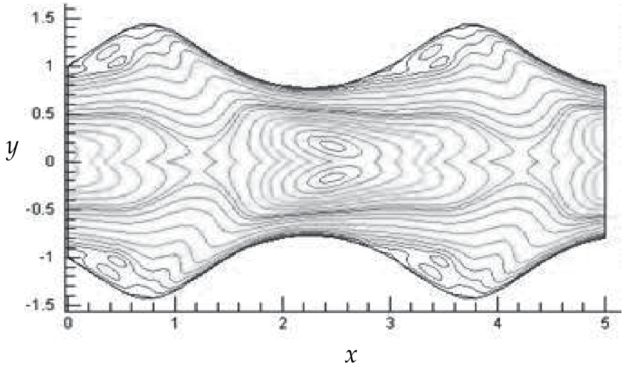


Fig. 16. Isovelocity lines for *Re*=20, *N*=0.2 and ε=0.3.

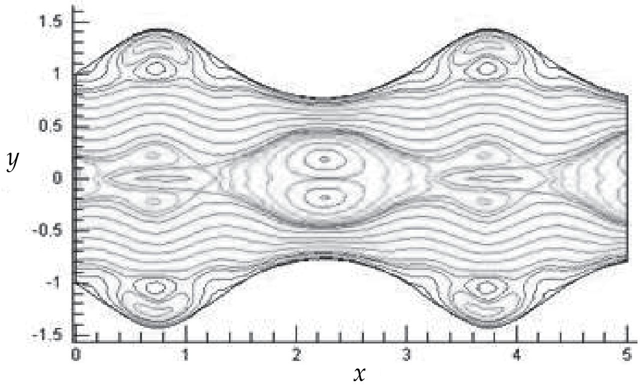


Fig. 17. Isovelocity lines for *Re*=100, *N*=0.1 and ε=0.3.

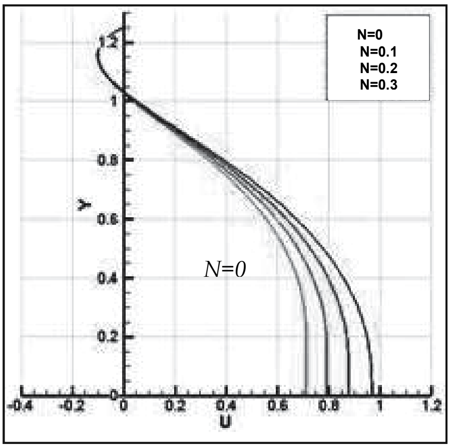


Fig. 18. Axial velocity profiles at x=0.75 for $\varepsilon = 0.2$ and Re =100.

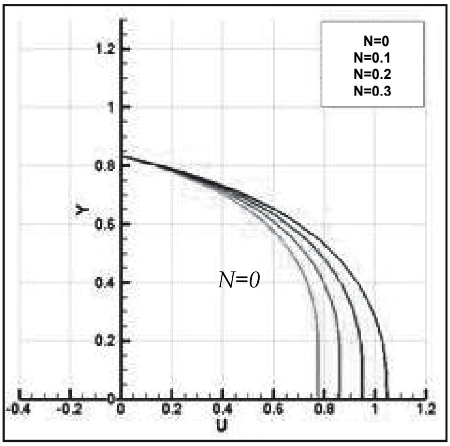


Fig. 19. Axial velocity profiles at x=2.25 for $\varepsilon = 0.2$ and Re =100.

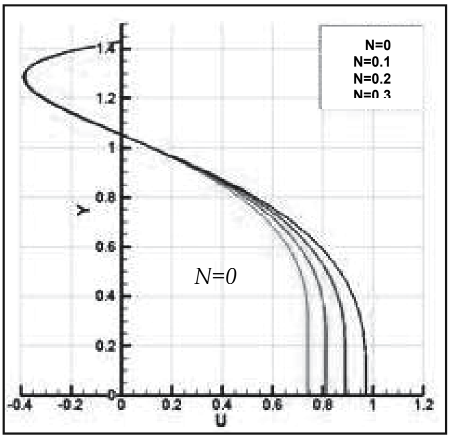


Fig. 20. Axial velocity profiles at x=0.75 for $\varepsilon = 0.3$ and Re =100.

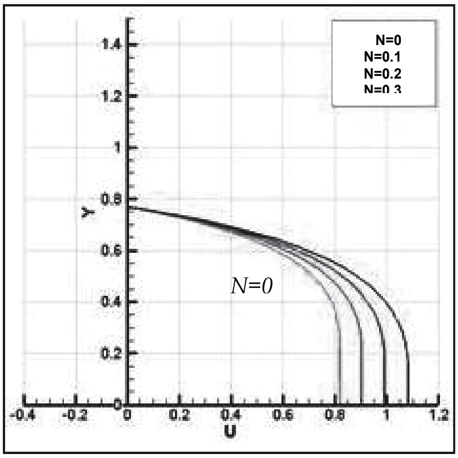


Fig. 21. Axial velocity profiles at x=2.25 for $\varepsilon = 0.3$ and Re =100.

4. Conclusion

The method here described can lead to very accurate solutions for the velocity field and related variables such as shear stress, rate of flow and pressure in a great variety of flows in tubes and channels. Symbolic software presently available, such Maple and MathCAD make it possible to obtain and compute higher order solutions that, in some cases, may have complex algebraic structures. The fact that for all cases here considered, ie cases where $n \geq 3$, ε is much less than unity (cf table 2), leads to a regular perturbation scheme that in most cases requires terms up to second order to achieve enough accuracy. The cases when $n = 1$ and $n = 2$ deserve special mention. For $n = 1$ the shape factor (1) describes an excentric circle, and for $n = 2$ an ellipse. In this last instance ε is not bounded and can take any finite value, which implies that the perturbation scheme would break down if $\varepsilon \geq 1$. So that, in this particular case, the method is limited to elliptical cross-sections of axes ratio close to unity. The method can be expanded to many more complex flow geometries. This possibility is implicit in the more general shape factor (3), which makes it necessary to develop a compound perturbation scheme, in terms of more than one perturbation parameter. The structure of the shape factor (1) determines that the analysis, especially for $n \geq 3$, is more sensible to the perturbation parameter for $r \geq 1$, ie close to the wall conduit. This requires a careful analysis of series convergency which should define the order of the higher order term considered. On the other hand, in the case of flow in straight tubes, in all cases studied, in a considerable region around the conduit axis, say for $r \leq 0.4$, the flow variables are independent of the boundary geometry and take the values of the corresponding flow in round tubes.

5. Acknowledgment

The authors acknowledge the financial support provided, at different stages of the work here presented, by FONDECYT-CONICYT and DICYT at the University of Santiago of Chile.

6. References

Batchelor G. K. (2000). *An Introduction to Fluid Dynamics*. Cambridge University Press.

Cetin B. and Li D. (2008). *Microfluidic continuous particle separation via ac- dielectrophoresis with 3d electrodes*. ASME International Mechanical Engineering Congress and Exposition.

Chen P.CH.; Wang H.; Park D.S.; Park S.; Nikitopoutos D.E.; Soper S. A. and Murphy M. C. (2008). *Protein adsorption in a continuous flow micro-channel environment*. ASME International Mechanical Engineering Congress and Exposition.

Forte J.A.; Sipahi R. and Ozturk A.(2008). *A novel device for nonmagnetic particle navigation using ferrofluids manipulated by magnetic fields*. ASME International Mechanical Engineering Congress and Exposition.

Gebauer P. and Bocek P.(2002). *Recent progress in capillary isotachophoresis,* pp. 3858-3864. Electrophoresis 23.

Letelier M.F. and Siginer D.A. (2003). *Secondary Flows of Viscoelastic Liquids in Straight Tubes*. pp. 5081-5095. International Journal of Solids and Structures, N°40.

Letelier, M.F. and Siginer, D.A. (2007). *On the Flow of a Class of Viscoinelastic-Viscoplastic Fluids in Tubes of Non-circular Contour,* pp. 873 – 881. International Journal of Engineering Science, Vol. 45, Issue 11.

Letelier, M.F.; Leutheusser, H.J.; Member, ASCE, and Marquez Z., (1995). *Laminar fluid transients in conduits of arbitrary cross section,* pp.1069-1074. Journal of Engineering Mechanics.

Letelier, M.F.; Siginer D.A. and Cáceres, C.G. (2002). *Pulsating Viscoelastic Flow in Tubes of Arbitrary Cross-Secction, Part I: Longitudinal Field*. International Journal of Non-Linear Mechanics.

Letelier, M.F.; Siginer, D.A.;. Stockle, J.S and Huilcan A. P. (2010). *Laminar flow in tubes of circular cross section with arbitrary axial variation. ASME* International Mechanical Engineering Congress and Exposition..

Mathies R.A. and Huang X.C.(1992). *Capillary array electrophoresis: an approach to high-speed, high-throughput DNA sequencing* pp. 167-169. Nature 359.

Pedley T. J. (2008). *The Fluid Mechanics of Large Blood Vessels*. Cambridge University Press.

Sommer G.J.; Chang D. S.; Jain A.; Langelier S. M.; Park S.; Rhee M.; Wang F.; Zeitun R. I. and Burns M. A. (2008). *Introduction to microfluidics,* pp.1- 33. In microfluidics for biological applications. W-C, Tian and E. Fine hout, editor, Springer.

Srivastava N.; Davenport RD.; Burns MA. (2005). *Nanoliter viscometer for analyzing blood plasma and other liquid samples,* pp. 383 - 392. Anal Chem. 77.

Stevenson, B. (2010). *Análisis del campo de velocidades y de las condiciones de existencia y características de la zona tapón en flujos plásticos en ductos de secciones arbitrarias.* Mechanical Engineering thesis, Master degree, University of Santiago of Chile.

Svensson, P.J. (2010). *Análisis de la zona tapón en flujo plástico en canales ondulados.* Mechanical Engineering thesis, University of Santiago of Chile.

Woolley A. and Mathies R. (1994). *Ultra-high-speed and fragment separations using microfabricated capillary array electrophoresis chips,* pp. 11348 - 11352. Proceedings of the National Academy of Sciences 91.

Xue, S.C.; Phan-Thien N. and Tanner, R.I. (1995). *Numerical study of secondary flows of viscoelastic fluid in straight pipes by an implicit finite volume method,* pp. 191-213. Journal of Non-Newtonian Fluid Mechanics, 59.

Yager P.; Edwards T.; Fu E.; Helton K.; Nelson K.; Tam MR.; Weigl BH.(2006) *Microfluidic diagnostic technologies for global public health,* pp.412 - 418. Nature 442.

20

CSA – Clinical Stress Assessment

Sepp Porta et al.*
[1]Institute of Applied Stress Research, Judendorf – Strassengel,
[2]Institute of Pathophysiology, Medical University of Graz,
[3]Institute of Mathematics and Scientific Computing, KFU Graz,
[4]Rehabilitation Clinique of the AUVA, Tobelbad,
[5]Theresianische Militärakademie, Wiener Neustadt,
[6]St. Anna Hospital, Herne,
[1,2,3,4,5]Austria
[6]Germany

1. Introduction

1.1 Theoretical background – Outlook to applications

What is stress?

Although stress may be defined seemingly differently by endocrinologists, physiologists, psychologists, a.s.o. prefering the tools of their own, specialized trades, the general denominator and most important fact to remember is, that stress never ever impacts somehow upon you or threatens you from somewhere. Stress is always and only your individual typical reaction to something beginning with a menace and ending perhaps with a slight challenge typical for everyday life. Your personal reaction to such provocations is called stress. The provocation itself is not at all a stress but is called "stressor". Those two technicalities are regrettably often confused by journalists, so that the word "stress" became an exceedingly wooly term.

But if we once agree upon the reactive nature of stress, further reasoning is simple: If your efforts to remove a provocation turn out to be too feeble, the provocation remains and your unsuccessful efforts become chronic – chronic stress ensues. Although perhaps a bit too feeble to remove the provocation, your stress efforts are still using up more than the portion of energy which you have allotted to the routine running of events. Thus your whole system needs more fuel over a longer time, which in its turn tends to exhaust your energy reserves. Your efforts grow feebler still – burnout threatens.

---

* Gertrud W. Desch, Harald Gell, Karl Pichlkastner, Reinhard Slanic, Josef Porta, Gerd Korisek, Martin Ecker and Klaus Kisters
1 Institute of Applied Stress Research, Judendorf – Strassengel, Austria,
2 Institute of Pathophysiology, Medical University of Graz, Austria,
3 Institute of Mathematics and Scientific Computing, KFU Graz, Austria,
4 Rehabilitation Clinique of the AUVA, Tobelbad, Austria,
5 Theresianische Militärakademie, Wiener Neustadt, Austria,
6 St. Anna Hospital, Herne, Germany.

It is interesting, that such reasoning describes on the one hand the well known development of chronically increased energy turnover into exhaustion and burnout. On the other hand it shows, that increasing exhaustion also increasingly curtails successful reactions to immanent provocations, meaning that those provocations cannot be fought with adequate reactions any more – an exhausted subject cannot mobilize enough reserves to fend off a challenge – there is not enough stress available to cope successfully. Thus one could appreciate the nonsense of statements repeated in journals over and over again, to "dismantle your stress" or "let your stress phase out". Far from getting rid of a personal reaction which may successfully release you from an impending menace, one has to fight the menace itself, which can be only done by successfully mobilizing ones reserves.

All those different reactions of the organism, due to differing workload intensities and different duration leading to the symptoms just described, can be quantified by a multiple parameter assessment called CSA.

Physiological aspects of stress and their utilization for stress assessment

Stress situations incite changes of Adrenaline and Noradrenaline which, in their turn effect variations in blood pH, CO2, O2, buffer parameters like BE or HCO3, lactate and blood glucose as well as electrolytes like K, Na, Ca and Mg. Since we could show as far back as 1991 (1, 2), that those stress hormone effects do correlate highly significantly with adrenaline and noradrenaline changes themselves, the tedious, costly and time consuming catecholamine determination by HPLC could – at least for the purposes with which we tend to deal in this chapter – be abandoned in favour of a much quicker method for determination of stress induced metabolic changes. Estimation of those metabolic effects has the additional benefit, that the obstacle of the receptor situation, which influences hormonal effects and thus restricts the meaningfulness of catecholamine determination as a tool for assessing stress intensity is avoided, since all of our parameters depict post receptor effects.

Those non hormonal parameters therefore show interdependent stress hormone effects, which, when determined simultaneously, can be laid over an organism like a data net by especially designed online software which can be taken as the basis of an individually adaptable multi parameter stress index.

In the simplest case of a physical workload e.g., we find workload dependent change in blood glucose, increase in lactate, accompanied by adaptive changes in pCO2 and/or HCO3 and baseexcess, softening the lactate impact upon pH. Moreover, typical shifts in Ca, Mg and K provide us with information about the intensity and duration of stress. When e.g. more sensitive parameters like pCO2 or less sensitive ones like HCO3 and blood glucose, are determined at the same time in the same sample, they can give a good idea about the duration of the stress, depending upon the relative involvement of the said parameters.

Likewise, change of electrolytes can tell – together with either pH or pCO2 about the momentary inclination to sportive performance, about the intensity of symapthoadrenal expectation situations or even about the individually felt efforts of competition and – in the long run – even help to diagnose and quantify a possible state of exhaustion. Chronic sympathoadrenal impact upon metabolism could also be detectable in psychopathological diseases like depressive disorders, where the question arises, whether the patients' mental exhaustion may not affect metabolic processes too.

Also, those interconnections of multiple metabolic effects can be used to uncover hitherto less well understood parameter interactions in metabolic diseases like the metabolic syndrome or even diabetes. There the quantity of electrolyte deficiency and its relation to the idiosyncratic behaviour of a chronically affected metabolism may open new aspects of diagnosis and therapy. Finally, diagnosis of mental and physical load leading to exhaustive stress can not only be used in managers and sports persons, but also to the purpose of being better able to judge upon correct treatment of livestock.

In most cases one has to take pains to collect pre- and post workload data.

The hardware used for such an assessment consists of well established determination systems, implemented in most ICUs all over the world. Due to the easy transportability, at least of those two examples shown below in fig.1 and fig.2., they have been used in assessment campaigns, ranging from determination of psychical workload of teachers in schools or managers in industrial plants to the evaluation of fitness of professional ski racing teams in mountain ranges and assessment of the impact of sleep deprivation in soldiers far away from human habitations.

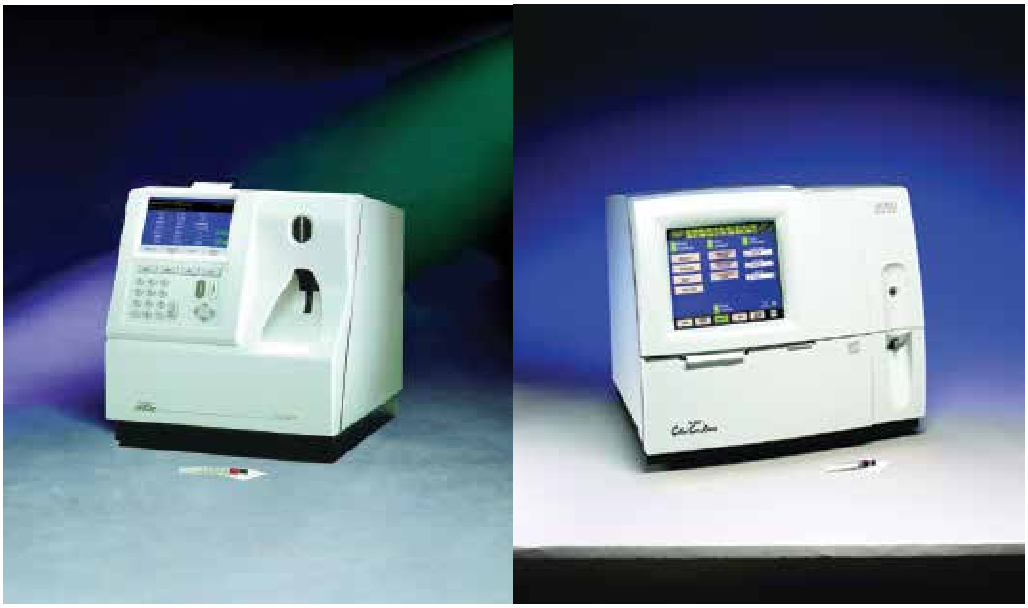


Fig. 1. NOVA CCX (Critical Care Express)

Fig. 2. NOVA Phox - M

Fig. 1. and 2. Two types of transportable ICU analyzers (dimensions estimable by the syringe in the foreground)

1.2 Practical implementation

Sampling and sample determination – Single persons

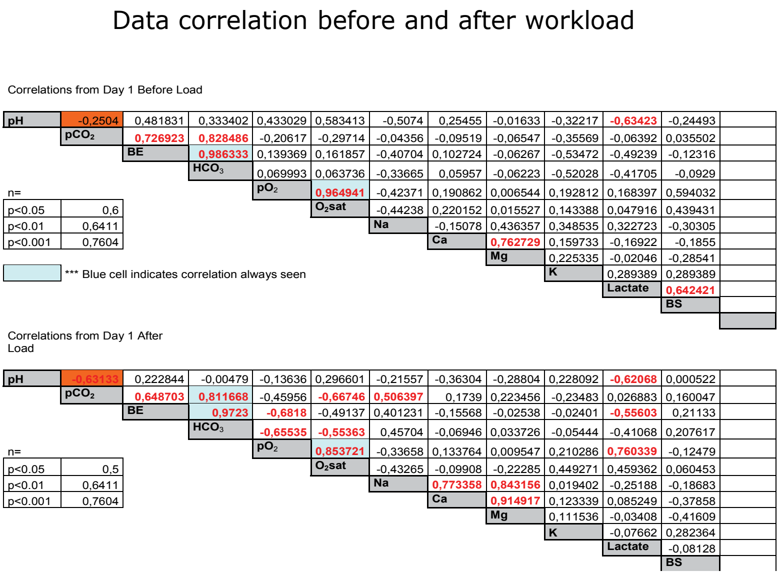
Thus, a persons' or an animals' workload, stress compatibility, duration of stress and also the intensity and the kind of stress can be determined within 3 minutes by collecting about 100 microliters of capillary blood, usually from the finger tip. The sample is routinely analyzed for pH, pCO2, pO2, O2saturation, ionized magnesium, ionized potassium, ionized calcium and ionized sodium, lactate, blood glucose, baseexcess and HCO3 (optionally hemoglobine and hematocrit) using a CCX (Crital Care Express, fig.1)) analyzer (NOVA Biomedical) or a Phox – M (fig.2) of the same producer (NOVA Biomedical), with about the same functions but smaller and even easier transportable and CSA (Clinical Stress Assessment) software. Both devices are widely applied all over the world in Intensive Care Units (ICUs), the software for online data evaluation and interpretation however has been developed by an Austrian corporation (PLK, Judendorf - Strassengel).

Healthy persons are usually checked before and after a standardized ergometric workload, mainly 80 Watts during 8 minutes, or before and after absolving sporting activities, or before and after a standardized psychical load like a Shapiro – test or any kind of training routine the impact of which stands in question. Additionally, the effects upon a persons' metabolism by so called wellness activities - from steam bath to sauna baths, massages etc. - can be investigated, applying the same protocol. In other fields of application even single determinations can be useful e.g. in the course of daily glucose profiles from diabetic patients in rehabilitation hospitals.

Sampling and sample determination – Groups

Such determinations can therefore characterize the reaction of a single person but they also can do the same with a whole group of people. In the latter case the group reaction may serve as a mirror of the typical demands of a certain task on a sample of persons. Thus information about e.g. the usefulness of a bout of training for defined purposes can be collected. All data won from the 100 microliter sample are analysed simultaneously within two minutes, so that total measuring time from blood sampling until the printout of the online processed data takes no more than 3 minutes. Calculations of averages, standard errors of means (SEM), delta values between the groups and linear correlations between all sampled parameters with their regression coefficients are software immanent. This means, that basic group statistics are available immediately after the testing of the last group member. The automatic correlation of every single parameter with all others comes in useful in many ways as we will see during the progress of the chapter. It allows us to look behind the equalizing group averages, thus enabling us to quantify the position of every group member from the point of view of a certain parameter combination. Moreover, the automatically emphasized numbers of significant regression coefficients in the correlation table creates a quickly recognizable pattern of typical group behaviour, as explained in fig. 3. Seen for the first time, this coefficient tables seem somehow crowded with data, but after a short acclimatisation one appreciates the quick oversight over all relevant interconnections of the measured parameters under different conditions:

In the first table (correlations day 1 before load) there are 7 significant linear correlations between stress related parameters. In the second table, describing the correlative situation of the same parameters and the same group, but this time after workload (military obstacle run – HIB), the number of correlations more than doubles to 15. Especially the increase of correlations with lactate and pCO2 after workload and the occasional difference in plus/minus signs are noteworthy. As we will endeavour to demonstrate, such correlative views can show at a glance, whether the present workload of a group still allows overcompensation (see below) or whether the group seems to be on the brink of exhaustion already. The usefulness of this tool in preventive medicine is obvious. The results of more than 2000 patients and experiments have been recorded in our data banks and published in about 60 papers and printed abstracts, thus providing comparative material for easier interpretation of results.



Data correlation before and after workload

Correlations from Day 1 Before Load

pH	-0,2504	0,481831	0,333402	0,433029	0,583413	-0,5074	0,25455	-0,01633	-0,32217	**-0,63423**	-0,24493		
	pCO_2	**0,726923**	**0,828486**	-0,20617	-0,29714	-0,04356	-0,09519	-0,06547	-0,35569	-0,06392	0,035502		
		BE	**0,986333**	0,139369	0,161857	-0,40704	0,102724	-0,06267	-0,53472	-0,49239	-0,12316		
			HCO_3	0,069993	0,063736	-0,33665	0,05957	-0,06223	-0,52028	-0,41705	-0,0929		
				pO_2	**0,964941**	-0,42371	0,190862	0,006544	0,192812	0,168397	0,594032		
					O_2sat	-0,44238	0,220152	0,015527	0,143388	0,047916	0,439431		
						Na	-0,15078	0,436357	0,348535	0,322723	-0,30305		
							Ca	**0,762729**	0,159733	-0,16922	-0,1855		
								Mg	0,225335	-0,02046	-0,28541		
									K	0,289389	0,289389		
										Lactate	**0,642421**		
											BS		

n=

$p<0.05$	0,6
$p<0.01$	0,6411
$p<0.001$	0,7604

*** Blue cell indicates correlation always seen

Correlations from Day 1 After Load

pH	**-0,63133**	0,222844	-0,00479	-0,13636	0,296601	-0,21557	-0,36304	-0,28804	0,228092	**-0,62068**	0,000522		
	pCO_2	**0,648703**	**0,811668**	-0,45956	**-0,66746**	**0,506397**	0,1739	0,223456	-0,23483	0,026883	0,160047		
		BE	**0,9723**	**-0,6818**	-0,49137	0,401231	-0,15568	-0,02538	-0,02401	**-0,55603**	0,21133		
			HCO_3	**-0,65535**	**-0,55363**	0,45704	-0,06946	0,033726	-0,05444	-0,41068	0,207617		
				pO_2	**0,853721**	-0,33658	0,133764	0,009547	0,210286	**0,760339**	-0,12479		
					O_2sat	-0,43265	-0,09908	-0,22285	0,449271	0,459362	0,060453		
						Na	**0,773358**	**0,843156**	0,019402	-0,25188	-0,18683		
							Ca	**0,914917**	0,123339	0,085249	-0,37858		
								Mg	0,111536	-0,03408	-0,41609		
									K	-0,07662	0,282364		
										Lactate	-0,08128		
											BS		

n=

$p<0.05$	0,5
$p<0.01$	0,6411
$p<0.001$	0,7604

Outprint of CSA – software immanent automatic correlations of all the parameters measured.
Upper triangle: before workload,
Lower triangle: after workload
Red numbers: significant correlation coefficients.

Fig. 3. Example of two tables of regression coefficients of a group before and after workload, useful for quick overall estimation of the changes of interparametric dynamics.

Ethical aspects

All participants in the investigations presented in this chapter consented to the anonymous use of their data after being carefully informed about the aim of the study according to the Helsinki Charter (*http://www.wma.net/e/ethicsunit/helsinki.htm.*). The most recent submission to and passing of our experimental intentions by the ethical commission of the Austrian Federal Ministry for Defence and Sports (BMfLVS) dated from Dec. 13th, 2010.

1.3 Application of CSA

Our own experience of CSA application ranges from investigations into the behaviour of teachers, students, triathletes, aircraft pilots, bungee jumpers, military combat groups, special police units, rescue teams, sleep deprivation, free radical research, patient's stress in rehabilitation clinics to the impact of wellness treatments, even to the exhaustion of farmers providing holidays for tourists additionally to their usual tasks.

The following four applications that will be discussed in the course of the chapter form a concentrated substrate, presented to make the reader think about own, customized applications:

1. Determination of the impact of sport training or training of so called "first responders", like military training units, special police groups, fire fighters and others. We maintain that it is not only possible to link changes of blood parameters with sportive success but also to predict success chances before competition or deployment.
2. Determination of the quantity of mental stress.
3. In the field of internal medicine mainly idiosyncrasies of diabetic metabolism, especially those due to the newly found importance of mineral deficiencies in type2 diabetics

which, by CSA application, we were able to link up with the deterioration of other metabolic disturbances.

4. Qualification and quantifications as well as predictions of success chances in competing animals like horses or camels and also prevention of cruelty to animals by stress documentations are one our next step of development.

ad 1: Determination of the impact of sport training or training of so called "first responders", like military training units, special police groups, fire fighters and others

Up to now the decision about a persons' fitness for a certain competition mainly rests upon the trainers' subjective adjudication, bolstered by lactate tests or even more demanding workout procedures and a more or less profound experience. Thus the availability of a hardly molesting, objective, in depth assessment of competitors and first responder personnel concerning their momentary ability to perform a certain task (3, 4, 5) comes in useful. Investigations in that area revealed not unexpectedly, that the metabolic situation of a person before a contest contributes decisively to the degree of the later success. Therefore it could be advantageous if we would be able to quantify the metabolic turnover before the contest in each case. Because, by quantifying pre contest metabolic situations, both the individual position within a group of contesters could be determined and eventually significant deviations of possible group outsiders could be marked down, understood and subsequently discussed with the person in question. That e.g. exhausted soldiers or sportsmen are no more able to perform satisfactorily is a truism. However, there are unsatisfactory performances which are less well explainable. The most common reasons are mostly privately known to the performer but not eagerly revealed to the trainer or group commander like lack of sleep due to entertainments during the previous evening.

But also unexpected bouts of good or bad performances occur, unexplainable to both performer and trainer. Moreover, the very same scores obtained easily by performer A could have been demanding for performer B, so that equal scoring may not mean equal potential at all. Even slight influences during the pre contest situation which are hardly felt and therefore frequently ignored, like temperature differences or even the changing of a routine can be reasons for a significant shift in interdependencies of electrolytes and metabolic parameters with measurable impact upon the performance to follow. Especially correlative changes between Mg, K, the Ca/Mg quotient or, to a lesser extent Ca alone with H^+ donors like lactate and the consecutive pH and blood buffer situation react rather sensibly to changes of the sympathoadrenal situation in man and even in horses (3,4,5,6). Since in our experience the demand for such determination focuses upon comparatively small groups and their individual members, we tried to offset the comparatively small number of experiments by software immanent statistics from diverse points of view to keep results controlled as strictly as possible.

Since our assessment provides us with at least 24 interconnected data per person (from determinations before and after workload), a more comprehensive description than that by the widely used lactate test or by catecholamine assay alone is possible.

It is remarkable, that nearly the same small amount of capillary blood which is still routinely used for lactate tests could easily yield eleven times more information instead of the single lactate determination. As an example we would like to show an investigation of 14 Ensigns of the Theresianische Militärakademie in Wiener Neustadt, Austria before and after a demanding military obstacle race of 3 – 5 minutes duration. It turned out that the metabolic changes in the experiments due to that military obstacle race were considerable, although the overall duration did not exceed five minutes:

Changes in group averages before and after the obstacle race (*military steeplechase, Hindernisbahn - HIB*) are shown in table 1: An expected significant change in stress dependent parameters like pH, pCO2, BE, HCO3 and K, due to severe exertions was visible. The averages of Mg and Ca however did not change significantly.

	pH	pCO_2	BE	HCO_3	pO_2	O_2sat	Na	Ca	Mg	K	Latate	BS
Averages before HIB	7,433	34,714	-1,114	23,400	68,293	94,164	145,886	1,135	0,514	4,339	1,793	102,357
+ / - SEM	0,005059	0,568578	0,419295	0,383162	1,5218	0,522737	0,429906	0,008238	0,008228	0,077598	0,11117	1,99892
Averages after HIB	7,257214	26,14286	-15,3643	12,01429	92,59286	96,29286	146,3714	1,140714	0,515	3,805	12,98462	151,2857
+ / -SEM	0,020643	0,53171	1,048002	0,72323	2,565461	0,215464	0,39103	0,009167	0,009931	0,093631	0,864019	6,807743

Average group values before and after sports +/- standard errors of means (SEM). Changes in pH, pCO2, BE, HCO3, pO2, O2sat., K, Ca, lactate and blood glucose are highly significant (two sided t- tests $(p<0,01)$, (7))

Table 1. CSA outprint, example of the change of group averages due to a military obstacle race (HIB).

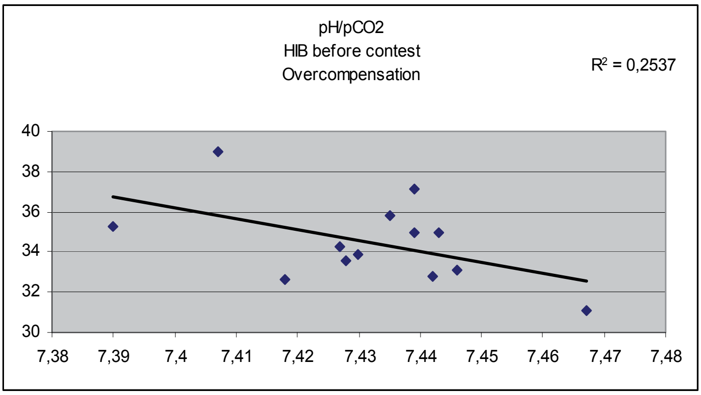
Beyond the information which the average group values provide, we were able to uncover four different facts by correlating appropriate parameters of the pre- contest situation.

a. The phenomenon of "overcompensation" in the pre contest phase can be detected and its effect upon the assessment of a persons' fitness and sportive compatibility can be discussed.
b. The impact of pre contest sympathoadrenal arousal upon the metabolism of each group member can be quantified. If outliers are present, they can be characterized and marked down for interviews
c. Connections between the now quantifiable position of each group member in the pre contest phase and the definite scoring during the contest can be established. Success/effort relationship can be characterized.

In other words, the answers to those four points provide the trainer with comprehensive information about the individual pre contest situation of each group member, whether the adrenergic arousal before contest remains within beneficiary boundaries or already uses up too much energy which would be more profitable employed later during the contest. They detect outsiders and even provide the trainer with scoring - chance predictions for the contest to follow. How is it done? First of all we have to become familiar with the usual state of affairs before a contest, the knowledge of which seems to be not usual at all. We call it "overcompensation". Different group members tend to express it in differing quantities.

a. Overcompensation

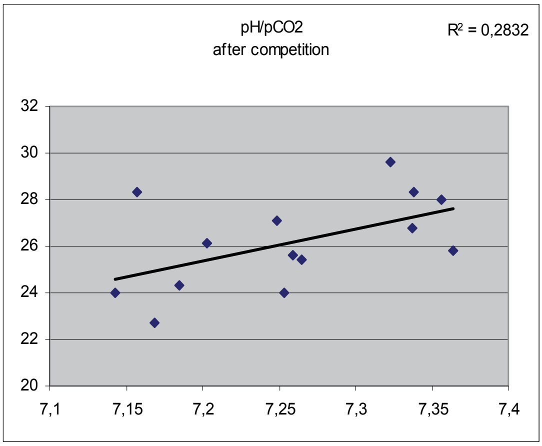
Slight to moderate mental or physical load frequently results in metabolic overcompensation. Simplified, the well known respiratory compensation of metabolic acidity can become over efficient, so that persons with increased lactate or other H^+ donors end up with a more alkaline pH brought about by a disproportionally increased breathing frequency, connected of course with an equally disproportional loss of CO2. The benefit of this seemingly wasteful behaviour is a kind of run up into higher pH regions to premeditate a later fall into dangerous acidity by an eventually more demanding workload in the immediate future which the organism seems to prudently forestall. Accumulation of O2 in the blood because of the more alkaline conditions points the same way. A good example of such a reaction is the pre contest situation before the military steeplechase concerning pH/pCO2 relationship. (fig. 4). This figure deals with and quantifies an "over successful" removal of acidity from the blood by increased breathing frequency.



Abscissa: pH
Ordinate: pCO2 in mmHg
P< 0,05, significant
Situation before the military steeplechase contest.

Fig. 4. Situation before contest: Increased breathing frequency leads to diminishing pCO2 and consecutive pH increase.

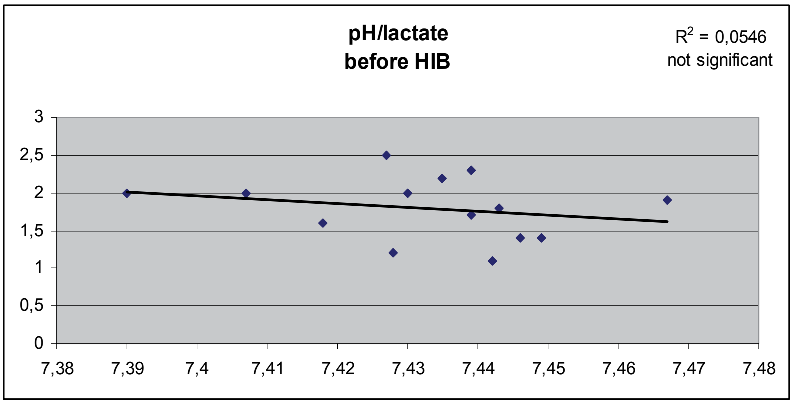
The highest breathing frequency, borne out by the most pronounced loss of CO2 leads to the most alkaline pH. This means, that here the most pronounced metabolic activity, expressed by the highest breathing frequency, paradoxically yields the highest pH values (fig. 4)
The picture is completely contrary to the familiar concomitant fall in pH and pCO2 during pronounced physical action (contest), which is shown in the next graph to underline the striking differences between the "warming up"(fig.4) and the real contest situation (fig.5):
pH / pCO2 after workload



Abscissa: pH
Ordinate: pCO2 in mmHg
P< 0,01, highly significant
Situation after the military steeplechase contest.

Fig. 5. During the demanding military exercise pH/pCO2 relation turns around, CO2 release is now unable to prohibit the fall in pH (note the extremely low pH and pCO2 values due to the heavy workload).

The adrenaline induced slight, individually different lactate increase in the pre contest situation should lead to a concomitant decrease in pH. But since at the same time H ions are indirectly got rid of by increased breathing, proportional pH decrease along with lactate increase is counteracted. Consequently, in the pre contest situation, pH and lactate do not correlate positively any more – as they do during the contest – due to the increased loss of CO2 during the said overcompensation, as shown in fig.6.



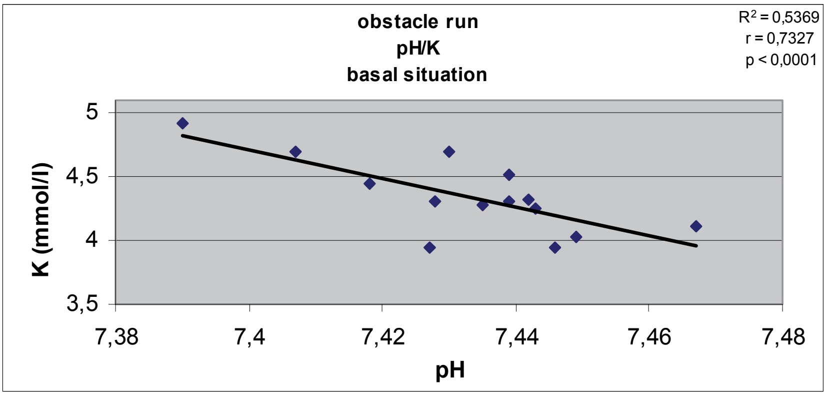
Abscissa: pH
Ordinate: lactate in mM/l
p>0,05, not significant
Situation before the military steeplechase contest.

Fig. 6. Due to successful CO2 control of pH before contest, pH/lactate relationship vanishes.

Consequently again, all our following correlative graphs of pre contest situations dealing at least with either pCO2 or pH have to be interpreted as situations, when the highest pH and the lowest pCO2 occur in the person with the most clearly increased metabolism. Having ourselves used those automatic correlative evaluations regularly, we came across overcompensation surprisingly frequently, mostly in pre contest situations or other moments of sympathoadrenal arousal. Therefore we would like to forward the supposition that overcompensation is a general feat of adaptation to probable demands in the future and thereby possibly an important part of evolutionary survival strategy.

b. Quantification of pre- contest conditions and characterization of outliers

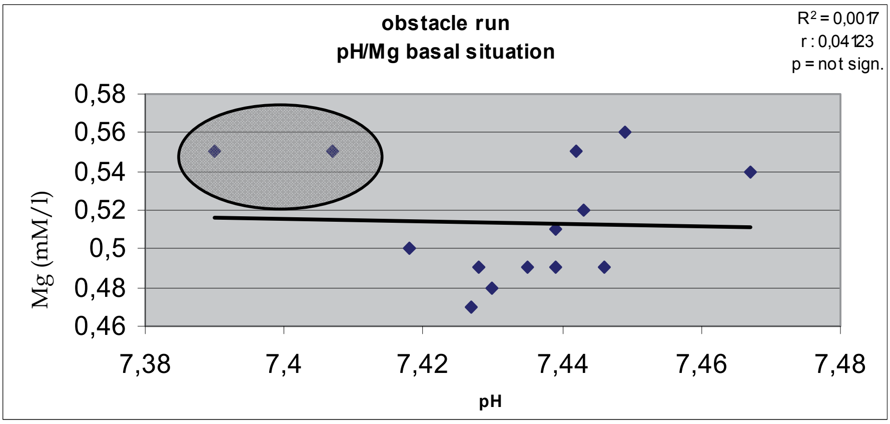
Following up our suppositions of a possible impact of the pre- contest situation upon the contest proper, we checked as a first step the K/pH proportions of the experimentees, because a possible sympatho – adrenal arousal in expectation of the contest may well lead to an individual increase of lactate values (the average increase being only a slight one, see tab.1), consequently to increased H ions, which would be exchanged with K ions from the tissue in a rate presumably proportional to the H ion production and therefore proportional to catecholamine impact. Indeed, a highly significant, but negative correlation between pH and Ionized K ensues, positioning the most pronounced K loss along with the highest metabolic turnover, characterized in this overcompensating situation by the highest pH and lowest K values. A combination of electrolyte- and metabolic parameters therefore are seemingly able to characterize typical group idiosyncrasies (overcompensation in this case) as well as the individual position of the participants within the group (fig.7):



Abscissa: pH
Ordinate: K in mM/l,
p<0,001, highly significant
Correlation between pH and potassium before sports.

Fig. 7. Sympathoadrenal arousal before contest incites overcompensation, also shown by significant inverse pH/K relationship.

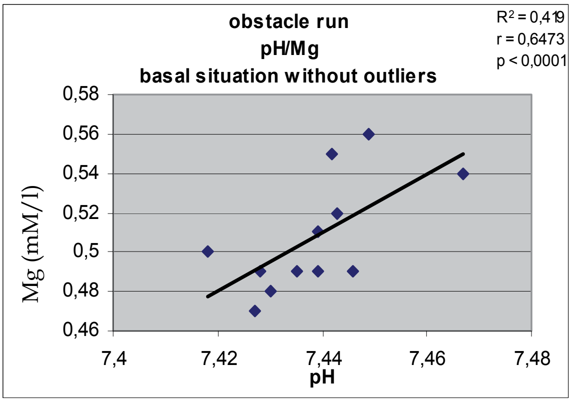
In such an adrenergic state of expectation, a further correlation can be expected, namely a proportional behaviour of pH and ionized magnesium, because adrenaline increase changes of pH via the mechanism mentioned above and can also increase ionized Mg in blood (8,9).
However, no significant correlation could be found. This could have been due to two outliers which are marked by an oval inclusion in the graph below (fig.8):



Abscissa: pH
Ordinate: Mg in mM/l,
p>0,05, not significant
points within the elliptic figure: presumptive outliers

Fig. 8. Positive pH/Mg correlation is disturbed by two outliers.

When the outliers are removed, a highly significant positive correlation between ionized Mg and pH in the blood of the pre-contest experimentees evolved (fig.9):



Abscissa: pH
Ordinate: Mg in mM/l,
p<0,001
Correlation between pH and magnesium before sports without outliers.

Fig. 9. Removal of the outliers leads to restored pH/Mg correlation.

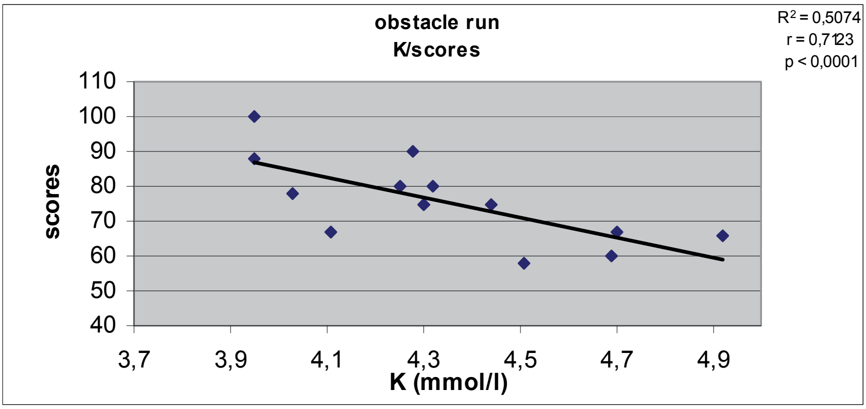
Similar as in fig.8, the highest Mg turnover (here Mg increase) goes along with the most pronounced metabolic turnover, again characterized by pH increase, due to overcompensation, provided the outliers have been removed. However, to characterize and/or remove the outliers out of purely statistical reasons is not correct, although it seems obvious, that they are not part of the sample. They both show high Mg values at concomitantly low pH which does not fit the group behaviour at all. On the contrary, this combination of parameters points towards an already most active metabolism before sports, which may not be able any more to meet the subsequently further increased energy turnover, necessary for high scoring. It follows, that they do not develop any sign of overcompensation, as does the rest of the group, since their high Mg values exist concomitantly with low pH, a feat that does nowhere occur in the rest of the overcompensating group. The pre – contest diagnosis of a prematurely increased metabolism of the outliers is consistent with some of the outliers' values *after* sports, forming a different multi parameter pattern:

1. very high BE values after sports
2. very high pCO2 values after sports
3. very high HCO3 values after sports
4. low scoring (ranking on 12th respectively 13th position within a group of 14).

According to this additional information, we felt justified to separate those two participants from the other members of the group, which – without those two - forms the mentioned well definable, highly significant pH/Mg relationship (fig.3). At least for practical reasons the information of their outlying position - which is only shown by the correlation analysis and would never have come forth by group average calculations alone - should on no account be discarded and at least used for closer observation and extended interviews with the two experimentees.

c. Prediction of success chances by quantitative evaluation of pre- contest conditions

Investigations into interdependencies between basal K and awarded scores (fig.10).For the first time we introduce non CSA values in our correlative interpretations – the awarded scores for the obstacle run, which are nearly identical with ranking of the period of time needed for its absolvation.



Abscissa: Basal K in mM/l
Ordinate: awarded scores
p<0,001

Fig. 10. Blood K levels before contest possibly predict chances of scoring at the contest proper.

Pre contest K concentrations and the scores awarded after steeplechase correlated negatively in a highly significant manner, meaning that persons with lowest pre- contest K values cherish the best chances for high scoring in the subsequent contest. Lowest K concentrations on the other hand coincide with highest pH levels in the overcompensating group (fig.1), which – under those circumstances – mark high metabolic activity. Within reason therefore, those who have been most successfully mentally "warming up" themselves, stand to be rewarded with better scoring chances. It is important, not to be led astray by the high scoring of contestants with an alkaline starting position, Slightly alkaline pH in this context is definitely not a sign of low metabolic turnover, but, as we have already been able to demonstrate, a feat of increased energy turnover by sympathoadrenal arousal, what we just called "mental warming up". Summing up the information of the graphs hitherto presented, the pre contest metabolic pattern of a presumable high scorer seems to be:

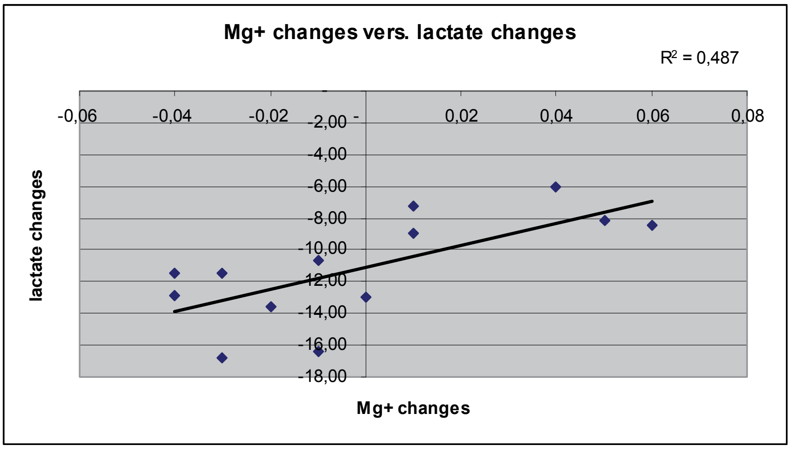
1. Low potassium (fig. 7)
2. High pH (fig 1))
3. High Mg (fig. 9)
4. Low pCO2 (fig. 4)

This multi parameter pattern shows, that increased metabolism characterized by the most pronounced loss of CO2, increases Mg clearance from the tissue and also increase blood pH. Since pre- contest pH is already increased by excessive CO2 loss, there is no urgent need for cation exchange, which is underlined by the low blood K values. Roughly spoken, K seemingly is allowed to stay put within the tissue, regardless of increased metabolism. Such a tissue reserve of readily exchangeable K ions however, could facilitate a more successful H ion removal into tissues later, during the demanding contest expected. That lactate values did not enter our multi parameter pattern more prominently can be explained by the failing correlation between lactate and pH. Increased breathing frequency obviously buffers direct lactate impact. Also – during the predominantly mental stress before contest – a participation of free fatty acids from catecholamine induced beta oxidation is to be expected and can indeed be roughly calculated by subtracting lactate values from baseexcess (both in mM/l)

d. Characterization of success/effort relationships

The role of certain electrolytes in blood, changed by the psychic arousal of the pre- contest situation could be shown already by correlating them with metabolic parameters like pH or pCO2. But they play a further role – again together with metabolic parameters - as indicators of competition success, part of which we have already exemplarily shown by the predictive power of K changes before competition. During our investigations of blood samples before and after the military obstacle run (HIB), however, we found connections of lactate and Mg changes with running times and scoring which contributed substantially to a better understanding of a contestants' attitude to the task. This may come in useful for basic research about the role of Mg in energy turnover but also has practical importance in such cases when one wants to check effort and performance of tasks where other means of objectivation cannot be applied or one is simply not present.

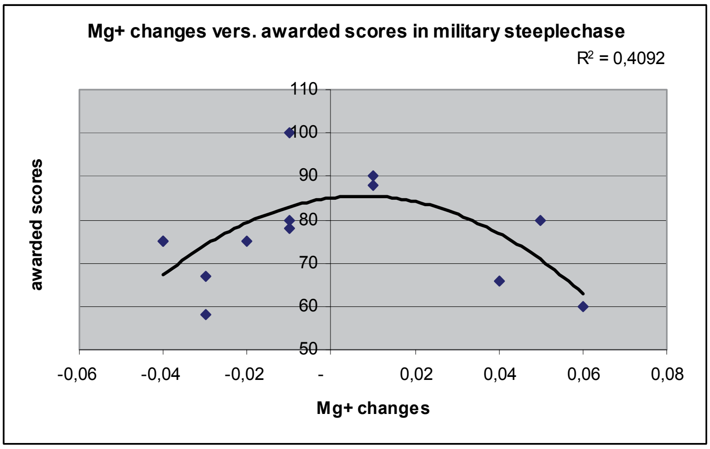
Therefore let us have a look into the Mg/lactate relationship, which we purport to be especially apt to reveal the individual attitude towards the contest: Although individual Mg+ changes were rather pronounced, no corresponding *average* Mg+ increase or decrease was visible because Mg+ changes often pointed in opposite directions. This individual behaviour of the Mg+ changes however correlates highly significantly with lactate changes. Fig.11 shows, that sometimes more information can be gained by using the – also automatically compiled – delta values, the changes between the values before and after workload and not the measured values themselves. Thus clear proportionalities between effort and electrolyte turnover could be shown.



Abscissa: Mg+ changes (delta values: Mg+ before exercise minus Mg+ after exercise)
Ordinate: Lactate changes (delta values: lactate before exercise minus lactate after exercise)
P< 0,001, highly significant

Fig. 11. Documentation of delta values sometimes are better able to reveal parameter interactions than absolute concentrations.

Higher lactate increases beyond 11 mM/l correspond with Mg+ increase, lactate changes below that value corresponds with Mg+ decrease in blood. That lactate changes do not correlate with both awarded scores and running time is not surprising, since it is common knowledge that subjects with better fitness may score higher with less lactate increase than less well trained subjects with higher lactate increase (3). But unexpectedly, Mg+ changes (automatically computed delta values) did correlate with awarded scores and running times as well, highly significantly in a polynomial curve of the 2nd order. Fig. 12 shows, that Mg changes not only correlate with lactate changes (effort) but also with awarded scores, but now in a different, polynomial manner.



Abscissa: Mg+ changes (delta values)
Ordinate: scores awarded
P< 0,001, highly significant

Fig. 12. Mg delta values this time plotted against scores, once more prove their usefulness.

Since changes in blood Mg+ are to be considered as subtractions of Mg+ influx from tissue and Mg+ redistribution (10) by clearance from the blood, people with the best balanced Mg+ in- and efflux show nearly zero deviation. And it is exactly this group that has been shown to have the best chances for highest scores and shortest running times. Moreover, the supposed unpredictability of positive or negative Mg+ changes, at least during short term exercise with high energy turnover, may be qualified by the existence of this polynomial correlation between both scores and running times and the Mg+ changes in question.

Distinctly increased Mg+ blood concentrations therefore, went along with the high lactate increases beyond 11 mM/l, mostly in participants with low scores. Contrarily, low scores and long running times went together with pronounced decreases in Mg+ along with comparatively low lactate increases otherwise seen in high scorers, but then with hardly any Mg+ changes.

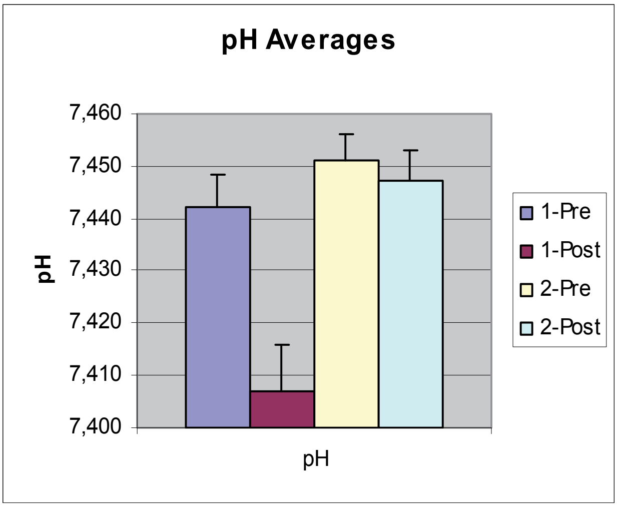
Thus, a typical combination of lactate and the obviously quicker Mg changes is characteristic not only for a certain score, but also provides more information about the reason for low scoring, when one combines the information of fig.4 and fig.5: High Mg+ increase is mostly associated with pronounced lactate increase and pronounced Mg+ decrease with a relatively small lactate increase for this demanding kind of exercise. All the high lactate increases therefore are on the side of Mg+ increase and vice versa. Moderate lactate increase, along with moderate Mg+ change seem to characterize the reaction of a well trained subject.

One is tempted to deduce, that in our short and demanding exercise the mark of an average delta value of 11mM/l of lactate may be the turning point between a preponderance of Mg+ clearance from the blood during the first stages of the exercise and a consecutively increased Mg+ influx into the blood, overcoming the clearance rate, because of considerably increased demand upon muscular tissue.

This turning point should be demonstrable by a polynomial curve, which indeed it has been. It furthermore may even serve as a kind of standardization mark, beyond which additional lactate increase (effort) does correspond less and less with effectiveness. A possible practical application of the combined information of fig.4 and fig.5 could be i.a. an objectivation of subjective adjudications of any persons' effort and success. Hitherto, mostly increases of Mg+ blood concentrations during short term exercise have been advocated, while Mg+ decrease has been mainly associated with longer lasting workloads (11, 12, 13) Biphasic Mg+ changes, at least during short term exercises have not been described up to now. As mentioned above, the results of such an investigation allow basic considerations about new aspects of the role of electrolytes as well as pave the way towards some practical progress in adjudication of success and effort relationship.

ad.2: Metabolic changes due to mental stress in depressive patients

Having been able to show that the predominantly psychically induced change in metabolic parameters before a contest can be measured and thus quantification of psychical arousal by metabolic determinations can be attempted at least proportionally, we would like to show an application of this idea at psychiatric in - patients. Our clinical study included 19 patients (17 females and 2 males) with a mean age of 44 years (range from 24 to 65, with a median of 44). All of them were suffering from major depressive disorders (Hamilton depression scale from 18 to 33). We compared them with a group of .46 subjects (35 males and 11 females, nearly equally aged) Before and after a slight ergometric effort (60 watts for 6 minutes) capillary blood samples were drawn as described above and the resulting group averages

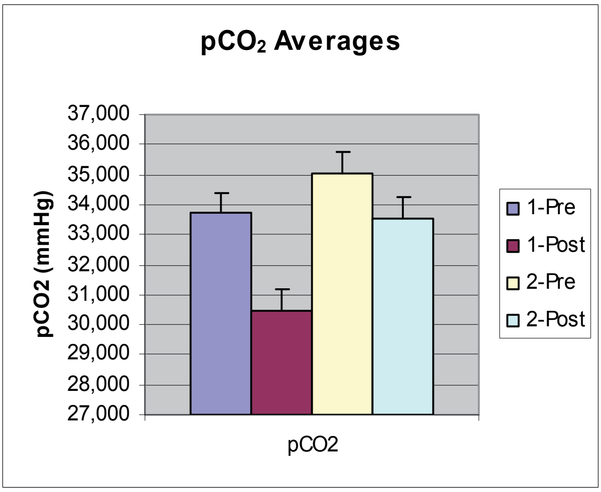


Abscissa: determination situations (see text)
Ordinate: pH

Fig. 13.

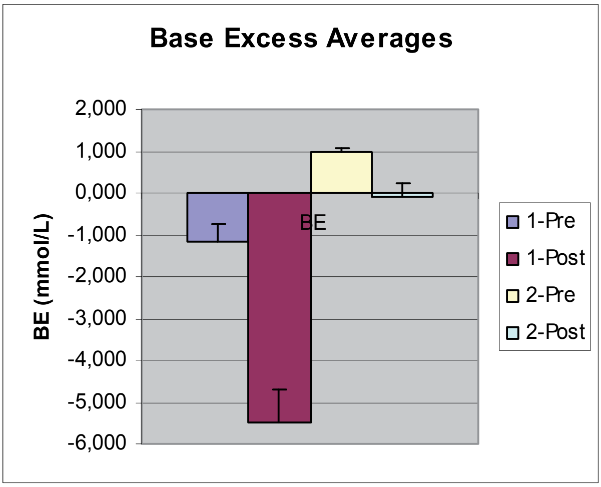
for pH, pCO2, Basexcess and Magnesium are shown in the next four automatically generated graphs. The first two columns of each graph show the group averages before (col.1) and after workload (col.2) of the psychiatric patients, columns 3 and 4 show the group averages of the equally treated healthy group

Fig. 13: Figs. 13, 14, 15 and 16 deal with the average changes of pH (fig.13), pCO2 (fig.14), base excess (fig.15) and magnesium (fig.16) of psychiatric in – patients (blue and red column) and a matched control group from our data banks (yellow and green columns). Remark the much more sensible reaction of the patients to workload.



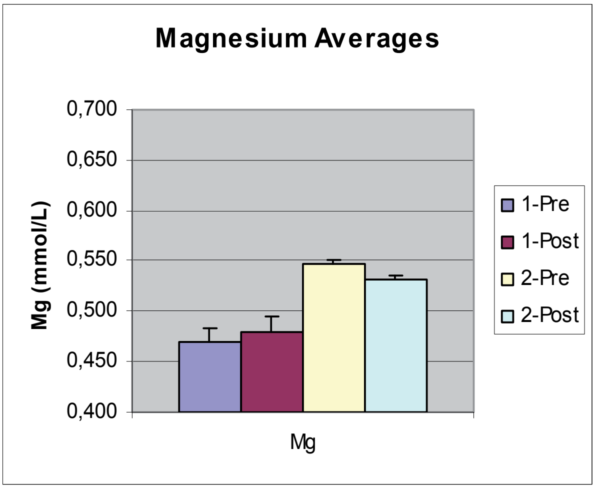
Abscissa: determination situations (see text)
Ordinate: pCO2 in mmHg

Fig. 14.



Abscissa: determination situations (see text)
Ordinate: Baseexcess in mM/l

Fig. 15.



Abscissa: determination situations (see text)
Ordinate: Ionized Mg in mM

Fig. 16.

With the exception of the baseexcess there were no significant differences in the basal situation between the depressive and the healthy group. Since the significantly lower buffer capacity, shown by the baseexcess values of the depressive group (fig. 15, col.1 and col. 3) cannot be due to acutely increased breathing, (fig. 14, col.1 and col.3) it has to be acknowledged as a chronic buffer diminishment of longer standing, developed in the course of the illness. This is underlined by the significantly and clearly more intensive reaction to the slight workload by the depressive patients. Their pCO2, their pH and their baseexcess react disproportionately sensitive to the moderate workload. Such an expected accumulation of over sensitive reactions to daily demands may well have been the reason for a chronic decrease of their total buffer capacity in the course of the illness. Accordingly, our investigations into the differences of metabolic reaction between depressive and healthy people yield – just by glancing at the automatically generated graphs and average statistics - two general results:

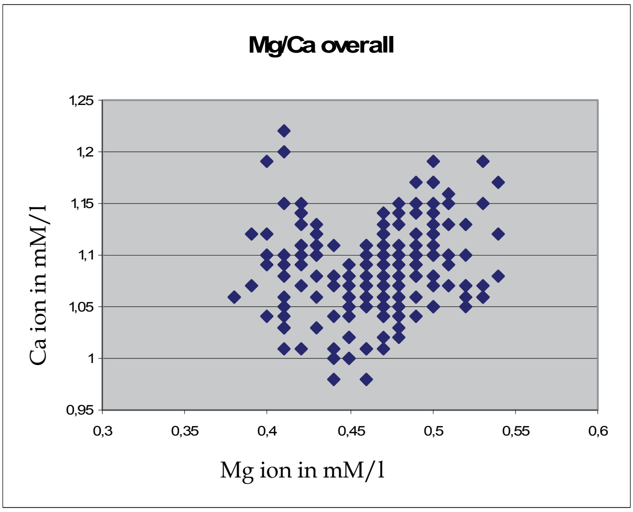
- The well known unwillingness of depressive patients or burnout patients to perform bodily feats is obviously not only due to depressive moods but also to a handicapped metabolism, demonstrable mainly by a disproportional reaction to provocation. Careful attempts to increase bodily fitness of the patients may be rewarded.
- Although metabolic tests can never be used diagnostically in mental illnesses, repeated checks of the metabolic situation could be helpful for documentation and quantification e.g. of the progress of medication.

ad 3: Idiosyncrasies of the diabetic metabolism, especially those due to the newly found importance of mineral deficiencies in type2 diabetics uncovered by CSA diagnosis

The blood glucose status of diabetic in - patients is routinely checked by a daily glucose profile, which consists of glucose determination from capillary blood at three different times. With nearly the same small amount of blood and the same effort we could determine not only glucose but also 11 additional parameters. The results of those investigations have been published in some full papers and several congress abstracts. We could show i.a. that not only type1 diabetics but also about 36% of the nearly tenfold higher number of type2 diabetics suffer from severe loss of electrolytes, especially from hypomagnesemia. However, the magnesium state of type 2 diabetics has not been considered to be crucially important for the patients' wellbeing up to now, since only easily treatable cramps were thought to ensue from magnesium deficiency. But by correlatively combining of some of our simultaneously determined parameters, we could show, that diabetic hypomagnesemia seems to be responsible not only for the said cramps, but for a whole series of negative influences upon the already strained diabetic metabolism:

Let us direct, e.g. our attention to some differences in metabolic behaviour in patients with Mg levels below and above the hypomagnesemic threshold of 0,45mM/l ionized Mg in blood (hypomagnesemic threshold according to the Austrian Consensus Conference as well as the Deutsche Gesellschaft für Ernährung – German Society for Nutrition) and exemplarily look at some facts accompanying those differences: As already mentioned above, severe deficits in ionized blood magnesium became increasingly conspicuous during investigations into interactions of blood glucose, buffers and electrolytes during daily glucose profiles of type2 diabetic patients, since we had the opportunity of magnesium determination with ion sensitive electrodes (NOVA CCX, CSA). This fraction, according to our knowledge, has not been compared yet with blood glucose metabolism in type2 diabetic patients to any larger extent.

Similarly, investigations about the behaviour of ionized calcium in type2 diabetics seem at least to be rare. Its average values in our patients are, like those of magnesium, very low. Also remarkably low were the base excesses of the patients, though lactate concentrations in blood did not exceed normal elevations found on moderately busy days.



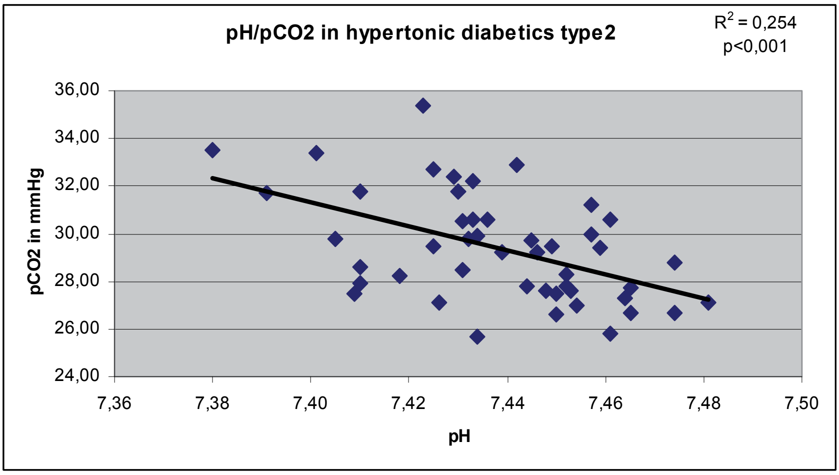
Abscissa. Ionized Ca in mM/l
Ordinate: ionized Mg in mM/l

Fig. 17. Remarkable change of relationship of Mg and Ca in the blood of diabetics nearly exactly at the point of the hypomagnesiemic threshold.

Magnesium and calcium averages give the impression to be inversely proportional to the concomitant blood glucose values, a feat that has been already mentioned by others together with magnesium and blood glucose or insulin sensitivity (17). But when we put together all single values of all our patients, regardless of sampling times, there was no significant inverse correlation between ionized magnesium and blood glucose. Still, the pattern of the individual points in the Ca/Mg graph was exceptional (fig.6). Exactly at the Mg value of 0,45 mM/l (the agreed hypomagnesemic threshold) the Ca/Mg relationship seemed to switch directions. Fig.17 shows the turnaround of the relationship of ionized Ca with ionized Mg. Obviously, only from the hypomagnesemic threshold (0,45mM/l) upwards a significant, positive correlation has developed.

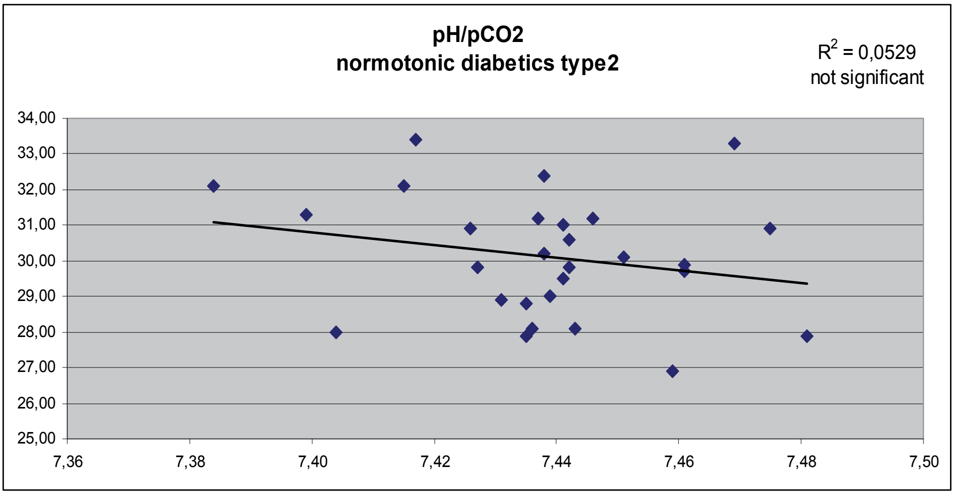
Even without indrawn regression lines one can observe, that the Ca/Mg ratio takes an opposite course nearly exactly above and below the hypomagnesemic threshold of 0,45 mM/l. We acknowledged this ambivalent behaviour by splitting the sample along this threshold of 0,45 mM/l into a high- and a low Mg subgroup. Consequently, we found a highly significant positive correlation between Mg and Ca in the higher Mg subgroup but no correlation at all in the lower subgroup. This finding encouraged us to look for more correlations, not within the overall sample but again within the subgroups above and below the hypomagnesemic threshold, trying to find at least some hints for the reason of the very low Mg values in our diabetics, where 36 % had Mg concentrations of 0,45mM/l and lower, since the mechanism of low magnesium values in NIDDMs seems to be unclear, Some authors (18,19) discuss a recurrent metabolic acidosis, along with episodes of osmotic diuresis as possibilities among others for magnesium diminishment in the diabetic patients, while stating that this diabetic hypomagnesemia seems to merit poor attention by physicians anyway. Shaffie et al (20) observed a lowering of bicarbonate along with low tissue pCO2 and hyperventilation.

Indeed, pCO2 values in our patient sample are low. Additionally, correlations between pH and pCO2 overall and in the subgroups showed a significantly inverse behaviour, most clearly expressed in the low Mg subgroup, with a slope more than double as steep (y=59x) as in the higher Mg subgroup (y=23,6x) , pointing to an increasing need for respiratory compensation along with diminishing Mg concentrations. Thus, pH seems to be kept at an average of 7, 43 by constant loss of CO2, obviously slightly overcompensating a steady input of anions.



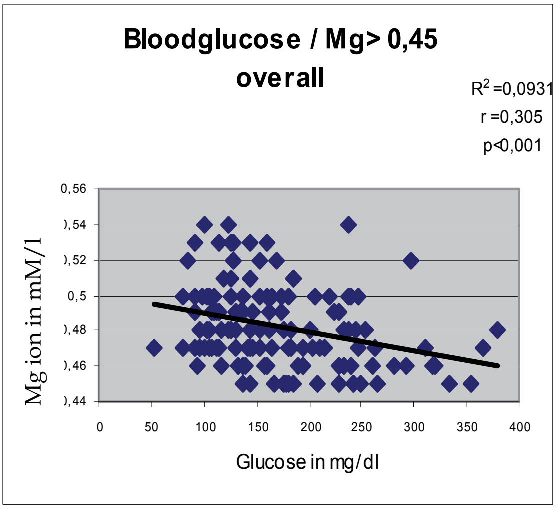
Abscissa. pH
Ordinate: pCO2 in mmHg
P<0,001 highly significant

Fig. 18. Increasingly lower pCO2 creates more and more alkaline pH in hypertonic diabetics.



Abscissa. pH
Ordinate: pCO2 in mmHg
P>0,05 not significant

Fig. 19. Normotonic diabetics do not show any relationship between pH and pCO2 (see fig. 18) at all.

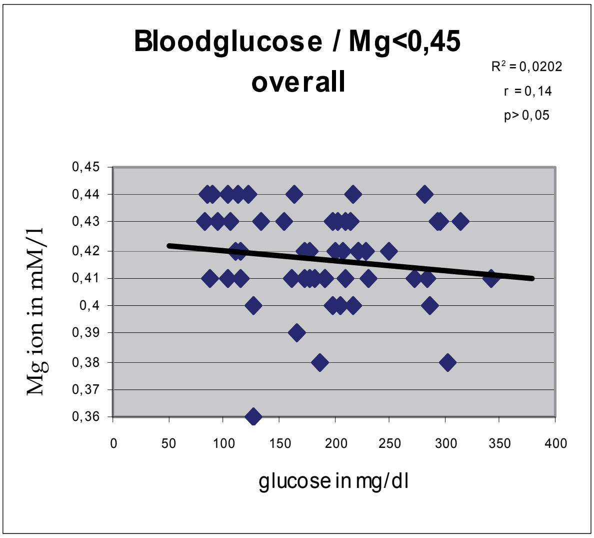


Abscissa. Glucose in mg/dl
Ordinate: ionized Mg in mM/l
P<0,001 highly significant

Fig. 20. Above the hypomagnesiemic threshold blood glucose increases along with lower Mg levels – a good argument for substitution.

Such chronic processes accompanied by a moderately increased breathing frequency may slowly but successfully waste the magnesium (and calcium) resources of the patient. The interesting observation, that patients with Mg concentrations below 0,45mM/l do not seem to show correlations with blood glucose any more, may indeed point towards a certain exhaustion, for which low Mg concentrations are usually characteristic. But low Mg in those patients may be not only a marker of increasing metabolic exhaustion, but could actively contribute to wasteful anaerobic glycolysis by limiting ATP – ADP turnover. At least in non diabetic patients, we could show that low magnesium concentrations before extensive liver surgery deteriorate the prognosis about the final outcome significantly (21). The most important result concerning type2 diabetics is unquestionably the highly significant negative correlation between Mg levels and blood glucose above the hypomagnesemic threshold.

It means that diabetic patients with lower Mg levels are prone to higher blood glucose values.



Abscissa. Glucose in mg/dl
Ordinate: ionized Mg in mM/l
P 0,05 not significant

Fig. 21. Below the hypomagnesiemic threshold the calculable Blood glucose/ Mg relationship vanishes, but the number of patients with high glucose values is remarkable.

Therefore we think that Mg determination, especially, that of the more active ionized fraction, should be included into the monitoring at least of hospitalized diabetic patients. When interpreted together with blood glucose levels and other CSA parameters, it reveals a much deeper insight into the metabolic state of the patient. In our opinion, increased knowledge of physicians about the impact of Mg deficiency upon the diabetic (and also non-diabetic) metabolism would increase the demand for magnesium determination and also for magnesium medication.

Consequently, we can see, that without increased effort, just by substituting a more up to date measuring device coupled with appropriate software, the daily glucose profile, a routine diagnostic method of very long standing, could be changed into a much more sensitive investigative tool, capable of quickly unearthing new knowledge about metabolic dynamics for the benefit of the patients.

ad 4: CSA application in animals (outlook)

About 15 years ago we investigated the catecholamine state of immobilized rats and of pigs in abattoirs. Immobilisation of animals led to dramatic increase in catecholamines and to vastly diminished stress compatibility (22, 23) Catecholamine levels of pigs before slaughtering in abattoirs have been found to be an incredible hundred thousand fold higher than normal.

Both, immobilisation and the pre- slaughtering situation are widely common in the treatment of pigs, since mother sows are often kept practically immobilized in very small cages. Objectivation of the effects of the demonstrated catecholamine increase upon the metabolic parameters provided by simple and scarcely molesting CSA testing could reveal at least electrolyte changes in blood, most probably vastly increased electrolyte input into muscle tissue. Low quality, watery meat could well be the outcome. In immobilized mother sows the metabolic effects of catecholamine elevation may also have a whole bunch of negative effects, easily imaginable (24). At the very least CSA metabolic investigation may provide prove, that cruelty to animals does not pay, in fact actually decreases profits of husbandry.

Concerning the application of CSA tests in the training of e.g. racing animals like horses, dogs or camels, it seems easy to adapt our methods and results from our investigations in humans.

It is obviously an asset for both animals and trainers to be able to adjudge the pre contest condition, thereby the contest chances and the metabolic changes during a given contest of a specific animal. The familiarity of a good trainer with animals in his care, by which up to now subjective judgement has been delivered, could be supplemented by objective testing. Hitherto surprising reactions during the contest may thus be more successfully avoided, momentary fitness state of the animal more correctly adjudicated, latent illnesses better anticipated as benefits for animal, trainer and owner.

2. General outlook

By providing therapists, trainers or commanders of first responder units with a multi – parameter pattern of stress hormone effects and their statistics within minutes, an objective support of decisions which range from selection of specific treatments, adaptation of training efforts or educated guesses about success chances of teams to detection of team *outliers* is provided.

Still, the decision making responsibility of the physician, the trainer or the commander remains untouched. On the one hand, the intricate pattern of the interwoven parameters provides a rather sensitive metabolic picture, which can be broadened by additional correlations with sociological and psychological scores (for which free spaces in our correlation tables are already provided). But the same multi parametric intricacy prohibits – at least at the moment – a mathematical overall calculation including all changing parameters according to their momentary importance. It turned out, that pure multiple regressions of the bulk of all data are not stable enough, so that factor analyses have to be applied in advance. Some smaller problems could be already solved in this way, but generally the best way to use our tool at the moment is a one days course in interpretation with lots of practical applications and some theoretical background, which in nearly all cases enables the user to proceed successfully on her or his own.

3. Summing up

Stress hormone changes correlate significantly with their also changing effects like lactate, blood gases, buffer capacities and electrolyte concentrations. About twelve of those parameters can be determined simultaneously by easily transportable ICU equipment from 100 microliter of capillary blood within three minutes and the data be transferred on line by CSA (Clinical Stress Assessment) software into data banks. Individual statistics or group statistics are automatically calculated, so hat a kind of data net is spread over the proband or a whole group of probands. Changing of sensitive parameters like pCO2 or Ca are indicative of acute effects, while changing of less sensitive values like HO3 or blood glucose point towards more chronic developments.

Application of CSA has been found fruitful i.a. to detect and quantify:

- The pre – contest stress situation of sports persons or pre deployment situation of first responders like soldiers or fire fighters, thereby being able to select outliers and calculate scoring changes preliminarily. A persons' success/effort relationship can be quantified.
- Since mental stress imprints its quantity in nearly all cases upon metabolic changes, those changes and their idiosyncrasies can be quickly measured and taken as sensitive pointers for e.g. a gradual effectiveness check of medical treatment.
- Inclusion of CSA application in diabetic check ups can provide the physicians with a more then 10 – fold increased information about blood glucose changes within the individual metabolic ambient of a patient, out of nearly the same small amount of blood within 3 minutes. It enabled us at least to find hitherto hardly suspected deficiencies and interactions of electrolytes and glucose in type2 diabetics, with a significant further deterioration in hypertonic diabetics.
- Concerning the predictive and in depth assessing capabilities of CSA concerning sports and first responders, it hits the eye that those very capabilities could be applied successfully in training of competing animals as well as in the control of unnecessary stress of husbandry situations.

4. Acknowledgment

The authors would like to thank G. Porta and T. Porta for their valuable assistance.

5. References

[1] Porta S., Emsenhuber W., Petek W., Pürstner P, Vogel W., Schwaberger G., Slawitsch G., Korsatko Detection and Evaluation of Persisting Stress Induces Hormonal Disturbances by a Post Stress Provocation Test in Humans. W. Life Sciences, Vol.53,pp.1583-1589, 1991).

[2] W. Emsenhuber, S. Porta, W. Petek, P. Pürstner, W. Vogel, P. Felsner, G. Schwaberger, P. Slawitsch, Retrospective Stress Measurement By Standardized Post-Stress Provocation, in „Stress: Neuroendocrine and Molecular Approaches" Edited by R. Kvetnansky, R. McCarty and J. Axelrod Gordon and Breach Science Publishers S.A., New York, USA, 1992

[3] Desch W., Schappacher W., Schappacher G., Wintersteiger R., Ecker M., Köhler U., Korisek G., Porta S. An attempt to quantify the influence of some IC parameters upon the levels of ionized Mg in blood. Trace Elem. Elec. 26; 2/2009 89 – 94

[4] S. Porta, K. Pichlkastner, H. Gell, W. Desch, J. Porta, and M. M. Bratu, Interdependencies of electrolyte and metabolic parameters can characterize handicaps and predict success in sports. Trace. Elem. Electr. 27, 3, 2010, 103 – 109

[5] S. Porta and K. Kisters, Electrolytes and sports – special role of magnesium, Editorial, Trace. Elem. Electr. 27, 3, 2010, 103 – 109

[6] JH Foreman, JK Waldsmith and RB Lalum. Physical, acid–base and electrolyte changes in horses competing in training. Preliminary and intermediate horse trials, Equine and Comparative Exercise Physiology (2004), 1:99-105 Cambridge University Press.

[7] Porta S, Gell H, Pichlkastner K, Cichocki G, Desch W, Schappacher W, Korisek G, Grieshofer P and Stelbrink U. Direct correlation between Mg++ changes and awarded scores in military steeplechase (HIB). Trace Elem.26,4/2009 177 – 180.

[8] Thomas L. "Labor und Diagnose", TH – Books Verlagsgesellschaft Frankfurt/Main 1998.

[9] S. Porta, W. Desch, G. Korisek, K. Kisters, J. Porta, H. Gell und M. M. Bratu. Eine neue Möglichkeit zur Erfassung von Stoffwechselbesonderheiten bei hypertonen Typ2 Diabetikern. Nieren und Hochdruckkrankht. 39, 5/2010, 220 – 230.

[10] Kisters K, Hoffmann O, Hausberg M, Quang Nguyen M, Micke O, Gremmler B. Plasma magnesium status and pulse pressure in essential hypertension. Trace Elem Elec 22; 2005; 67 – 70

[12] Westmoreland D, Porta S, Leitner G, Knapp M, Spencer K, Merback J, Leitner T. The effect of magnesium supplementation on exercise-induced plasma magnesium shifts and lactic acid accumulation in female youths. Trace Elem Elec. 2004a; 21: 95-98.

[13] Westmoreland D, Orland U, Porta S. The diurnal rhythm of plasma magnesium and its probable influence on the exercise-induced plasma magnesium shift. Trace Elem Elec. 2004b; 21: 185-189.

[15] S. Porta, G. Korisek, B. Machan, H. Gaggl, W. Desch, J Porta, G. Schappacher, K. Kisters Changes in electrolyte- and metabolic interactions below the hypomagnesaemic threshold in Type 2 Diabetics. Trace elements and electrolytes 27/3 119 – 124, 2010-.

[16] S. Porta, K. Kisters, G. Korisek, G. Desch, J. Porta and H. Gell, Differences in electrolyte mismanagement between normotonic and hypertonic Type2 diabetics detectable by correlative capillary blood evaluation, Trace. Elem. Electr. 28/1 2011, 31 – 36.

[17] Barbagallo, M, Dominguez, LJ. Magnesium metabolism in Type2 diabetes mellitus, metabolic syndrome and insulin resistance. Arch. Biochem. Biophys. 2007, 458: 40 - 47

[18] Sales CH, Pedrosa LdeF: Magnesium and diabetes mellitus: their relation. Clin. Nutr. 25, 2006, 554 – 5.

[19] Phuong – Chi T. Pham, Phuong – Mai T. Pham, Son V. Pham, Miller J-M, Phuong – Thu T. Pham. Hypomagnesaemia in patients with Type2 diabetes. Clin J Am Soc Nephrol 2: 2007 366-373.

[20] Shaffiee M.A, Kamel S. Kamel, Halperin M.L: A conceptual approach to the patient with metabolic acidosis. Nephron (Suppl.1): 46 – 55 92; 2002.

[21] Bacher H, Mischinger HJ, Cerwenka H, Werkgartner G, El-Shabrawi A, Supancic A, Porta S. Liver ischemia, catecholamines and preoperative condition influencing postoperative tachycardia in liver surgery. Life Sciences 1999; 66(1):11-18.

[22] Porta S., Kvetnansky R., Oprsalova O., Emsenhuber W., Leitner G, Plasma Epinephrine Elevation Increases Stress induced Catecholamine Increase Even After an Interval of 24 Hours. 7th International Catecholamine Symposium, Amsterdam June 22-26, 1992

[23] Porta S., Leitner G., Kvetnansky R., Kaciuba Uscilko H., Nazar K., Teter A., Emsenhuber W. Cumulative Effects of Repeated Stress Detectable Even After Very Long Intervals. International Symposium On Stress And Adaption, Amsterdam June 19, 1992.

[24] Porta S., Ehrenberg A., Helbig J., Classen H.G. ,Egger G., Weger M., Zimmermann P., Weiss U., Time and Dose Dependent Influence of Magnesium-Aspartate-Hydrochloride Treatment Upon Hormonal and Enzymatic Changes as well as Alterations in Meat Quality Due to Slaughtering Stress in Pigs.Mg. Bull. 12/3, 1993, 21 – 24.

21

Neurotechnology and Psychiatric Biomarkers

William J. Bosl
Harvard Medical School,
Children's Hospital Boston Informatics Program,
Boston University School of Medicine Behavioral Neuroscience Program,
USA

1. Introduction

Neuroscience as a scientific discipline has enjoyed enormous growth and success in the past decade. Some have called the early 20th century the golden age of physics, the latter half of the 20th century a period when the genomic revolution blossomed and predict that the early 21st century will be a period when brain sciences achieve remarkable success. While understanding the basic mechanisms of the brain and how they relate to thought and behavior may be the foundation for applications to medicine, there is a great need for technological innovation if more than academic results are to be achieved. The need for neurotechnology is already great. Traumatic brain injury and damaged limbs require prosthetic devices that can be controlled in some way by willful volition. Ideally, direct connections between thought and action are desirable to restore natural functions. Mental health is a branch of medicine that has long been relegated to a secondary status within medicine. The reasons for this may be many, but they certainly include the difficulty of understanding and measuring brain activity in a quantitative way and relating those measurements to behavior and cognitive activity. As healthcare costs continue to spiral out of control in both developed and developing regions of the world, the need for engineers to become involved in neuroscience and neurotechnology research and development has never been greater. Innovative engineering ideas, with a view toward practical application and affordable cost have much to contribute to clinical applications of brain science. A key contribution of neuroengineering will be innovative methods for quantitative measurement of brain activity and mapping of those measurements to behavior and thought. The term psychiatric biomarkers will be used here in this broad sense to indicate quantitative measurements of the brain and the algorithms necessary to interpret them in psychiatric or psychological descriptions or diagnoses.

One important way in which psychiatric biomarkers differ from other physiological biomarkers is in that the mapping from biomarker to symptom or disease is much more complex. A biomarker for a specific cancer, for example, may be a gene mutation that is in some way directly involved in the disease progression itself. The relationship between biomarker and the manifestation of interest – cancer, in this case – is rather simple and direct. That is not to say that the gene or the gene expression patterns are simple to find, but only that the conceptual relationship between the marker and disease is simple to understand. In the case of psychiatric biomarkers, the phenomena of interest, thought and behavior, are complex, emergent phenomena of brain neural activity. The relationship between neural firing patterns and the communication deficits that are clearly evident in a person with a mental disorder is not at all clear, even if we posit that all thought is indeed dependent upon neural activity. The relationship in this case is somewhat like the relationship between letters of the alphabet and a metaphor in great literature. Certainly metaphors depend on spelling and grammar, but the concept is much more than spelling and grammar. Similarly, the complex patterns of neural activity that distinguish the way a person with autism responds to someone speaking directly to them from someone considered "normal" are quite complex.

Normal and abnormal behavior are differentiated by subtle, complex patterns of activity that a trained expert observes or discovers through systematic diagnostic tests. If brain function and behavior are mirrors of each other, as is commonly accepted (Cowan and Kandel, 2001; Hyman, 2007; Kandel, 1998; Singh and Rose, 2009), then biomarkers of mental disorders may be hidden in subtle, complex patterns of neurobiological data. There is a growing realization that the neurophysiological mechanisms that underlie brain function cannot be understood by pure reduction to physiological causes (Stam, 2005; Ward, 2003). The dynamics of the brain is inherently nonlinear, exhibiting emergent dynamics such as chaotic and transiently synchronized phenomena that may be central to understanding the mind-brain relationship (Varela et al., 2001) or the 'dynamic core' (Le Van Quyen, 2003). The behaviors and thoughts that characterize mental dysfunction may be emergent phenomena or complex patterns of physiological processes, especially neural processes. For example, major depression or the communication deficits present in a child with autism are emergent phenomena that reflect complex patterns in brain function that differ from some socially-defined norm. The task of the neuroengineer is to create new technology to measure and interpret the patterns of brain activity that connect brain measurements to observed behavioral patterns.

A key challenge in cognitive neuropsychiatry is to discover the neural correlates underlying behavior. To be clinically useful, these discoveries must be accompanied by technology that enables brain activity to be measured and interpreted safely, inexpensively and easily. The explosive growth of neuroimaging studies that link functional brain activity to behavior promises exciting opportunities for measuring nonlinear brain activity that may indicate abnormalities or allow response to therapy to be monitored. While several imaging modalities are available for neuroscience *research*, most have significant limitations that prohibit their use as routine *clinical* tools. Cost and ease of use are essential qualities for clinically useful tools, which may not be as important or relevant in a scientific research context. Neuroengineers must be cognizant of these constraints when considering the intended use of the technology.

Measurements of brain electrical activity with electroencephalography (EEG) have long been a valuable source of information for neuroscience research, yet this rich resource may be under-utilized for clinical applications in neurology and psychiatry (Niedermeyer, 2003; Niedermeyer and Lopes da Silva, 2005). To fully exploit this data, methods for discovering subtle nonlinear patterns and deeper understanding of the relationship between emergent signal features, neurophysiology and behavior are needed. Near infrared spectroscopy (NIRS) has recently been introduced as a safe, portable alternative for measuring blood oxygen level dependent (BOLD) response in infants (Irani et al., 2007; Muehlemann et al., 2008). One of the primary advantages of NIRS, like EEG, is that it is safe for all ages, relatively inexpensive and portable. As a new brain-imaging tool, much remains to be discovered about the value and limitations of NIRS as a clinical instrument. In addition, coupling EEG and NIRS may have some advantages for clinical use and remains to be explored by researchers. Many of the advances in non-invasive functional brain measurement are being driven by the brain computer interface community, where mobility and cost requirements limit the technologies that can be adopted (Dornhege, 2007). Neuropsychology and cognitive neuropsychiatry can learn from this community, while adapting the methods to the particular needs of behavioral, affective and cognitive assessment.

In this chapter some relevant information concerning our current understanding of complex network organization and implications for finding neural correlates of behavior will be reviewed with goal of motivating engineers to consider contributing their skills to developing new neurotechnology. Considerable attention will be given to EEG measurements as one of the most promising technologies for clinical application to neuropsychiatry. Novel methods for extracting information from EEG signals are beginning to appear, taking advantage of advances in the physics of nonlinear systems and signals, complex network theory and machine learning algorithms. The need for innovative neurotechnology to meet the need for mental and neurological healthcare in developing regions of the world is great, but the promise is even greater. The primary goal of this chapter is to provide information to enable researchers interested in brain disorders and mental health to become involved in creating innovative neurotechnology for clinical use.

2. The brain as a complex system

2.1 Complex networks

The human brain contains on the order of 10^{11} neurons and more than 10^{14} synaptic connections (Kandel et al., 2000). Although sparsely connected, each neuron is within a few synaptic connections of any other neuron (Buzsáki, 2006). This remarkable connectivity is achieved by a kind of hierarchical organization that is not fully understood in the brain, but is ubiquitous in nature, called scale-free or complex networks (Barabasi, 2009; Bassett and Bullmore, 2006; Ravasz and Barabasi, 2003). Complex networks are characterized by dense local connectivity and sparser long-range connectivity (Barabasi, 2009) that is fractal or self-similar at all scales.

Many brain disorders appear to be associated with abnormal brain connectivity that may vary between different regions and different scales (Bassett and Bullmore, 2009; Craddock et al., 2009; Noonan et al., 2009). Examples include autism (Belmonte et al., 2004; Noonan et al., 2009), schizophrenia (Raghavendra et al., 2009; Uhlhaas et al., 2008; Whittington, 2008), depression (Li et al., 2008; Sheline et al., 2009) and epilepsy (Douw et al., 2010; Percha et al., 2005). Methods for estimating neural connectivity or changes in neural connectivity might be effective diagnostic biomarkers for abnormal connectivity development that is associated with brain dysfunction. The electrical signals produced by neural networks are believed to contain information about the neural network structure on several scales in the vicinity of the EEG sensor (Raghavendra et al., 2009; Stam, 2005; Zavaglia et al., 2008).

Novel analysis methods from nonlinear systems theory are able to extract information from EEG signals that reflect the underlying network organization. System invariants will be encoded in the time series produced and measured by EEG sensors (Fuchs et al., 2007; Gao and Jin, 2009). Multiscale entropy (MSE) is one invariant measure of system dynamics that has been shown to particularly sensitive to changes in physiological systems (Costa et al., 2005b; Hu et al., 2009b; Takahashi et al., 2010), including mental disorders such as

Alzheimer's Disease (Abasolo et al., 2006), schizophrenia (Takahashi et al., 2010), the effect of antipsychotic drugs (Takahashi et al., 2010) normal aging (Bruce et al., 2009) and autism spectrum disorders (Bosl et al., 2011).

A comparison of functional network properties using fMRI showed that children and young-adults' brains have similar "small-world" organization at the global level, but differ significantly in hierarchical organization and interregional connectivity (Supekar et al., 2009). The networks measured with fMRI in this study are those that are formed among correlated voxels, which may represent a different spatial scale from that measured by single EEG sensors. Transient or sustained generalized synchronization between EEG sensors is another measurement of functional connectivity in the brain. A simultaneous study of EEG and fMRI signals in patients with a degenerative type of epilepsy showed that in the non-myoclonus state, subtle abnormalities that were detected in EEG signals did not affect fMRI, suggesting that EEG measurements of connectivity may measure different connectivity or may be more sensitive to temporal synchronization that occurs on a time scale less than that of fMRI (on the order of one second).

Neuroscience has made great progress using linear methods of spectral (Fourier) analysis, it is likely that much more information is contained in the complex patterns of brain activity (see, for example, (Bruce et al., 2009). EEG may be under-utilized for clinical applications in neurology and psychiatry (Niedermeyer, 2003; Niedermeyer and Lopes da Silva, 2005) particularly now as new developments in complex systems and multivariate nonlinear time series analysis may allow previously unexplored information to be extracted from EEG signals (Kulisek et al., 2008; Mizuhara et al., 2005; Stam, 2005; Varela et al., 2001). Sensors such as EEGs measure the coordinated electrical response of many neurons to produce time series that reflect the dynamics of this complex system. In order to extract salient information from this data, methods appropriate for analyzing nonlinear time series are required. Although many useful techniques for nonlinear time series and system analysis have been developed in other disciplines (Bosl, 2000; Braha et al., 2006; Elnashaie and Grace, 2007; Holland, 1995; Stauffer, 2006), it is not immediately clear which are most appropriate for neuroscience research. Nevertheless, recent research results are quite promising.

The complexity of EEG signals was found in one study to be associated with the ability to attend to a task and adapt to new cognitive tasks; a significant difference in complexity was found between normal subjects and those with diagnosed schizophrenia (Li et al., 2008). Schizophrenic patients were found to have lower complexity than normal controls in some EEG channels and significantly higher interhemispheric and intrahemispheric cross mutual information values (another measure of signal complexity) than normal controls (Na et al., 2002).

Methods for chaotic signal and phase synchronization analysis arose from a need to rigorously describe physical phenomena that exhibited what was formerly thought to be purely stochastic behavior but was then discovered to represent complex, aperiodic self-organized dynamics (Pikovsky et al., 2001). The analysis of signal complexity and interaction between signals leading to transient synchronization may reveal information about local neural complexity and long-range communication between brain regions (Buzsáki, 2006; Stam, 2005; Varela et al., 2001).

The synchronization patterns of complex networks have been shown to be closely related to the topology of the network (Arenas et al., 2006) and are related to brain connectivity (Sakkalis et al., 2008). EEG signals are believed to derive from pyramidal cells aligned in parallel in the cerebral cortex and hippocampus (Sörnmo and Laguna, 2005), which act as many interacting nonlinear oscillators (Nunez and Srinivasan, 2006). As a consequence of the scale-free network organization of neurons, EEG signals exhibit complex system characteristics reflecting the underlying network topology, including transient synchronization between frequencies, short and long range correlations and cross-modulation of amplitudes and frequencies (Gans et al., 2009). A great deal of information about interrelationships in the nervous system likely remains hidden because the linear analysis techniques currently in use fail even to detect them (Drongelen, 2007).

2.2 Mental processes as emergent phenomena

Much of modern scientific medicine is reductionist, involving a search for ultimate basic causes of disease. This paradigm for scientific research follows naturally from the extraordinary success of physics in the last century, with its search for the fundamental laws of the universe. But that model is giving way to a new vision of the universe as a complex dynamical system, one in which fundamental laws may in fact be emergent properties of the system (Laughlin, 2005).

Emergent properties are those that result from the organization of individual parts and do not exist apart from the organizational whole. The saying that the whole is more than the sum of the parts is a description of an emergent property. The process of cell division, for example, is an emergent process that cannot be explained or studied using quantum mechanics – even though quantum mechanics is a good description of how atoms interact, and a cell is made up of many atoms, each obeying the fundamental physics of quantum theory. Similarly, the difference between a well-written high school essay for a college admissions committee and a Pulitzer Prize winning novel is not to be found in grammar and spelling, even though proper spelling and grammar are essential to the meaning of literature. Literature, genre and metaphor are emergent phenomena, more than words and sentences.

Developing neurotechnology, including devices and analysis methods, may be the most challenging subfield of biomedical engineering because the phenomenon of interest, a mental state or a complex set of behaviors that may indicate a diagnosis of a mental disorder, is an emergent phenomenon. Human behavior is controlled by the brain, which is ultimately a complex network of neurons that transmit electrochemical signals. Thought and behavior cannot be understood or measured by studying neurons (or genes) alone. Psychiatric biomarkers that focus on complex system properties may be the most informative measurements for assessing mental state. We now present a survey of complex system properties that can be computed from time series of brain electrical activity. Neural activity is electrochemical activity. Taking into account degradation due to the skull and scalp and the introduction of noise by the electronic sensors, EEG may be the most direct measurement of brain function that is possible. Thought might be considered an emergent phenomenon of neural electrical activity.

3. Methods

3.1 Univariate measures

Many different methods for computing the complexity of time series have been defined and used successfully to analyze EEG data (Chen et al., 2009; Kuusela et al., 2002). Sample entropy, a measure of nonlinear time series complexity, was significantly higher in certain regions of the right hemisphere in pre-term neonates that received skin-to-skin contact than in those that did not, indicating faster brain maturation (Scher et al., 2009). Sample entropy has also been used as a marker of brain maturation in neonates (de la Cruz et al., 2007) and was found to increase prenatally until maturation at about 42 weeks, then decreased after newborns reached full term (Zhang et al., 2009). A study of the correlation dimension (another measure of signal complexity) of EEG signals in healthy subjects showed an increase with aging, interpreted as an increase in the number of independent synchronous networks in the brain (Stam, 2005).

Intuitively, complexity is associated with structural richness, depth, patterns upon patterns, incorporating correlations over multiple spatio-temporal scales (Costa et al., 2005b). There is no consensus on a definition of complexity, but algorithms have been developed to attempt to give meaning to complexity. In the context of time series analysis, the concept of entropy is relevant. The use of entropy measures to describe the information content of time series began with the publication of Shannon's Mathematical Theory of Communication (Shannon and Weaver, 1949). In an intuitive sense, information is a measure of the difference in uncertainty before and after a measurement. In the context of time series, information is related to the predictability of the series. Entropy is a mathematical function of the probability that the next point in a sequence or time series will be a certain values, given the previous (Baddeley, 2000). Several different entropy measures can be defined algorithmically, including Shannon entropy, spectral entropy, approximate entropy, Lempel-Ziv complexity and sample entropy (Sabeti et al., 2009), each with certain advantages for particular time series characteristics (length, amount of noise, for example). The sample entropy has been used for a number of investigations of physiological signals. Changes in sample entropy appear to correlate with aging and pathological conditions in the context of cardiac health (Bruce et al., 2009; Costa et al., 2008; Norris et al., 2008) and for normal brain development (Zhang et al., 2009) and to distinguish certain mental disorders such as schizophrenia (Sabeti et al., 2009; Takahashi et al., 2010).

The multiscale entropy (MSE) analysis is one method for computing the complexity of a time series that builds on the sample entropy and expands the concept. It has been used to analyze a number of physiological processes (Costa et al., 2005b; Hornero et al., 2009; Norris et al., 2008; Takahashi et al., 2010). The multiscale entropy algorithm incorporates two steps. The first is a coarse-graining procedure that uses successive averaging of a time series to create new coarse-grained time series. For a window size τ, $\tau = 1, 2, \ldots j$, the jth coarse-grain series, y^{τ}_{j}, is computed by averaging non-overlapping windows:

$$y_j^{\tau} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \tag{1}$$

where x_i is the original time series of length N and τ is the scale factor satisfying $1 < \tau < N/\tau$. A schematic illustration of this process is shown in equation 2.

$$\begin{array}{llllllllllllllllll}
\text{Scale 1 :} & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 & s_9 & s_{10} & s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} & \text{original time series} \\
\text{Scale 2 :} & x_1 & & x_2 & & x_3 & & x_4 & & x_5 & & x_6 & & x_7 & & x_8 & & x_k = (s_i + s_{i+1}) / 2 \\
\text{Scale 3 :} & y_1 & & & y_2 & & & y_3 & & & y_4 & & & y_5 & & & & y_k = (s_i + s_{i+1} + s_{i+2}) / 3 \\
\text{Scale 4 :} & z_1 & & & & z_2 & & & & z_3 & & & & z_4 & & & & z_k = (s_i + s_{i+1} + s_{i+2} + s_{i+3}) / 4
\end{array} \tag{2}$$

The coarse-graining method for extracting signal variability on different scales used by (Costa et al., 2005b) seems to be a heuristic procedure without any solid theoretical foundation. Other procedures can be substituted that may be justified on similar grounds. Perhaps the most immediate alternative would be to use the median rather than the mean value in each coarse graining step. This would have the effect of emphasizing the variability of the original signal rather than smoothing out such variability. Another procedure would be to select every kth point from the original series, where k is the desired scale, and use a pre-selected window size to compute an average value at the kth point. A systematic discussion and computational experiments have yet to be done for the coarse graining procedure that is central to the multiscale entropy algorithm.

The second step is to then compute the entropy of each of the coarse-grain time series y^{τ}_{j}, using some entropy measure. The sample entropy is the most common entropy formulation to be used for analyzing physiological signals (Costa et al., 2005b). A useful variation to the original multiscale entropy algorithm uses the modified sample entropy defined in (Xie et al., 2008). The practical effect of using the modified sample entropy is the computed entropy values are more robust to noise and results are more consistent with short time series. In brief, the similarity functions A^m and B^m defined by equations (7) and (9) in (Xie et al., 2008) are computed for each coarse-grained time series defined in equation 1. The modified multiscale entropy (mMSE) is then defined as the series of modified sample entropy values at each of the coarse grain scales. This method was used for complexity analysis of EEG time series as a biomarker for autism risk (Bosl, et al. 2011).

An alternative to the MSE is the scale dependent Lyapunov exponent (SDLE) algorithm described in (Gao, 2007; Gao et al., 2006). This measure of complexity is stable for short, noisy time series and reportedly is able to distinguish a number of different types of chaotic motion, including noise-induced chaos, stochastic oscillations and others, which entropy measures are not able to do. SDLE has not yet been used to analyze EEG signals in young children or infants. The SDLE algorithm is based on following the time evolution of all pairs of vectors in phase space that satisfy a given embedding restriction. This results in a rather straightforward algorithm for computing the SDLE. The SDLE is reportedly better at distinguishing noise from chaotic dynamics in time series. SDLE was shown to be a more effective measure of heart disease than sample entropy and MSE (Hu et al., 2009a). Similarly, SDLE was shown to be more effective in retrospectively identifying changes in EEG signal complexity just prior to the onset of epileptic seizures than MSE, but few other studies of SDLE with EEG time series have been done. This is a potentially promising measure to be investigated further.

3.2 Detecting nonlinearity in time series

Living systems exhibit a fundamental propensity to move forward in time. This property also describes physical systems that are far from an equilibrium state. For example, heat moves in only one direction, from hot to cold areas. In thermodynamics, this property is related to the requirement that all systems must move in the direction of higher entropy. Time irreversibility was found to be a characteristic of healthy human heart electrocardiogram (EKG) recordings and was shown to be a reliable way to distinguish between actual EKG recordings and model EKG simulations (Costa et al., 2008). EKG signals from patients with congestive heart disease were found to have lower time irreversibility indices than healthy patients (Costa et al., 2005a). Interestingly, time irreversibility of EEG signals has been associated with epileptic regions of the brain and this measure has been

proposed as a biomarker for seizure foci (Gautama et al., 2003). Time irreversibility may be used as a practical test for nonlinearity in a time series.
As an illustration, a time irreversibility index (t_{rev}) was computed for different resolutions of the EEG time series using the algorithm of (Costa et al., 2008). The third column of Figure 1 shows t_{rev} for several different linear and nonlinear time series. Of particular note is that only the sine wave time series and both random time series have nearly zero irreversibility indices, while the index for the nonlinear logistic series and the representative EEG signal are both nonzero on all scales shown.

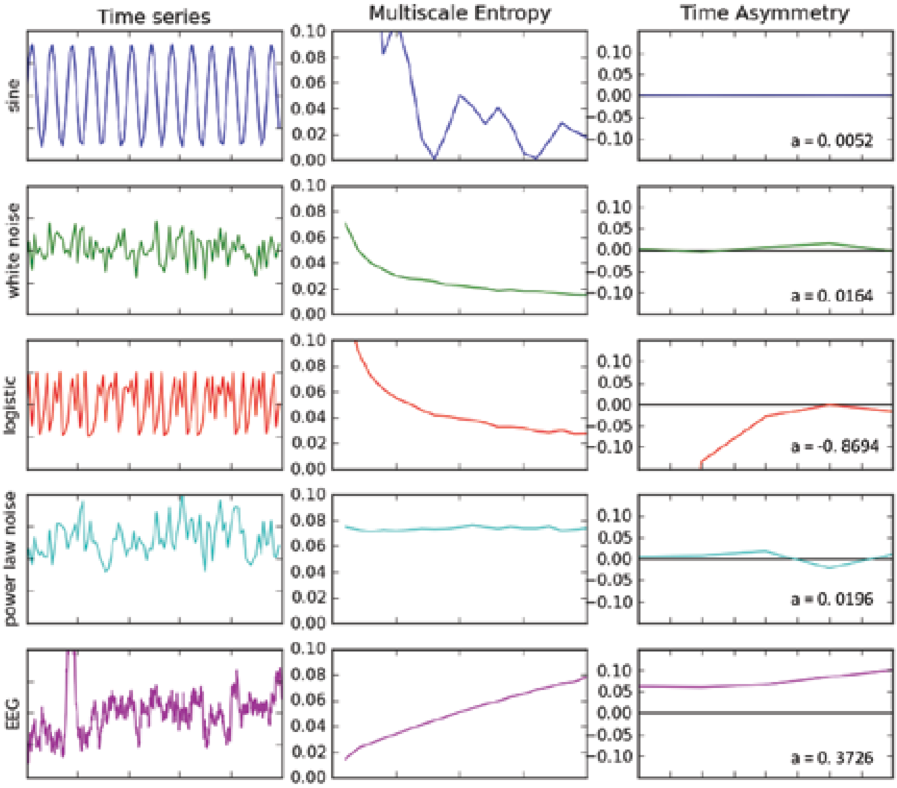


Fig. 1. Characteristics of five different time series are shown. Column 1 shows the time series amplitudes. Column two is the multiscale entropy, where the horizontal axis is the coarse graining scale, from 1 to 20. Column 3 is the multiscale time asymmetry value. The value of a in the lower right corner of the time asymmetry plot is the value of the time asymmetry index summed over scales 1 to 5. A non-zero time asymmetry value is a sufficient condition for nonlinearity of a time series.

After computing multiple resolutions of the EEG time series using the multiscale algorithm shown in equation 2, an estimate of the time irreversibility for each resolution is computed by noting that a symmetric function or time series will have the same number of increments as decrements. That is, the number times $|x_{i+1} - x_i| > 0$ will be approximately the same as the number of times $|x_{i+1} - x_i| < 0$. Thus, an estimate of the time series symmetry (or reversibility) was found by summing increments and decrements and dividing by the length of the series. A reversible time series will have a value of zero. For a series of 5000 points, as used in figure 1, $t_{rev} > 0.1$ is a significant indicator of irreversibility and thus of nonlinearity (Schreiber and Schmitz, 1997). This information is used only to indicate that nonlinear information is contained in the EEG time series that is not used in linear analysis methods, suggesting that the MSE may contain more diagnostically useful information than power spectra analysis alone.

Additional methods for characterizing nonlinear signals may be derived from recurrence plot analysis, to be discussed separately below. Communications and electrical engineers may be especially well suited to research in analyzing brain activity and applying methods from communication signals analysis to find information that can be correlated to behavioral and cognitive assessment data. Integration of new results from both nonlinear time series analysis and complex network research may prove to be a fruitful approach for engineers interested in finding patterns in neural activity that are correlated to complex behavioral patterns that psychiatrists and psychologists use to characterize mental health.

3.3 Synchronization

While signal complexity is a property of a single time series or EEG channel, transient synchronized activity is a measure of the interaction between different channels and an indication of communication and coordination between different brain regions. Synchronization may be used as a marker for diagnosing underlying mental disorders that involve aberrant long-range connectivity in the brain and may also reveal causal mechanisms (Whittington, 2008). The complexity of synchronization patterns appears to change during network development and reflects different neural wiring schemes and levels of cluster organization (Fuchs et al., 2007).

The synchronization patterns of complex networks have been shown to be closely related to the topology of the network (Arenas et al., 2006) and are related to brain connectivity (Sakkalis et al., 2008). Synchronization between sensors is an indicator of connectivity between brain regions on a scale commensurate with the sensor spacing. EEG signals are believed to derive from pyramidal cells aligned in parallel in the cerebral cortex and hippocampus (Sörnmo and Laguna, 2005), which act as many interacting nonlinear oscillators (Nunez and Srinivasan, 2006). Synchronization between gamma activity (typically defined to be 30-50 Hz range) is believed to be involved in long-range communication between brain regions. A possible link between gamma activity and the hemodynamic response measured by fMRI was found in a study of auditory response. Distinct activations in the gamma frequency range were found in subcortical structures, including the anterior cingulate cortex (ACC) and thalamus (Mulert et al., 2007).

Synchronized oscillations are transiently stable, thus form and decay rapidly. Synchrony can result from a common input oscillator, such as in cardiac synchronization. It can also be an emergent, self-organized phenomenon that is related to the network structure itself. The latter is of particular relevance to the search for psychiatric biomarkers that are associated with complex behaviors. The complexity of synchronization patterns appears to change during network development and reflects different neural wiring schemes and levels of cluster organization (Fuchs et al., 2007). It is thus reasonable to suppose that the developing brain will show different but characteristic synchronization patterns at different developmental stages. While the fundamental neurophysiological correlates of these patterns may be difficult to ascertain, they nevertheless may serve as a marker for normal and abnormal brain functional development. The emergence of a "social brain network" during early childhood was found in a study of evoked response potentials (ERPs) in 3-, 4-, and 12-month-old infants viewing faces of different orientation and direction of eye gaze (Johnson et al., 2005), suggesting a particular pattern of brain connectivity that develops in early childhood. Default mode networks (DMNs) found in adults and (negatively) associated with particular cognitive and sensorimotor activities were lacking in a study of (premature) 3 month old infants (Fransson et al., 2007). However, resting state networks have recently been shown to emerge in the first year of life, suggesting development of brain networks and their potential disruption in neurodevelopmental disorders (Supekar et al.; Uddin et al., 2010).

There is no consensus at this time on the best methods for determining nonlinear synchronization in neurological data and a number of different algorithms have been proposed (Kreuz et al., 2007; Sakkalis et al., 2009). Although strong signal synchronization would likely be detected by any of several nonlinear synchronization measures, two measures that are based on different algorithmic approaches are chosen here (see (Sakkalis et al., 2009) for more thorough discussion of some of the relative merits of each).

Two methods outlined here for **c**omputing bivariate synchronization matrix are: (1) the synchronization likelihood (SL) method (Montez et al., 2006); and (2) the instantaneous cross modulation from the circular phase Hilbert transform-based synchronization index (HI), which is robust to signal noise and short time windows (Gans et al., 2009).

Synchronization indices may be searched for correlation in each frequency bands using centered moving averages (Bashan et al., 2008). This approach will find weak or strong correlations with time lags. For each pair, the relative phase index can be computed and stored in a correlation matrix. At each time, some channels may be synchronized. A clustering algorithm can be applied to all channels at a single (averaged) time segment. The result is analogous to gene expression profile clustering (Ramoni et al., 2002). Statistical significance of clusters can be determined by assigning a numerical label to each channel involved in a cluster and the fractional overlap of clusters in different individuals computed. Synchronized clusters may also exhibit very low frequency oscillations, with frequencies of 1 to 0.01 hertz. These have been found in fMRI studies of default mode networks (Broyd et al., 2009; Greicius et al., 2008; Uddin et al., 2010).

Synchronization likelihood (SL) is a method based on the assumption that neurons are highly nonlinear devices, hence methods from chaotic dynamical systems may effectively capture the relevant dynamics of the system (Sakkalis et al., 2009). It is an unbiased generalized synchronization method that relies on detection of simultaneously occurring patterns that may differ in two time series. A method for automatically computing all but two user parameters for the SL algorithm has been developed and will be used here (Montez et al., 2006).

Instantaneous cross modulation (synchronization) of EEG channels can be computed using the Hilbert transform method (Gans et al., 2009). This method is robust to noise and detects synchronization across all frequency bands. The n:m cyclic relative phase index $y_{1,2}$ between two signals, $f_1(t)$ and $f_2(t)$, at a specific time t is computed over a time interval using a sliding window as:

$$\Psi_{1,2}^{n,m}(t) = \left| n\phi_1(t) - m\phi_2(t) \right|, \mod 1 \tag{3}$$

where f(t)=arctan(H(y)/y) and H(y) is the Hilbert transform of the time series y. This approach is stable for nonstationary data (Gans et al., 2009), which is appropriate for our data. The mod 1 term ensures that significant phase differences will be detected even in the presence of noise-induced phase jumps. In most cases n=m=1 is commonly assumed, though cross correlation of signals with n not equal m will be used here. Two signals are defined to be synchronized when $\Psi_{1,2}$ is less than a specified constant. This algorithm is stable for nonstationary data and will detect synchronization without the need to distinguish between noise and chaos (Gans et al., 2009). A sliding window will be used to compute sync over 5 minutes.

A number of methods have been used for determining synchronization in neurological data and a number of different algorithms have been proposed (Kreuz et al., 2007; Sakkalis et al., 2009). Although useful, many of these methods have difficulties with nonstationary, nonlinear signals and either fail to find true synchrony or introduce spurious synchronization (Fine et al., 2010). Spurious synchronization due to volume conduction effects can be removed by applying a spatial algorithm to ICA decomposition (Hironaga and Ioannides, 2007). ICA eliminates volume conduction effects while maintaining the same time resolution, thus still allowing generalized synchronization to be computed.

As Fourier spectrum can only give meaningful interpretation to linear and stationary processes, its application to data from nonlinear and nonstationary processes is problematical. A relatively new method for extracting *instantaneous* phase and frequency information from both linear and nonlinear, chaotic signals is the Huang-Hilbert transform (Huang and Wu, 2005; Huang et al., 2009). Determination of instantaneous phase and frequency is usually accomplished using the Hilbert transform method (Kreuz et al., 2007). However, this is only appropriate for monofrequency analytic signals that have a single center of rotation in the complex plane. The Empirical Mode Decomposition (EMD) introduced by Huang makes no assumptions about linearity. The EMF decomposes a nonlinear, nonstationary time series into adaptively determined characteristic time scales of each of the components (Huang et al., 2009). These component functions are termed intrinsic mode functions (IMF) and are analogous to Fourier components in a traditional linear decomposition. The IMFs computed using the empirical mode transform (EMF) have the property of a single center of rotation in the complex plane, ideally satisfying the requirements for the application of the Hilbert transform to determine instantaneous phase and frequency (Fine et al., 2010; Huang et al., 2009).

After computation of IMFs for each EEG channel, the IMF components with the highest power will be used to compute an instantaneous phase coherence matrix, R, using a sliding window. Following (Bialonski and Lehnertz, 2006), R is computed:

$$R_{ij} = \left(\frac{1}{w} \sum_{t=0}^{w-1} e^{i(\phi_{it} - \phi jt} \right) \text{, mod 1} \tag{4}$$

where w is the number of time samples in the time series segment or window, i and j designate the channel number (or the IMF component of the channel) and f(t)=arctan(H(y)/y) and H(y) is the Hilbert transform of the IMF component. The mod 1 term ensures that significant phase differences will be detected even in the presence of noise-induced phase jumps. The Hilbert transform obtains the best fit of a sinusoid to each IMF at every point in time, identifying an *instantaneous frequency* (IF), along with its associated *instantaneous amplitude* (IA). The IF and IA provide a time-frequency decomposition of the data that is highly effective at resolving non-linear and transient features. This algorithm is stable for nonstationary data and will detect synchronization without the need to distinguish between noise and chaos (Gans et al., 2009). An example of bivariate synchronization between two EEG sensors in the right medical parietal region is shown in figure 2. The synchronization likelihood in this example was computed using the only the first three IMFs from each sensor, without searching for cross band synchronization.

To identify synchronized clusters of EEG channels, a method based on an eigenvector space method, using eigenvalues of R, can be used, following the algorithm developed and applied in (Allefeld and Bialonski, 2007; Bialonski and Lehnertz, 2006; Fine et al., 2010). The outcome of this algorithm will be synchronized clusters of EEG channels. These may be mapped onto scalp plots and the identified clusters compared to default mode networks that have been identified in young children (Sauseng and Klimesch, 2008; Supekar et al., 2009; Supekar et al., 2010). It will be of particular interest to determine if synchronization clusters are significantly correlated with functional networks in the brain and are biomarkers of abnormalities in brain network function (Assaf et al., 2010; Kennedy and Courchesne, 2008). To date, most research on functional brain networks, including the default mode network, has relied on functional MRI. Networks determined by fMRI reflect only the hemodynamic or metabolic response of neurons (Power et al., 2010). This can be considered a kind of amplitude correlation but not true synchronization of brain regions.

If synchronization of electrical activity can be shown to be an alternative measure of brain network activity it would open up much more exploration of the role of brain networks in cognitive activity, brain computer interfaces and neuropsychiatric disorders. Aberrations to default mode networks have been implicated in a number of brain disorders (Broyd et al., 2009) including post traumatic stress syndrome (Daniels et al., 2010), social phobias (Gentili et al., 2009), depression (Sheline et al., 2009), ADHD (Uddin et al., 2008), autism (Di Martino et al., 2009), and schizophrenia (Lagioia et al., 2010). fMRI is far too expensive to be used routinely as a clinical screening and monitoring tool. Yet the apparent widespread role of synchronized brain networks in many neuropsychiatric disorders suggests that a less expensive and easy to administer technology for analyzing brain networks would be widely useful in clinical practice.

3.3 Recurrence quantitative analysis

Several univariate measures of time series complexity and a number of approaches for computing the degree of synchronization between signals have been used to analyze EEG data. Applications of these methods to psychiatric care and mental health continue to show promise. A more general framework for characterizing the dynamics of complex systems may be to construct recurrence plots (Marwan et al., 2007) and compute quantitative properties. The idea to use recurrence plots as a representation of complex system dynamics was first proposed by (Eckmann et al., 1987) in the late 1980s. The original tool presented a graphical means for visualizing differences in system dynamics. Methods for quantifying the small scale structures in recurrence plots were devised and shown to be capable of revealing system parameters and transitions that are not easily obtained by other methods (Marwan et al., 2002; Zbilut et al., 2002). Some dynamical parameters, such as K2 entropy and mutual information can also be derived from recurrence plots without RQA methods by computing the distribution of line lengths in the recurrence plot (Marwan et al., 2007). Readers are referred to the references for reviews of this unifying approach to nonlinear systems analysis.

(Schinkel et al., 2007) demonstrated that a single measure from RQA analysis could detect a change in the N400 response in single trials when subjects were presented with an oddball task, suggesting that RQA may be a sensitive measure of transient brain states. Few studies have been done using RQA to determine more stable brain functional characteristics. This may be a promising new field for research on EEG biomarkers of psychiatric disorders.

4. Clinical applications

4.1 Infant brain development and autism spectrum disorders

Autism spectrum disorder (ASD) constitutes a heterogeneous developmental syndrome that is characterized by a triad of impairments that affect social interaction, communication skills, and a restricted range of interests and activities (APA, 2000), with highly variable outcomes. Studies have consistently shown that early intervention leads to better long-term outcomes. But early intervention is predicated on early detection. Behavioral measures have thus far proven ineffective in diagnosing autism before about 18 months of age, in part because the behavioral repertoire of infants is so limited. Neural development may precede overt behavioral observations and thus provide an earlier marker for emerging autistic behaviors. Yet, measuring functional brain development is difficult because few noninvasive methods are available for infants and it is not clear what features to measure that are biomarkers of normal development. As discussed above, multiscale entropy computed from EEG time series has been shown to be a particularly informative analysis tool for physiological signals.

Complex mental disorders such as autism are associated with abnormal brain connectivity that may vary between different regions and different scales (Noonan et al., 2009). Estimation of changes in neural connectivity might be an effective diagnostic biomarker for abnormal connectivity development that leads to ASD behaviors. The electrical signals produced by neural networks are believed to contain information about the network structure (Raghavendra et al., 2009; Stam, 2005; Zavaglia et al., 2008). The physics of complex networks suggests that system invariants will be encoded in the time series produced and measured by EEG sensors (Fuchs et al., 2007; Gao and Jin, 2009). Multiscale entropy (MSE) discussed above is one invariant measure of system dynamics that has been shown to particularly sensitive to changes in physiological systems (Costa et al., 2005b; Hu et al., 2009b; Takahashi et al., 2010), including mental disorders such as Alzheimer's Disease (Abasolo et al., 2006), schizophrenia (Takahashi et al., 2010), the effect of antipsychotic drugs (Takahashi et al., 2010) and normal aging (Bruce et al., 2009). Preliminary results suggest that MSE may be a biomarker for autism endophenotypes (Bosl et al., 2011) and may provide an earlier diagnosis than behavioral assessments.

MSE values were computed for 79 different infants: 46 at high risk for ASD (hereafter referred to as HRA) based on having an older sibling with a confirmed diagnosis of ASD and 33 controls, defined on the basis of a typically developing older sibling and no family history of neurodevelopmental disorders. The study participants were part of an on-going longitudinal study and for this analysis visits were evaluated at regular intervals. Infants were seated on their mothers' laps in a dimly lit room while a research assistant engaged their attention by blowing bubbles. This procedure was followed to limit the amount of head movement made by the infant that would interfere with the recording process. Continuous EEG was recorded with a 64-channel Sensor Net System. The data were amplified, filtered (band pass 0.1-100.0 Hz), and sampled at a frequency of 250 Hz. Data were manually reviewed to remove sections with obvious artifacts and continuous, clean 20 second segments were identified to compute alert, resting state MSE values.

Results shown in figure 2 were computing by averaging all channels in a given region, including left/right hemispheres, left/right frontal and total scalp MSE values. The plots are derived by averaging regional MSE values for all infants at a given age. These reveal a significant difference between high risk and typically developing infants (Bosl et al., 2011). The data suggest that there are not only significant differences between the high risk group and typically developing controls, but MSE follows characteristic trajectories during development. To our knowledge, these preliminary results are the first demonstration of a complex systems analysis of EEG data for biomarkers in infants at risk for a complex neurodevelopmental disorder. More details about this study and the use of machine learning to discover diagnostically significant patterns in MSE data can be found in (Bosl et al., 2011).

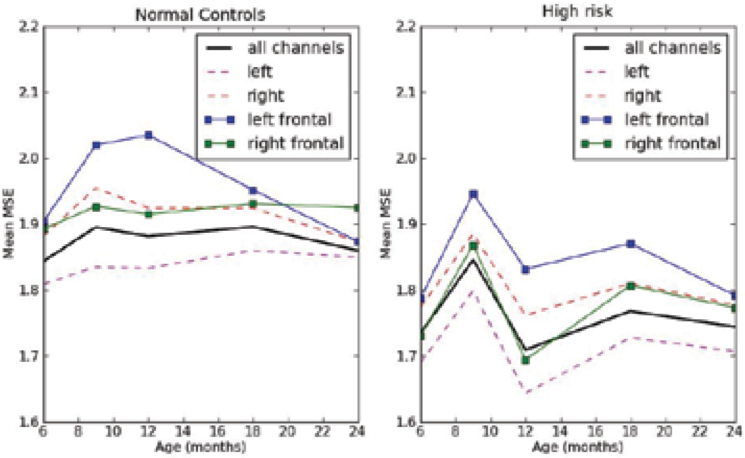


Fig. 2. The change in mean MSE over all channels is shown for each age. Averaging over all channels reveals that in general MSE is higher than the normal controls than in the high risk group, but regional differences cannot be seen.

The trajectory of the curves between 6 and 12 months in figure 2 appear to be as informative as information at any specific age. This is a period of important changes in brain function that are foundational for the emergence of higher level social and communicative skills that are at the heart of ASD. A key characteristic of autism is lack of social perception; autistic spectral disorders are also associated with abnormal brain connectivity (Belmonte et al., 2004; Kleinhans et al., 2008). A number of major cognitive milestones typically occur beginning at around the age 9 months and perhaps earlier in girls. These include joint perception (Behne et al., 2005) and loss of the ability to discriminate certain categories of faces (Pascalis et al., 2002). The latter developments are especially significant because they reveal how *socially-grounded experiences* influence changes in the neurocognitive mechanisms that underlie speech and face processing. In a prospective study, (Ozonoff et al., 2010) found that social communicative behaviors in infants that later developed ASD declined dramatically between 6 and 18 months when compared to typically developing infants. The results shown here may indicate that MSE is sensitive to neural correlates underlying the observed social and cognitive developments.

Currently there are no objective medical tests for diagnosing autism. According to the American Academy of Pediatrics (AAP), data strongly suggest better outcomes in children whose conditions are diagnosed early and participate in early intervention programs. Children who are diagnosed early and given intensive behavioral intervention can often be mainstreamed and live normal, productive lives. The AAP recommends that *all* children be screened for autism at 18 and again at 24 months. Unfortunately, behavioral assessment is time consuming and requires specialized training, both of which pose a problem for routine screening. Because atypical brain development is likely to precede abnormal behavior by months or even years, a critical developmental window for early intervention may be missed if diagnosis is based entirely on a behavioral phenotype. If reliable methods can be developed to lower the age of diagnosis, and if insight is gained into the biological mechanisms that underlie the disorder, it may be possible to develop intervention strategies that can be implemented during the first year of life. Complex systems analysis of EEG may enable a relatively inexpensive and reliable way to estimate the risk of developmental disorders within the first year of life.

4.2 Schizophrenia

Early studies of signal synchronization in patients diagnosed with schizophrenia used a *linear* measure of complexity based on global correlation of all channels (Wackermann et al., 1993) and synchronization based on linear correlation between amplitudes (Saito et al., 1998). Although these methods will necessarily miss nonlinear signal characteristics, they found significant differences between patients and normal controls that were interpreted to reflect decreased coordination between different brain regions, a characteristic of schizophrenia (Uhlhaas et al., 2008; Whittington, 2008).

One interpretation of biological complexity is that it reflects a systems' ability to quickly adapt and function in a changing environment (Costa et al., 2005b). The complexity of EEG signals was found in one study to be associated with the ability to attend to a task and adapt to new cognitive tasks; a significant difference in complexity was found between normal subjects and those with diagnosed schizophrenia (Li et al., 2008). Schizophrenic patients were found to have lower complexity than normal controls in some EEG channels and significantly higher interhemispheric and intrahemispheric cross mutual information values (another measure of signal complexity) than the normal controls (Na et al., 2002).

While signal complexity is a property of a single time series or EEG channel, transient synchronized activity is a measure of the interaction between different channels and an indication of communication and coordination between different brain regions. Synchronization may be used as a marker for diagnosing underlying mental disorders such as schizophrenia and may also reveal causal mechanisms (Whittington, 2008). The complexity of synchronization patterns changes during network development and reflects different neural wiring schemes and levels of cluster organization (Fuchs et al., 2007).

Abnormalities in phase synchronization between multiple bands have been found to be sensitive biomarkers for mental dysfunction in schizophrenic patients (Uhlhaas et al., 2008; Whittington, 2008). Unfortunately, similar abnormalities in synchronous activity have been found associated with a number of other mental disorders (Uhlhaas et al., 2008), so further research is required to discover if more refined patterns of synchrony exist for discriminating different disorders or subtypes. A developmental perspective may be useful here. For example, while many attempts to correlate cortical thickness with intelligence have failed, recent research demonstrated that specific characteristic growth *trajectories* of cortical thickness from infancy to early teen years were highly correlated with above or below average intelligence (Shaw et al., 2006), suggesting that growth curves of brain function may contain more information than any combination of measurements at one specific age. This would require routine brain measurements become part of the medical record and algorithms that recognize abnormal trends would need to be used to interpret data after regular checkups.

5. Neurotechnology for global mental health

Mental healthcare in developing nations has long been overlooked by organizations concerned with global health needs. A number of reasons can be cited for this, but one significant factor is the difficulty of accurately diagnosing and classifying mental disorders and the relatively intensive need for human expertise to administer assessments and care. Precision healthcare is a term coined by Clayton Christensen to describe objective medical diagnosis and prescription that is enabled by new technologies. Innovative neurotechnologies that enable basic mental diagnosis to be administered in a cost effective manner in clinics that may be staffed by technicians is urgently needed in developing nations. Even if the diagnostic precision is not as high as might be possible in big city hospital in the developed world, low-cost basic care would serve to alleviate a great deal of personal suffering as well as the economic burden imposed.

The need for neurotechnology to enable precision mental healthcare has never been greater. The developed world is spending greater and greater amounts of money on healthcare and costs threaten their economic vitality. Developing nations will benefit tremendously from even basic neurological and mental healthcare.

While considerable resources have been devoted to finding cures for infectious diseases such as HIV and malaria, relatively little attention has been given to the neurological and cognitive effects of these infections on young patients that recover from the acute infection. A majority of HIV infected patients manifest HIV-associated neurocognitive disorders despite receiving highly active antiretroviral therapy (Van Rie et al., 2009). The developing brains of children who have acquired HIV through vertical infection are particularly vulnerable to central nervous system insults either directly, through secondary infections or from side effects of antiretroviral therapies (LLorente et al., 2009). For example, school age children with perinatally acquired HIV infections receiving antiretroviral therapy displayed lower cognitive function than either HIV affected or HIV unaffected children. Although the antiretroviral therapy reduced HIV symptoms, the cognitive deficits were not improved (Puthanakit et al., 2010).

These clinical observations inform the need to obtain a better understanding of how and when HIV affects cognitive function and to develop novel therapeutic drug candidates to prevent or interfere with progression of HIV associated neurocognitive impairment (Robertson et al., 2010). Some treatments for the primary or secondary infections, may be more harmful to vulnerable brain tissue in children that have been compromised by HIV infection than in adults (LLorente et al., 2009). Research on the effects of HIV infection on the developing brain is a topic of increasing importance as new antiretroviral treatments become more widely available causing many more HIV-1 infections to be treated as a chronic condition to be managed rather than as a fatal disease (Abubakar et al., 2008). Few longitudinal studies have been done of HIV infected children in developing regions who have been receiving antiretroviral therapy. The methods for measuring and analyzing brain electrical activity developed here may help to fill the need for objective, cross-cultural measures of cognitive function to estimate the effects of HIV on brain development in children.

It is hoped that neuroengineers will adopt methods outlined in this chapter to discover biomarkers for monitoring the progression of cognitive development and demonstration their utility in the context of HIV associated neurocognitive disability (HAND). Protocols are needed for early assessment and diagnosis of mental and cognitive impairment due to HIV infection, secondary infections and drug therapy. These may also be used as a clinical tool for widespread screening of HIV infected or exposed infants that may enable developmental cognitive impairments to be diagnosed before other HIV symptoms are apparent. Such tools would represent a major step forward in the effort to understand the long-term impact of chronic HIV infection and other infections on cognitive development in children and to provide a new approach for monitoring the cognitive effects of long term management of HIV as a chronic condition in developing regions of the world.

6. Conclusion

Neuroscience has been called the next frontier science after the breakthroughs achieved by physics in the mid-twentieth century and genetics and biophysics at the end of the 20th century and early 21st century. The brain is of great scientific interest because of the emergent phenomenon of thought that arises from the vast complexity of this organ. While the question of how the brain thinks is likely to remain a challenging scientific and philosophical topic for a long time, the need for practical innovations to meet the healthcare needs of the world is more urgent. Two issues drive this need at the present time. First, of course, is the tremendous amount of suffering that is the result of brain injury and mental disease. Mental disorders in particular are a difficult challenge for medicine because in most cases there is not a single cause that can easily be diagnosed and “cured”. Rather, the perspective throughout this chapter has been that mental disorders are complex emergent phenomena and must therefore be analyzed as such.

This is an exciting and promising time for clinical neuroscience research due to the convergent development of several technologies at this time in history. Although EEG equipment has been around since the 1920s, the creation of inexpensive, wireless, high quality EEG headsets by innovative companies such as Emotiv Systems and Neurosky opens up research possibilities with EEG equipment for many more neuroscientists. As the cost comes down and quality goes up, these tools become practical for routine clinical use. The rapid development of the physics of complex networks and systems, as well as mathematical methods for analyzing complex (chaotic) time series that are produced by complex networks enables new information from the signals measured with inexpensive EEG devices to be extracted and studied as markers for neural correlates of behavior. Finally, the continuing advance of pattern recognition algorithms from the artificial intelligence community enables more subtle pattern-markers to be discovered in complex EEG signal features.

The high cost of healthcare threatens the economies of developed nations and prevents many people in developing regions from obtaining the psychiatric care that they need. On both of these fronts, innovative neurotechnologies for early detection and monitoring of mental disorders are urgently needed. The ready availability of mobile communication devices introduces a platform for making information technology based mental healthcare available to many people that have no access at this time. Integration of EEG devices and complex systems methods with clinical decision support tools and mobile device-based health records promises both greatly improved neuropsychiatric healthcare and lower costs. It is hoped that the methods introduced in this chapter will inspire a new generation of bioengineers concerned with mental health to create new tools for clinical neuroscience while closing the gap between neurology and psychiatry.

7. Acknowledgements

This research was supported by grants from the National Institute of Mental Health and the Simons Foundation.

8. References

Abasolo, D., Hornero, R., Espino, P., Alvarez, D., and Poza, J. (2006). Entropy analysis of the EEG background activity in Alzheimer's disease patients. Physiol Meas *27*, 241-253.

Abubakar, A., Van Baar, A., Van de Vijver, F.J., Holding, P., and Newton, C.R. (2008). Paediatric HIV and neurodevelopment in sub-Saharan Africa: a systematic review. Trop Med Int Health *13*, 880-887.

Allefeld, C., and Bialonski, S. (2007). Detecting synchronization clusters in multivariate time series via coarse-graining of Markov chains. Phys Rev E Stat Nonlin Soft Matter Phys *76*, 066207.

Arenas, A., Diaz-Guilera, A., and Perez-Vicente, C.J. (2006). Synchronization reveals topological scales in complex networks. Phys Rev Lett *96*, 114102.

Assaf, M., Jagannathan, K., Calhoun, V.D., Miller, L., Stevens, M.C., Sahl, R., O'Boyle, J.G., Schultz, R.T., and Pearlson, G.D. (2010). Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients. Neuroimage *53*, 247-256.

Baddeley, R. (2000). Information Theory and the Brain (New York, Cambridge University Press).

Barabasi, A.L. (2009). Scale-free networks: a decade and beyond. Science *325*, 412-413.

Bashan, A., Bartsch, R., Kantelhardt, J.W., and Havlin, S. (2008). Comparison of detrending methods for fluctuation analysis. Physica A *387*, 5080-5090.

Bassett, D.S., and Bullmore, E. (2006). Small-world brain networks. Neuroscientist *12*, 512-523.

Bassett, D.S., and Bullmore, E.T. (2009). Human brain networks in health and disease. Curr Opin Neurol *22*, 340-347.

Behne, T., Carpenter, M., Call, J., and Tomasello, M. (2005). Unwilling versus unable: infants' understanding of intentional action. Dev Psychol *41*, 328-337.

Belmonte, M.K., Allen, G., Beckel-Mitchener, A., Boulanger, L.M., Carper, R.A., and Webb, S.J. (2004). Autism and abnormal development of brain connectivity. J Neurosci *24*, 9228-9231.

Bialonski, S., and Lehnertz, K. (2006). Identifying phase synchronization clusters in spatially extended dynamical systems. Phys Rev E Stat Nonlin Soft Matter Phys *74*, 051909.

Bosl, W.J. (2000). Modeling Complex Crustal Processes. In GeoComplexity and the Physics of Earthquakes, J.R. Rundle, D.L. Turcotte, and W. Klein, eds. (Washington, D.C., American Geophysical Union).

Bosl, W.J., Tager-Flusberg, H., and Nelson, C.A. (2011). EEG Complexity as a Biomarker for Autism Spectrum Disorder. BMC Medicine *9*.

Braha, D., Minai, A., and Bar-Yam, Y. (2006). Complex engineered systems : science meets technology (Berlin ; New York, Springer).

Broyd, S.J., Demanuele, C., Debener, S., Helps, S.K., James, C.J., and Sonuga-Barke, E.J. (2009). Default-mode brain dysfunction in mental disorders: a systematic review. Neurosci Biobehav Rev *33*, 279-296.

Bruce, E.N., Bruce, M.C., and Vennelaganti, S. (2009). Sample entropy tracks changes in electroencephalogram power spectrum with sleep state and aging. J Clin Neurophysiol *26*, 257-266.

Buzsáki, G. (2006). Rhythms of the brain (Oxford ; New York, Oxford University Press).

Chen, W., Zhuang, J., Yu, W., and Wang, Z. (2009). Measuring complexity using FuzzyEn, ApEn, and SampEn. Med Eng Phys *31*, 61-68.

Costa, M., Goldberger, A.L., and Peng, C.K. (2005a). Broken asymmetry of the human heartbeat: loss of time irreversibility in aging and disease. Phys Rev Lett *95*, 198102.

Costa, M., Goldberger, A.L., and Peng, C.K. (2005b). Multiscale entropy analysis of biological signals. Phys Rev E Stat Nonlin Soft Matter Phys *71*, 021906.

Costa, M.D., Peng, C.K., and Goldberger, A.L. (2008). Multiscale analysis of heart rate dynamics: entropy and time irreversibility measures. Cardiovasc Eng *8*, 88-93.

Cowan, W.M., and Kandel, E.R. (2001). Prospects for neurology and psychiatry. Jama *285*, 594-600.

Craddock, R.C., Holtzheimer, P.E., 3rd, Hu, X.P., and Mayberg, H.S. (2009). Disease state prediction from resting state functional connectivity. Magn Reson Med *62*, 1619-1628.

Daniels, J.K., McFarlane, A.C., Bluhm, R.L., Moores, K.A., Clark, C.R., Shaw, M.E., Williamson, P.C., Densmore, M., and Lanius, R.A. (2010). Switching between executive and default mode networks in posttraumatic stress disorder: alterations in functional connectivity. J Psychiatry Neurosci *35*, 258-266.

de la Cruz, D.M., Manas, S., Pereda, E., Garrido, J.M., Lopez, S., De Vera, L., and Gonzalez, J.J. (2007). Maturational changes in the interdependencies between cortical brain areas of neonates during sleep. Cereb Cortex *17*, 583-590.

Di Martino, A., Ross, K., Uddin, L.Q., Sklar, A.B., Castellanos, F.X., and Milham, M.P. (2009). Functional brain correlates of social and nonsocial processes in autism spectrum disorders: an activation likelihood estimation meta-analysis. Biol Psychiatry *65*, 63-74.

Dornhege, G. (2007). Toward brain-computer interfacing (Cambridge, Mass., MIT Press).

Douw, L., de Groot, M., van Dellen, E., Heimans, J.J., Ronner, H.E., Stam, C.J., and Reijneveld, J.C. (2010). 'Functional connectivity' is a sensitive predictor of epilepsy diagnosis after the first seizure. PLoS One *5*, e10839.

Drongelen, W.v. (2007). Signal processing for neuroscientists : introduction to the analysis of physiological signals (Burlington, Mass., Academic Press).

Eckmann, J., Kaphorst, S.O., and Ruelle, D. (1987). Recurrence plots of dynamical systems. Europhysics Letters *5*, 973-977.

Elnashaie, S., and Grace, J.R. (2007). Complexity, bifurcation and chaos in natural and man-made lumped and distributed systems. Chemical Engineering Science *62*, 3295-3325.

Fine, A.S., Nicholls, D.P., and Mogul, D.J. (2010). Assessing instantaneous synchrony of nonlinear nonstationary oscillators in the brain. J Neurosci Methods *186*, 42-51.

Fransson, P., Skiold, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., and Aden, U. (2007). Resting-state networks in the infant brain. Proc Natl Acad Sci U S A *104*, 15531-15536.

Fuchs, E., Ayali, A., Robinson, A., Hulata, E., and Ben-Jacob, E. (2007). Coemergence of regularity and complexity during neural network development. Dev Neurobiol *67*, 1802-1814.

Gans, F., Schumann, A.Y., Kantelhardt, J.W., Penzel, T., and Fietze, I. (2009). Cross-modulated amplitudes and frequencies characterize interacting components in complex systems. Phys Rev Lett *102*, 098701.

Gao, J. (2007). Multiscale analysis of complex time series : integration of chaos and random fractal theory, and beyond (Hoboken, N.J., Wiley-Interscience).

Gao, J.B., Hu, J., Tung, W.W., and Cao, Y.H. (2006). Distinguishing chaos from noise by scale-dependent Lyapunov exponent. Phys Rev E Stat Nonlin Soft Matter Phys *74*, 066204.

Gao, Z., and Jin, N. (2009). Complex network from time series based on phase space reconstruction. Chaos *19*, 033137.

Gautama, T., Mandic, D.P., and Van Hulle, M.M. (2003). Indications of nonlinear structures in brain electrical activity. Phys Rev E Stat Nonlin Soft Matter Phys *67*, 046204.

Gentili, C., Ricciardi, E., Gobbini, M.I., Santarelli, M.F., Haxby, J.V., Pietrini, P., and Guazzelli, M. (2009). Beyond amygdala: Default Mode Network activity differs between patients with social phobia and healthy controls. Brain Res Bull *79*, 409-413.

Greicius, M.D., Kiviniemi, V., Tervonen, O., Vainionpaa, V., Alahuhta, S., Reiss, A.L., and Menon, V. (2008). Persistent default-mode network connectivity during light sedation. Hum Brain Mapp *29*, 839-847.

Hironaga, N., and Ioannides, A.A. (2007). Localization of individual area neuronal activity. Neuroimage *34*, 1519-1534.

Holland, J. (1995). Hidden Order: How Adaptation Builds Complexity (New York, Addison-Wesley Publishing Co.).

Hornero, R., Abasolo, D., Escudero, J., and Gomez, C. (2009). Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease. Philos Transact A Math Phys Eng Sci *367*, 317-336.

Hu, J., Gao, J., Tung, W.W., and Cao, Y. (2009a). Multiscale analysis of heart rate variability: a comparison of different complexity measures. Ann Biomed Eng *38*, 854-864.

Hu, J., Gao, J., Tung, W.W., and Cao, Y. (2009b). Multiscale Analysis of Heart Rate Variability: A Comparison of Different Complexity Measures. Ann Biomed Eng.

Huang, N.E., and Wu, Z. (2005). An Adaptive Data Analysis Method for nonlinear and Nonstationary Time Series:

The Empirical Mode Decomposition and Hilbert Spectral Analysis (Greenbelt, MD, NASA Goddard Space Flight Center).

Huang, N.E., Wu, Z., Long, S.R., and Arnold, K.C. (2009). On Instantaneous Frequency. Advances in Adaptive Data Analysis *1*, 177-229.

Hyman, S.E. (2007). Can neuroscience be integrated into the DSM-V? Nat Rev Neurosci *8*, 725-732.

Irani, F., Platek, S.M., Bunce, S., Ruocco, A.C., and Chute, D. (2007). Functional near infrared spectroscopy (fNIRS): an emerging neuroimaging technology with important applications for the study of brain disorders. Clin Neuropsychol *21*, 9-37.

Johnson, M.H., Griffin, R., Csibra, G., Halit, H., Farroni, T., de Haan, M., Tucker, L.A., Baron-Cohen, S., and Richards, J. (2005). The emergence of the social brain network: evidence from typical and atypical development. Dev Psychopathol *17*, 599-619.

Kandel, E.R. (1998). A new intellectual framework for psychiatry. Am J Psychiatry *155*, 457-469.

Kandel, E.R., Schwartz, J.H., and Jessell, T.M. (2000). Principles of neural science, 4th edn (New York, McGraw-Hill Health Professions Division).

Kennedy, D.P., and Courchesne, E. (2008). The intrinsic functional organization of the brain is altered in autism. Neuroimage *39*, 1877-1885.

Kleinhans, N.M., Richards, T., Sterling, L., Stegbauer, K.C., Mahurin, R., Johnson, L.C., Greenson, J., Dawson, G., and Aylward, E. (2008). Abnormal functional connectivity in autism spectrum disorders during face processing. Brain *131*, 1000-1012.

Kreuz, T., Mormann, F., Andrzejak, R., Kraskov, A., and Lehnertz, K. (2007). Measuring synchronization in coupled model systems: A comparison of different approaches. Physica D *225*, 29-42.

Kulisek, R., Hrncir, Z., Hrdlicka, M., Faladova, L., Sterbova, K., Krsek, P., Vymlatilova, E., Palus, M., Zumrova, A., and Komarek, V. (2008). Nonlinear analysis of the sleep EEG in children with pervasive developmental disorder. Neuro Endocrinol Lett *29*, 512-517.

Kuusela, T.A., Jartti, T.T., Tahvanainen, K.U., and Kaila, T.J. (2002). Nonlinear methods of biosignal analysis in assessing terbutaline-induced heart rate and blood pressure changes. Am J Physiol Heart Circ Physiol *282*, H773-783.

Lagioia, A., Van De Ville, D., Debbane, M., Lazeyras, F., and Eliez, S. (2010). Adolescent resting state networks and their associations with schizotypal trait expression. Front Syst Neurosci *4*.

Laughlin, R.B. (2005). A different universe : reinventing physics from the bottom down (New York, Basic Books).

Le Van Quyen, M. (2003). Disentangling the dynamic core: a research program for a neurodynamics at the large-scale. Biol Res *36*, 67-88.

Li, Y., Tong, S., Liu, D., Gai, Y., Wang, X., Wang, J., Qiu, Y., and Zhu, Y. (2008). Abnormal EEG complexity in patients with schizophrenia and depression. Clin Neurophysiol *119*, 1232-1241.

LLorente, A.M., Lopresti, C., and Satz, P. (2009). Neurobehavioral and Neurodevelopmental Sequelae Associated with Pediatric HIV Infection. In Handbook of Clinical Child Neuropsychology, C.R. Reynolds, and E. Fletcher-Janzen, eds. (New York, Springer Science + Business Media), pp. 635-669.

Marwan, N., Romano, M.C., Thiel, M., and Kurths, J. (2007). Recurrence plots for the analysis of complex systems. Physics Reports *438*, 237-329.

Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., and Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. Phys Rev E Stat Nonlin Soft Matter Phys *66*, 026702.

Mizuhara, H., Wang, L.Q., Kobayashi, K., and Yamaguchi, Y. (2005). Long-range EEG phase synchronization during an arithmetic task indexes a coherent cortical network simultaneously measured by fMRI. Neuroimage *27*, 553-563.

Montez, T., Linkenkaer-Hansen, K., van Dijk, B.W., and Stam, C.J. (2006). Synchronization likelihood with explicit time-frequency priors. Neuroimage *33*, 1117-1125.

Muehlemann, T., Haensse, D., and Wolf, M. (2008). Wireless miniaturized in-vivo near infrared imaging. Opt Express *16*, 10323-10330.

Mulert, C., Leicht, G., Pogarell, O., Mergl, R., Karch, S., Juckel, G., Moller, H.J., and Hegerl, U. (2007). Auditory cortex and anterior cingulate cortex sources of the early evoked gamma-band response: relationship to task difficulty and mental effort. Neuropsychologia *45*, 2294-2306.

Na, S.H., Jin, S.H., Kim, S.Y., and Ham, B.J. (2002). EEG in schizophrenic patients: mutual information analysis. Clin Neurophysiol *113*, 1954-1960.

Niedermeyer, E. (2003). The clinical relevance of EEG interpretation. Clinical EEG (electroencephalography) *34*, 93-98.

Niedermeyer, E., and Lopes da Silva, F.H. (2005). Electroencephalography : basic principles, clinical applications, and related fields, 5th edn (Philadelphia, Lippincott Williams & Wilkins).

Noonan, S.K., Haist, F., and Muller, R.A. (2009). Aberrant functional connectivity in autism: evidence from low-frequency BOLD signal fluctuations. Brain Res *1262*, 48-63.

Norris, P.R., Stein, P.K., and Morris, J.A., Jr. (2008). Reduced heart rate multiscale entropy predicts death in critical illness: a study of physiologic complexity in 285 trauma patients. J Crit Care *23*, 399-405.

Nunez, P.L., and Srinivasan, R. (2006). Electric fields of the brain : the neurophysics of EEG, 2nd edn (New York, Oxford University Press).

Ozonoff, S., Iosif, A.M., Baguio, F., Cook, I.C., Hill, M.M., Hutman, T., Rogers, S.J., Rozga, A., Sangha, S., Sigman, M., *et al.* (2010). A prospective study of the emergence of early behavioral signs of autism. J Am Acad Child Adolesc Psychiatry *49*, 256-266 e251-252.

Pascalis, O., de Haan, M., and Nelson, C.A. (2002). Is face processing species-specific during the first year of life? Science *296*, 1321-1323.

Percha, B., Dzakpasu, R., Zochowski, M., and Parent, J. (2005). Transition from local to global phase synchrony in small world neural network and its possible implications for epilepsy. Phys Rev E Stat Nonlin Soft Matter Phys *72*, 031909.

Pikovsky, A., Rosenblum, M., and Kurths, J. (2001). Synchronization : a universal concept in nonlinear sciences (Cambridge, Cambridge University Press).

Power, J.D., Fair, D.A., Schlaggar, B.L., and Petersen, S.E. (2010). The development of human functional brain networks. Neuron *67*, 735-748.

Puthanakit, T., Aurpibul, L., Louthrenoo, O., Tapanya, P., Nadsasarn, R., Insee-ard, S., and Sirisanthana, V. (2010). Poor cognitive functioning of school-aged children in thailand with perinatally acquired HIV infection taking antiretroviral therapy. AIDS Patient Care STDS *24*, 141-146.

Raghavendra, B.S., Dutt, D.N., Halahalli, H.N., and John, J.P. (2009). Complexity analysis of EEG in patients with schizophrenia using fractal dimension. Physiol Meas *30*, 795-808.

Ramoni, M.F., Sebastiani, P., and Kohane, I.S. (2002). Cluster analysis of gene expression dynamics. Proc Natl Acad Sci U S A *99*, 9121-9126.

Ravasz, E., and Barabasi, A.L. (2003). Hierarchical organization in complex networks. Phys Rev E Stat Nonlin Soft Matter Phys *67*, 026112.

Robertson, K., Liner, J., Hakim, J., Sankale, J.L., Grant, I., Letendre, S., Clifford, D., Diop, A.G., Jaye, A., Kanmogne, G., *et al.* (2010). NeuroAIDS in Africa. J Neurovirol *16*, 189-202.

Sabeti, M., Katebi, S., and Boostani, R. (2009). Entropy and complexity measures for EEG signal classification of schizophrenic and control participants. Artif Intell Med *47*, 263-274.

Saito, N., Kuginuki, T., Yagyu, T., Kinoshita, T., Koenig, T., Pascual-Marqui, R.D., Kochi, K., Wackermann, J., and Lehmann, D. (1998). Global, regional, and local measures of complexity of multichannel electroencephalography in acute, neuroleptic-naive, first-break schizophrenics. Biol Psychiatry *43*, 794-802.

Sakkalis, V., Doru Giurc Neanu, C., Xanthopoulos, P., Zervakis, M.E., Tsiaras, V., Yang, Y., Karakonstantaki, E., and Micheloyannis, S. (2009). Assessment of linear and nonlinear synchronization measures for analyzing EEG in a mild epileptic paradigm. IEEE Trans Inf Technol Biomed *13*, 433-441.

Sakkalis, V., Tsiaras, V., Michalopoulos, K., and Zervakis, M. (2008). Assessment of neural dynamic coupling and causal interactions between independent EEG components from cognitive tasks using linear and nonlinear methods. Conf Proc IEEE Eng Med Biol Soc *2008*, 3767-3770.

Sauseng, P., and Klimesch, W. (2008). What does phase information of oscillatory brain activity tell us about cognitive processes? Neurosci Biobehav Rev *32*, 1001-1013.

Scher, M.S., Ludington-Hoe, S., Kaffashi, F., Johnson, M.W., Holditch-Davis, D., and Loparo, K.A. (2009). Neurophysiologic assessment of brain maturation after an 8-week trial of skin-to-skin contact on preterm infants. Clin Neurophysiol *120*, 1812-1818.

Schinkel, S., Marwan, N., and Kurths, J. (2007). Order patterns recurrence plots in the analysis of ERP data. Cogn Neurodyn *1*, 317-325.

Schreiber, T., and Schmitz, A. (1997). Discrimination power of measures for nonlinearity in a time series. Phys Rev E Stat Nonlin Soft Matter Phys *55*, 5443-5447.

Shannon, C.E., and Weaver, W. (1949). A Mathematical Theory of Communication. Bell System Technical Journal *27*, 379-423.

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., Evans, A., Rapoport, J., and Giedd, J. (2006). Intellectual ability and cortical development in children and adolescents. Nature *440*, 676-679.

Sheline, Y.I., Barch, D.M., Price, J.L., Rundle, M.M., Vaishnavi, S.N., Snyder, A.Z., Mintun, M.A., Wang, S., Coalson, R.S., and Raichle, M.E. (2009). The default mode network and self-referential processes in depression. Proc Natl Acad Sci U S A *106*, 1942-1947.

Singh, I., and Rose, N. (2009). Biomarkers in psychiatry. Nature *460*, 202-207.

Sörnmo, L., and Laguna, P. (2005). Bioelectrical signal processing in cardiac and neurological applications (Amsterdam ; Boston, Elsevier Academic Press).

Stam, C.J. (2005). Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. Clin Neurophysiol *116*, 2266-2301.

Stauffer, D. (2006). Biology, sociology, geology by computational physicists, 1st edn (Amsterdam ; Boston, Elsevier).

Supekar, K., Musen, M., and Menon, V. (2009). Development of large-scale functional brain networks in children. PLoS Biol *7*, e1000157.

Supekar, K., Uddin, L.Q., Prater, K., Amin, H., Greicius, M.D., and Menon, V. Development of functional and structural connectivity within the default mode network in young children. Neuroimage *52*, 290-301.

Supekar, K., Uddin, L.Q., Prater, K., Amin, H., Greicius, M.D., and Menon, V. (2010). Development of functional and structural connectivity within the default mode network in young children. Neuroimage *52*, 290-301.

Takahashi, T., Cho, R.Y., Mizuno, T., Kikuchi, M., Murata, T., Takahashi, K., and Wada, Y. (2010). Antipsychotics reverse abnormal EEG complexity in drug-naive schizophrenia: a multiscale entropy analysis. Neuroimage *51*, 173-182.

Uddin, L.Q., Kelly, A.M., Biswal, B.B., Margulies, D.S., Shehzad, Z., Shaw, D., Ghaffari, M., Rotrosen, J., Adler, L.A., Castellanos, F.X., *et al.* (2008). Network homogeneity reveals decreased integrity of default-mode network in ADHD. J Neurosci Methods *169*, 249-254.

Uddin, L.Q., Supekar, K., and Menon, V. (2010). Typical and atypical development of functional human brain networks: insights from resting-state FMRI. Front Syst Neurosci *4*, 21.

Uhlhaas, P.J., Haenschel, C., Nikolic, D., and Singer, W. (2008). The role of oscillations and synchrony in cortical networks and their putative relevance for the pathophysiology of schizophrenia. Schizophr Bull *34*, 927-943.

Van Rie, A., Dow, A., Mupuala, A., and Stewart, P. (2009). Neurodevelopmental trajectory of HIV-infected children accessing care in Kinshasa, Democratic Republic of Congo. J Acquir Immune Defic Syndr *52*, 636-642.

Varela, F., Lachaux, J.P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. Nat Rev Neurosci *2*, 229-239.

Wackermann, J., Lehmann, D., Dvorak, I., and Michel, C.M. (1993). Global dimensional complexity of multi-channel EEG indicates change of human brain functional state after a single dose of a nootropic drug. Electroencephalogr Clin Neurophysiol *86*, 193-198.

Ward, L.M. (2003). Synchronous neural oscillations and cognitive processes. Trends Cogn Sci *7*, 553-559.

Whittington, M.A. (2008). Can brain rhythms inform on underlying pathology in schizophrenia? Biol Psychiatry *63*, 728-729.

Xie, H.-B., He, W.-X., and Liu, H. (2008). Measuring time series regularity using nonlinear similarity-based sample entropy. Physics Letters A *372*, 7140-7146.

Zavaglia, M., Astolfi, L., Babiloni, F., and Ursino, M. (2008). The effect of connectivity on EEG rhythms, power spectral density and coherence among coupled neural populations: analysis with a neural mass model. IEEE Trans Biomed Eng *55*, 69-77.

Zbilut, J.P., Thomasson, N., and Webber, C.L. (2002). Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals. Med Eng Phys *24*, 53-60.

Zhang, D., Ding, H., Liu, Y., Zhou, C., and Ye, D. (2009). Neurodevelopment in newborns: a sample entropy analysis of electroencephalogram. Physiol Meas *30*, 491-504.

22

Life Support System Virtual Simulators for Mars-500 Ground-Based Experiment

Eduard Kurmazenko[1], Nikolay Khabarovskiy[1], Guzel Kamaletdinova[1],
Evgeniy Demin[2] and Boris Morukov[2]
[1]*Joint-Stock Co. 'NIIchimmash',*
[2]*SSC RF-IMBP RAS,*
Russia

1. Introduction

A Mars manned mission is practically impossible without resolving some problems on the ground with test subjects involved related to crew life-support psychological stability, fitness to work during a long-duration, self-sustained space mission. One of the problems to be resolved in spaceflight is the crew's health and fitness work. These factors are possible to investigate in a ground experiment to make more effective preparation for interplanetary missions including a Mars mission. Another problem lies in failure-free functioning of on-board systems and first of all the Integrated Regenerative Life-Support System (IRLSS) in the Mars-500 project. The crew plays a key role in maintaining system operability and reliability the entire mission.

In order to make long-duration, self-sustained interplanetary missions a reality it is necessary to provide crew support and its activities under conditions essentially different from those of earth orbital flights.

Specifics of interplanetary flights include:

- long duration (over 960 day) missions with the crew being in confined space that demands:
    - expansion of functions of IRLSS functions related to crew personal;
    - prompt parameters of the crew's environment under spaceflight conditions;
    - the necessity of carrying out the crew's medical control and strain relief on-board the spacecraft.
- Self-sustained manned flight is characterized by:
    - lack of renewal expendables units;
    - the systems incorporated in the IRLSS architecture shell ensure trouble-free performance over the entire flight with minimal spare parts and expendables required;
    - the necessity of decision-making by crew as identification and localization of possible off-normal situations related to the IRLSS due to limited intervention on the part of the ground mission control center to control the crew's actions.

Investigations into prolonged influence of the conditions of self-sustained interplanetary flight on crewmembers' intellectual faculties in operation of the IRLSS are of prime importance to make a mission a success and cannot be predicted on the basis of theoretical and experimental data, ground experiments conducted shall be aimed first of all at evaluating the operator's effective activities and his psychological condition (attention, vigilance, perception, memory, thinking training) in combination with physiological parameters of the central and vegetative nervous systems.

Actual participation of test subjects in serving operational standard system at first phases of ground simulation of spaceflight to Mars is extremely complicated and economically unprofitable. A more rational approach is the application of standard system virtual simulators interacting with simulation models for both environment and crew as a load component and integrated in a single Hardware/Software Complex for Serving Operational Systems by crew (HSCSOS) intended for system functioning in normal, off-normal and emergency situations resulted in failure of some systems and deviation of environment controllable parameters from specified values. Those situations may include human factor (crew members' decline in fitness to work, activities, etc.).

The purpose of this charter is analysis of all possible approaches to development of similar complexes based on simulation taking into account a long-duration stay of Man in confined space. The analysis results may be used in development of similar hardware/software complexes to analyze complicated human-machine interaction and specialist training for various-purpose Man-Made Ecosystems (MMES).

2. The simulating object: Its engineering architecture and properties

The crews, environment, IRLSS placed in Pressurized Manned Modules (PMM) define a MMES of Interplanetary Spacecraft intended for crewmembers' support and activity as well as other biological object under conditions essentially different from those on Earth (Kurmazenko E.A. at al., 2009).

The distinguished features of the MMES in comparison with natural ecosystem include:

- the necessity of creating system architecture based on processes with intensity significantly exceeding the intensity of natural transformation processes;
- the system architecture should incorporate finite number of engineering devices and units with built-in or embedded man-made technologies;
- the capability of stable functioning is governed by the value of a substance reserve fund of the substances disposed in the PMM confined volume.

The IRLSS as an abiotic component of the MMES intended for long-duration manned missions may be defined as a sophisticated engineering system with devices interacting in time and space to provide crew support based on metabolic product recovery, a minimum of spare parts and expendables to create the conditions which ensure that the crew will be provided with physical and mental stability to a specified degree of reliability (Pravetskiy V.N. et al., 1981).

In generalized state space the MMES as a complicated integral system may be given by

$$\mathrm{C} \equiv \mathrm{C}\left(\left|C\right|, \bar{C}, \underline{C}\right). \tag{1}$$

The initial system decomposition allows two interrelated and interacted components to be signed out the IRLSS and external environment:

$$\mathrm{S} \equiv \mathrm{S}\left(\left|S\right|, \bar{S}\right); \tag{2}$$

$$\mathrm{E} \equiv \mathrm{E}\left(|E|, \bar{E}, \underline{C}\right), \tag{3}$$

$$\mathrm{C} = \mathrm{S} \cap \mathrm{E}; \tag{4}$$

$$\bar{C} \equiv \bar{S}\left(|S|, |E|, \bar{E}\right) \cap \bar{E}\left(|E|, \bar{E}, |S|\right), \tag{5}$$

in accordance with the purpose of the **C** system over each time of functioning.

In the expressions (1)÷(5) the following notation is adopted: **C, S, E**=the state functions of the MMES, IRLSS and an environment systems respectively; $|C|,|S|,|E|$ =the technological structure its systems; $\bar{C},\bar{S},\bar{E}$ =the system functioning regularities; $\underline{C}$ =the purpose of the MMES.

An IRLSS is selected as a baseline for HSCSOS development consists of (Figure 1):

- a System for Water Recovery from Humidity Condensate (WRS-AC) based on sorption/catalytic process to remove of organic and inorganic contaminants from condensate;
- a System for Water Reclamation from Urine (WRS-U) based on the low-tem-premature distillation and sorption/catalytic process for the removal of organic and inorganic contaminants from urine condensate;
- a System for Oxygen Generation (OGS) by water electrolysis from electrolyte solution (the OGS-1 is a virtual simulator and the OGS-2 is an Electrical Trainer (ET) of standard system enabling a number of manual operations);
- a Oxygen Solid Fuel Generator (OSFG) based on sodium perchlorates;
- a Bottle Filling System (BFS) for extra-vehicular activity based on post-purification and compression of oxygen produced by OGS;
- a Trace Contaminant Control System (TCCS) based on sorption/catalytic process for the removal of trace contaminants from the pressurized manned module cabin atmosphere;
- a Carbon Dioxide Removal System (CDRS) based on regenerable absorbent vacuum desorption;
- a Carbon Dioxide Concentration System (CDCS) based on regenerable absorbent vapor desorption;
- a Carbon Dioxide Reduction System (CRS) based on carbon dioxide conversion to methane by hydrogen;
- an Atmosphere Leakage Make-up System (ALMS) by means of nitrogen supply with supply correction based on the oxygen partial pressure value;
- Atmosphere Leakage Tracing System (ALTS).

The individual systems as part of the IRLSS interact via interfaces between the systems and environment. The analysis results of the operational interfaces shows the HSCSOS architecture shall incorporated the environment components such as the atmosphere of the PMM and the crew taking into account specifics of its functioning in interplanetary flight conditions as well as on-board Power Supply System (PSS) and a convective/radiation Thermal Control System (TCS).

The operator's skills which crewmembers shall posses in the critical phases of long self-sustained mission owing to localization of the off-nominal situations (ONS) related the IRLSS functioning imply a high degree of sophistication and may be disturbed under influence of many kinds of stresses. Therefore, when off-normal situation localization simulation is carried out it is necessary to approach it as closely as possible.

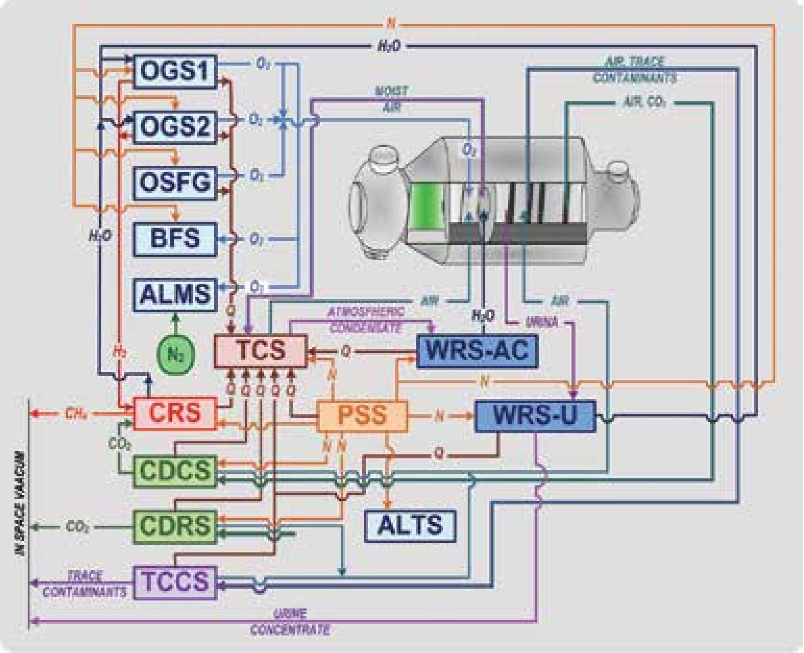

N
H₂O
OGS1
OGS2
OSFG
BFS
ALMS
O₂
O₂
O₂
O₂
O₂
MOIST AIR
AIR TRACE CONTAMINANTS
AIR, CO₂
O₂
H₂O
H₂
N₂
AIR
ATMOSPHERIC CONDENSATE
H₂O
URINA
AIR
Q
TCS
Q
WRS-AC
Q Q Q Q Q
N
N
CH₄
CRS
PSS
N
WRS-U
CO₂
N N
Q
CDCS
CDRS
ALTS
CO₂
IN SPACE VAACUM
TRACE CONTAMINANTS
TCCS
URINE CONCENTRATE

Fig. 1. Integrated Regenerative Life-Support System baseline selected for HSCSOS.

3. Approach to the HSCOCS software development

Approach to the HSCSOS software development is based on the application of the simulation modeling for analysis for the IRLSS system performance in the normal, off-normal and emergency situations, the monitoring of the environmental controllable parameter values and efficient assessment of the crew's actions when off-normal situations are being localized. The simulation model is a logic-mathematical description of the object which may be applied to staging computational experiments in designing, analyzing and assessing object functioning.

Simulation modeling as a specific type of modeling is applied when:

- an experiment upon an actual object is difficult or impossible;
- an analytical model cannot be built (there is cause/effect relationships, how-linearity, stochastic variables);
- time dependent system mode of operation.

The analysis results of the IRLSS performance show that the HSCSOS software architecture shall include the following interacting units (Figure 2):

- IRLSS individual system functioning Virtual Simulators (VS);
- a simulation model of the PMM atmosphere integrating the VSs as a whole and providing the monitoring of crew's environment controllable parameters;
- a crew simulation model as component of environment governing the loads on individual systems;
- procedures for generating probable off-normal situation (ONS) in operation of the IRLSS system, ONS identification and preparation of guidelines and rules for crew in localization of given ONS;

- procedure for crewmember's action efficiency assessment in localization of ONS;
- a specially-created database of monitoring and measurement of crewmember's physiological parameters.

In complex development such as the HSCSOS it is most important for formation of the closed formalized description to select the approaches for generation a closed formalized description of the above specific components of the complex architecture in order to obtain required and sufficient information on system performance analysis in normal and off-normal situations.

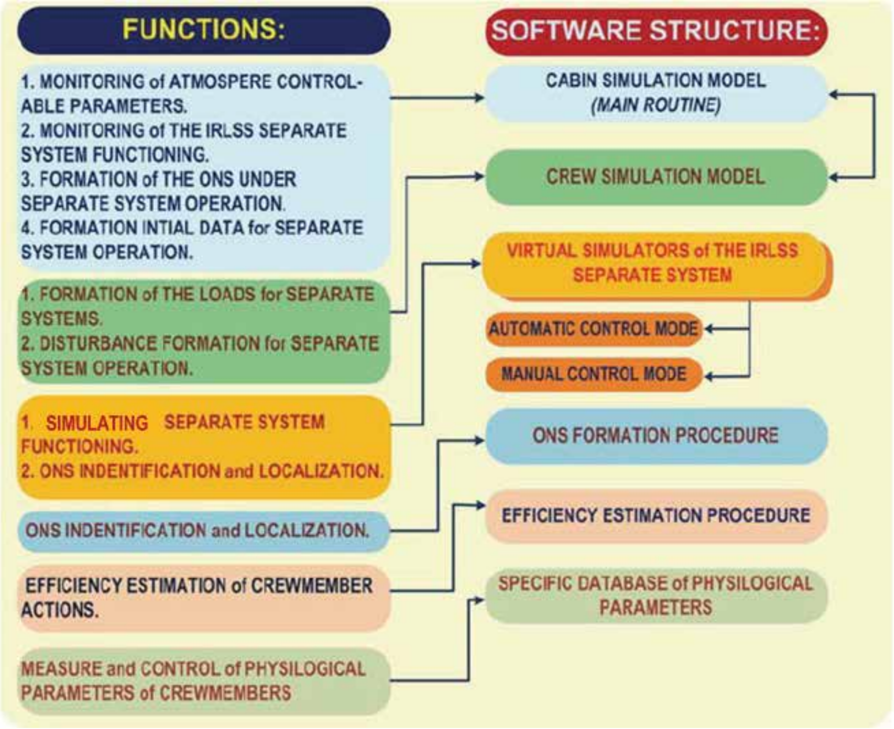

FUNCTIONS:
1. MONITORING of ATMOSPERE CONTROL-ABLE PARAMETERS.
2. MONITORING of THE IRLSS SEPARATE SYSTEM FUNCTIONING.
3. FORMATION of THE ONS UNDER SEPARATE SYSTEM OPERATION.
4. FORMATION INTIAL DATA for SEPARATE SYSTEM OPERATION.
1. FORMATION of THE LOADS for SEPARATE SYSTEMS.
2. DISTURBANCE FORMATION for SEPARATE SYSTEM OPERATION.
1. SIMULATING SEPARATE SYSTEM FUNCTIONING.
2. ONS INDENTIFICATION and LOCALIZATION.
ONS INDENTIFICATION and LOCALIZATION.
EFFICIENCY ESTIMATION of CREWMEMBER ACTIONS.
MEASURE and CONTROL of PHYSILOGICAL PARAMETERS of CREWMEMBERS
SOFTWARE STRUCTURE:
CABIN SIMULATION MODEL (MAIN ROUTINE)
CREW SIMULATION MODEL
VIRTUAL SIMULATORS of THE IRLSS SEPARATE SYSTEM
AUTOMATIC CONTROL MODE
MANUAL CONTROL MODE
ONS FORMATION PROCEDURE
EFFICIENCY ESTIMATION PROCEDURE
SPECIFIC DATABASE of PHYSILOGICAL PARAMETERS

Fig. 2. The architecture and basic functions of the HSCSOS software some components.

3.1 An approach applied to generating formalized descriptions of virtual simulators

Formalized descriptions of individual system virtual simulators based on the detailed level as the aggregate are applied (Kurmazenko E.A. at al., 1997; Kurmazenko E.A. at al., 2008).

In this case an individual system is presented as an aggregate which implies generation of a closed mathematical description including (Figure 3 a):

- the set {X}=alphabet of state in parameters;
- the set {Y}=alphabet of state output parameters;
- the set {Z}=alphabet of state inner parameters;
- the set {U}=alphabet of state controlling parameters;
- the set {W}=alphabet of outer and inner perturbation actions.

When applying the simulation models the alphabet of in parameters, out parameters and controlling parameters shall be correspond to controllable parameters of the system being simulated. The alphabet of perturbation actions is governed by time-varying controllable parameter values of the environment with a specific system in operation and crew present in the PMM atmosphere. The inner perturbation actions are mainly governed by controllable parameter values of the system being simulated. In order to formulate the alphabet of inner states an approach based on a functional description of this description is applied.

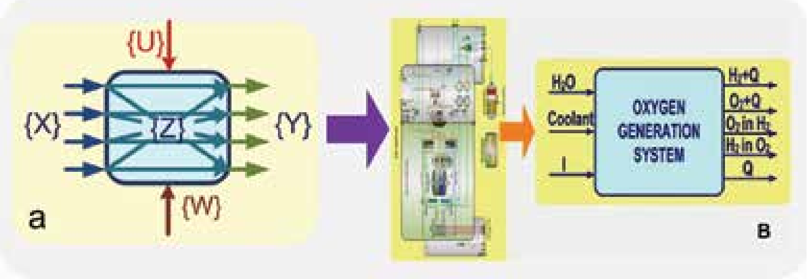


Fig. 3. The system presented as the aggregate: a = design aggregate schematic; b = the OGS presentation as aggregate.

As an example generation of a formalized description of the inner state alphabet for the OGS VS is presented on Figure 3 b. In order to formulate the alphabet of inner states the following basic assumptions are made:

- A formalized description shall take into consideration only values of controllable parameters of inflow and outflow which govern the regulatory and behavior of system functioning;
- thermal/physical properties of electrolyte–produced gas mixtures and coolant over the temperature range investigated are assumed constant;
- owing to a short time of transient process for a current the mass flows of oxygen, hydrogen in oxygen, hydrogen, oxygen in hydrogen and water vapor are described by algebraic equations;
- the main sources of oxygen in hydrogen and hydrogen in oxygen are electrolyzer water supply headers in which an uncontrollable electrolysis process takes place.

With consideration for the given assumptions the formalized description of the inner state alphabet may be presented as:

$$I_{el}^{\tau} = I \frac{R_{col}^{\tau}}{R_{col}^{\tau} + R_{el}^{\tau}}\ ; \tag{6}$$

$$U^{\tau} = e_a^0 + e_c^0 + a_a + a_c + 0{,}001\left(b_a + b_c\right)\frac{I_{el}^{\tau}}{S_{ed}} + I_{el}^{\tau} R_{el}^{\tau}\ ; \tag{7}$$

$$I_{col} = I - I_{el}\ ; \tag{8}$$

$$G_{O_2(H_2)}^{\tau} = \frac{n}{2F}\left[2A_{O_2(H_2)} I_{el}^{\tau} + \left(A_{O_2} + A_{H_2}\right) I_{col}^{\tau}\right]; \tag{9}$$

$$G_{H_2 \to O_2(O_2 \to H_2)}^{\tau} = \frac{n A_{O_2}\left(A_{H_2}\right)}{2F}, \tag{10}$$

where: A_{O_2}, A_{H_2} =chemical equivalents of oxygen and hydrogen, kg/mol, respectively; a_a, a_c, b_a, b_c =Tafel's constants for the anode and cathode; e_a^0, e_c^0 = theoretical potentials of the anode and cathode, V for a_a and a_c and Vm²/A, for b_a and b_c, respectively; F=Faraday constant, Kl/mol; $G^{\tau}_{O_2(H_2)}$ =the mass flow-rate of gas being produced in the electrolyzer oxygen (hydrogen) compartment, kg/s; $G^{\tau}_{H_2 \to O_2(O_2 \to H_2)}$ = the mass flow-rate of oxygen (hydrogen) produced in the hydrogen (oxygen) compartment; I, I_{el}, I_{col} = the total electrolyzer current, current through the electrolytic cell and current in the header, A; n= the quantity of an electrolytic cells; $R^{\tau}_{col}, R^{\tau}_{el}$ = the electric resistance of the header and electrolyzer, Ohm; S_{ed} =the electrolytic cell surface area, m²; U^{τ} = the electrolyzer voltage, V.

The average electrolyzer temperature T_{el}, K, as a function of the supply current I_{el} is determined from regressive dependence

$$T_{el} = f_1(T_{cool})\left(-0.3305 + 0.0034 I_{el}^2 + 2.085 I_{el}^{0.5} + \frac{21.29}{I_{el}^{0.5}}\right), \tag{11}$$

obtained as a result of processing of the data of a computer experiments conducted by using the OGS detailed simulation model.

The hydrogen (oxygen) moisture content downstream the separator, kg/kg

$$d^{\tau}_{H_2(O_2)} = \frac{\left[27.6 + 0.23\left(T^{\tau}_{sep} - 273\right)^{1.5}\right]}{m_{sep}} \cdot \frac{\mu_{H_2O}}{\mu_{H_2(O_2)}}, \tag{12}$$

where $\mu_{H_2(O_2)}, \mu_{H_2O}$ =the molar masses of hydrogen (oxygen) and water, kg/mol.

The temperature of the mass flows of hydrogen and oxygen downstream from the separator T_{sep}, K, is determined by regressive dependence as

$$T_{sep} = f_2(T_{cool})\left(-15.387 + 0.249 I_{el} + \frac{173.184}{I_{el}^{0.5}} + \frac{319.898}{I_{el}}\right). \tag{13}$$

The temperature functions $f_1\ (T_{cool})$ in the relationship (11) and $f_2\ (T_{cool})$ in the relationship (13) are determined as

$$f_1(T_{cool}) = \left[1 + 0.82(T_{el} - 273) + 0.03(T_{el} - 273)^2\right] T_{el}^{-1}; \tag{14}$$

$$f_2(T_{cool}) = \left[1 + 0.9(T_{el} - 273)\right] T_{sep}^{-1}, \tag{15}$$

where T_{cool}=the coolant temperature, K.

The similar approach to generation of formalized descriptions of system functioning is adopted for other virtual simulators.

3.2 Approach used for formation of the PMM atmosphere formalized description

When generation the formalized PMM atmosphere description the following basic assumptions are made (Kurmazenko E.A. at al., 1998):

- the PMM atmosphere is considered as an open thermodynamic system;
- man-made atmosphere is considered as a mixture of ideal gases the heat capacity of which is governed by its chemical composition and temperature–independent;
- trace contaminants due to their low content do not affect the generation of total pressure in the PMM and thermal/physical properties of man-made gaseous atmosphere.

Considering the assumptions made the nonlinear equations of mass balances for the basic components (oxygen, carbon dioxide, nitrogen and water vapor) and trace contaminants as well as the non-linear equation of internal energy balance for the PMM atmosphere reference volume may be written as the equations in deviations:

$$M_{PMMa}(\tau) = \sum_{i=1}^{i=4} M_{PMMi}(\tau); \tag{16}$$

$$M_{PMMi}(\tau) = M_{PMMi}(\tau - \Delta\tau) + \sum_{j=1}^{j=n}\left(\pm G_{ij}\Delta\tau\right); \tag{17}$$

$$M_{TC}(\tau) = M_{TC}(\tau - \Delta\tau) + \sum_{k=1}^{k=s}\left(\pm G_{TCk}\Delta\tau\right) \tag{18}$$

$$U(\tau) = U(\tau - \Delta\tau) \pm \sum_{l=1}^{l=p} c_{p_i} G_{il}\Delta\tau \pm \sum_{m=1}^{m=t} q_m \Delta\tau \tag{19}$$

In the equations (16)÷(19) the following notation is adopted: $M_{PMMa}, M_{PMMi}, M_{TC}, U$=the value of atmosphere total mass, i-basic component mass, k-trace contaminant mass, kg, and atmosphere internal energy, J, respectively; G_{ij}, G_{TCk}=i- basic component mass flow-rate and k=trace contaminant mass flow-rate entering and leaving the reference volume, kg/s; q_m=heat flows due to heat conduction entering and leaving the volume under consideration, W; c_{p_i} = the i- basic component specific heat capacity, J/kg °C; $\tau - \Delta\tau, \tau, \Delta\tau$ = previous time, current time and integration step in time, respectively, s.

The current values of mass flow-rates of the atmosphere basic components and trace contaminants upstream and downstream the reference volume as well as heat flows entering and leaving together with mass flows of atmosphere components and heat conduction are determined at each integration step by the current values of the ingoing and outgoing flows with the system performance virtual simulator values.

3.3 Approach used for formation of the 'crew' unit formalized description

When generating a formalized description of the 'crew' unit the following assumptions are made:

- a single crewmember is considered as the structure of interrelated functioning systems in which incoming mass and energy flows are converted into outgoing mass and heat flows, and activity;

- the cosmonaut's energy expenditure when doing various kinds of activity is balanced by caloric value of food ration and total value of energy expenditure;
- potable water is consumed with food;
- the main factor that governs the basic point is the crew's activity defined by the spaceflight program;
- in order to describe mass and heat flows in the 'crew' unit an international model of a conventional human where the mass and heat flows are proportional to energy expenditures (Adamovich BA., and Gorshenin V.A., 1997). In doing so, the coefficients of this model are corrected based on the results of the computational experiments used the detailed simulation model on the basis which the human organism main functional systems (Figure 4), governing the mass/exchange with the environment are simulated (Kurmazenko E.A. at al., 2000).

The initial data used for generating a formalized description of the 'crew' unit also include an activity/rest cycloramas for every crewmember.

Oxygen consumption G_{O_2}, g/h, is a function of energy expenditure N

$$G_{O_2} = a_0 N \,, \tag{20}$$

in which coefficient a_0 varies in the range from 0.28 to 0.31, g/kcal.

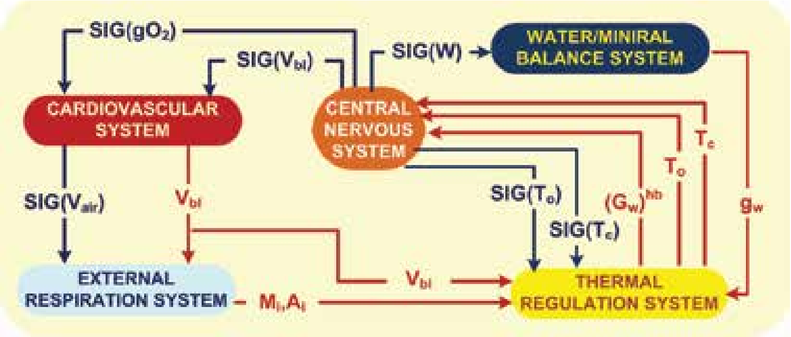


Fig. 4. Enlarged flowchart of human body main functional systems, where: $\mathrm{SIG}(gO_2)$, $\mathrm{SIG}(V_{bl})$, $\mathrm{SIG}(V_{air})$, $\mathrm{SIG}(W)$, $\mathrm{SIG}(T_0)$, $\mathrm{SIG}(T_c)$ =signal functions of the Central Nervous System controlled by the oxygen consumption, blood flow, breath volume, water consumption and Thermal Regulation System in dependence from values of ambient temperature and human's body core temperature, respectively; $V_{bl}, M_i, A_i, (G_w)^{hb}, g_w, T_0, T_c$ =the controllable parameters for signal functions.

Carbon dioxide released G_{CO_2} is a function of energy expenditure N is determined as

$$G_{CO_2} = a_0 K_r N \,, \tag{21}$$

with the value of the respiratory coefficient K_r is determined from empirical dependence

$$K_r = -0.801 + 0.142 \exp\left(\frac{-96.432}{A}\right). \tag{22}$$

The moisture losses as a result of perspiration and respiration *J*, g/h, are also of energy expenditure *N* is determined as

$$J = 91.7 - 0.19N\ . \tag{23}$$

The quantity of urine donated *U*, g/h, is a function of energy expenditure *N*

$$U = 0.19N\ . \tag{24}$$

Trace contaminant realized TC_i to be considered as a first order approximation are proportional energy expenditure *N*

$$TC_i = A_i N\ . \tag{25}$$

In calculation of trace contaminants realized the data presented in work (Savina, V.P., & Kuznetsova, T.I., 1980) are used. The data processed for the mixed ration (50 %-natural food products and 50 % sublimated food products) as a function of the ambient temperature are given in Table 1.

Trace contaminant	**Temperature-dependency specific secretion intensity A_i, mg kcal /h for:**	
	t=20	**20< t ≤40**
Ammonia and its compounds	0.0144	0.0144+0.0031t
Ketones	0.0577	0.0577+0.0003t
Carbon monoxide	0.195	0.195+0.0186t
Aldehydes	0.005	0.005+0.0005t
Inorganic acids	0.018	0.018+0.0017t
Total Alcohols	0.012	0.012+0.0005t
Hydrocarbons	0.008	0.008+0.0008t
Methane	0.033	0.033+0.0003t
Acetaldehyde	0.003	0.003+0.0001t
Methanol	0.008	0.008+0.0001t
Ethanol	0.011	0.011+0.001t
Dimethylamine	0.007	0.007+0.0006t
Acetone	0.046	0.046+0.0043t

Table 1. Initial data for simulation of trace contaminant secretion.

4. HSCSOS software implementation

The HSCSOS architecture includes the hardware and software components integrated in a single hardware/software complex.

A general view of the Mars-500 project Ground Experimental Stand (GES) and HSCSOS hardware arrangement are given on Figure 5 and Figure 6, respectively.

The HSCSOS hardware features two work places located in the PMM: Operator's Terminal in PMM-150 (OT1) and Operator's Terminal in PMM-50 (OT2), and the Instructor's Work Place (IWP) located in the Control Experimental Center (CEC).

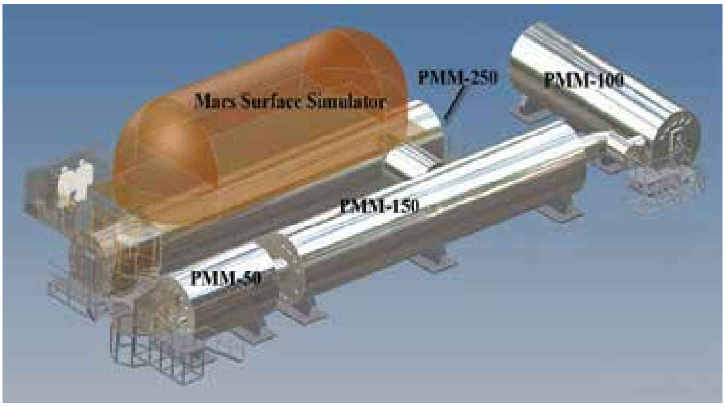

Mars Surface Simulator
PMM-250
PMM-100
PMM-150
PMM-50

Fig. 5. The general view of the Mars-500 project Ground Experimental Stand.

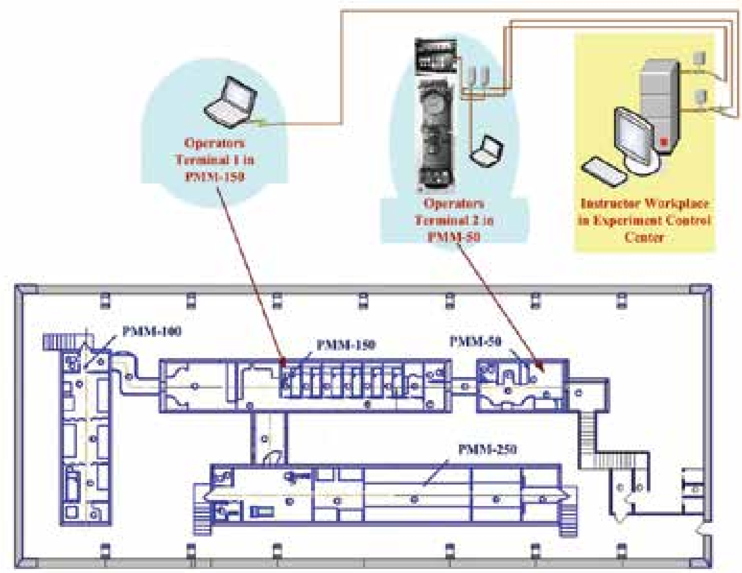

Operators Terminal 1 in PMM-150
Operators Terminal 2 in PMM-50
Instructor Workplace in Experiment Control Center
PMM-100
PMM-150
PMM-50
PMM-250

Fig. 6. HSCSOS hardware arrangement at the Mars-500 project Ground Experimental Stand.

According to the HSCSOS architecture the complex software based on 'client –server' technology of data processing and transfer has been developed. The software includes a server and a client sections. The server and client sections are interrelated via specialized technology of data transfer DataSocet integrated in the programming environment LabView, v. 8.5.

The server section contains various calculation modules of 'PMM atmosphere', and 'crew' simulation models, system virtual simulators, off-normal situation generation module, a crew audio warning module, a crew action efficiency assessment module and a protocol module.

The server section provides computations and interrelation with the client section for displaying the parameter values in order to classify the environment and individual system performance and translate control command from client section for correction of system production rates or ONS localization.

The client section displays the data obtained from the server section, generates control signals to transfer them to the server section. In addition, the client section allows the operator to identify and trace ONS develops in the HSCSOS and the time spent to conduct some operations in servicing the systems and localizing ONS. It is not necessary to permanently interconnect the server and client sections. The operator's terminal can be connected in case if the performance requires intervention.

4.1 'PMM atmosphere' simulation model software implementation

The 'PMM atmosphere' simulation model routine is a main program. This simulation model is the basis one to ensure integration of system virtual simulators in the HSCSOS architecture and monitoring of the crew environment parameters (Figure 7).

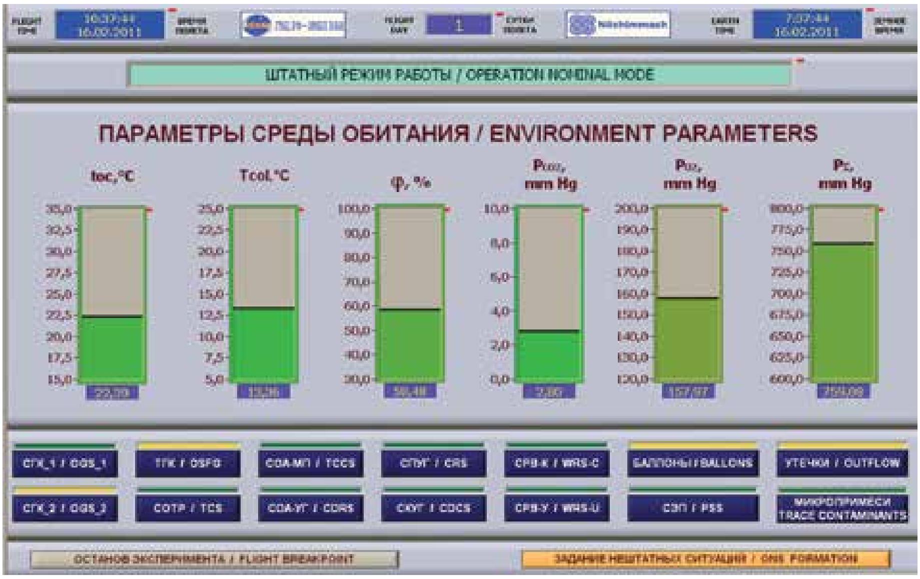

ШТАТНЫЙ РЕЖИМ РАБОТЫ / OPERATION NOMINAL MODE
ПАРАМЕТРЫ СРЕДЫ ОБИТАНИЯ / ENVIRONMENT PARAMETERS
toc,°C
Tcol,°C
φ, %
Pco2, mm Hg
Po2, mm Hg
PΣ, mm Hg
СГК_1 / OGS_1
ТГК / OSFG
СОА-МП / TCCS
СПУГ / CRS
СРВ-К / WRS-C
БАЛЛОНЫ / BALLONS
УТЕЧКИ / OUTFLOW
СГК_2 / OGS_2
СОТР / TCS
СОА-УГ / CORS
СКУГ / CDCS
СРВ-У / WRS-U
СЭП / PSS
МИКРОПРИМЕСИ TRACE CONTAMINANTS
ОСТАНОВ ЭКСПЕРИМЕНТА / FLIGHT BREAKPOINT
ЗАДАНИЕ НЕШТАТНЫХ СИТУАЦИЙ / ONS FORMATION

Fig. 7. Control panel of the 'PMM atmosphere' simulation model, where: t_{oc}, °C, T_{col}, °C, φ, %, p_{CO_2},mm of Hg, p_{O_2},mm of Hg,p_{Σ},mm of Hg = controllable parameter values.

The crew's environment parameter values are displayed on the main routine control panel both as, values and in color when the parameter value is changed. When the '*TRACE CONTAMNANTS*' key is pressed there appears an extra window on the routine control panel to display current trace contaminant concentrations (Figure 8).

In the lower part of the program front panel there are pushbuttons enabling call-in the front panel the corresponding program simulating the functioning of the system when the line appears.

At the bottom of the routine control panel there are call-in keys for control panels of inputs systems simulating system performance. In doing so, color indication '*SYSTEM STATUS CONDITION*' appears. Two keys, located below, fully assessable from the instructor's work place and disconnected at the operator's terminal have inscriptions '*ASSIGNMENT OF OFF-NORMAL SITUATIONS*' and '*EXPERIMENT SHUTDOWN*'.

Experimental verification of the model is carried out on the basis of algorithm assessment for inconsistency, analysis results of computational experiment related to calculation of the balance relationships and ergonomic requirements.

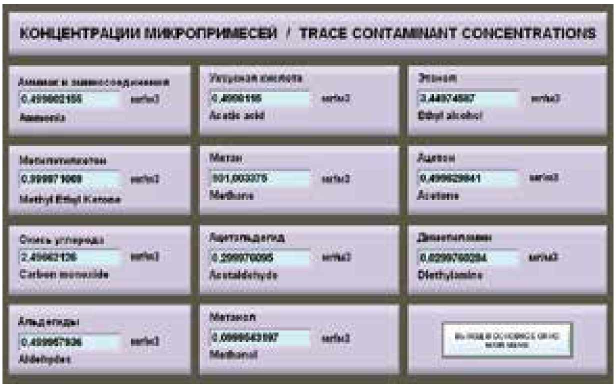

КОНЦЕНТРАЦИИ МИКРОПРИМЕСЕЙ / TRACE CONTAMINANT CONCENTRATIONS
Ammonia
Acetic acid
Ethyl alcohol
Methyl Ethyl Ketone
Methane
Acetone
Carbon monoxide
Acetaldehyde
Diethylamine
Aldehydes
Methanol

Fig. 8. 'Trace Contaminant Concentration' extra windows on control panel 'PMM atmosphere' simulation model.

4.2 'Crew' simulation model software implementation

The 'Crew' simulation model as part of the HSCSOS software simulates mass/energy exchange between the crew and the environment as the results of which the loads required for functioning of virtual simulators are generated. This model is also as a loading component which is the source of disturbance.

The given subroutine operates in the background mode and is not directly displayed when the main 'PMM atmosphere' routine is in operation although being its subroutine (Figure 9).

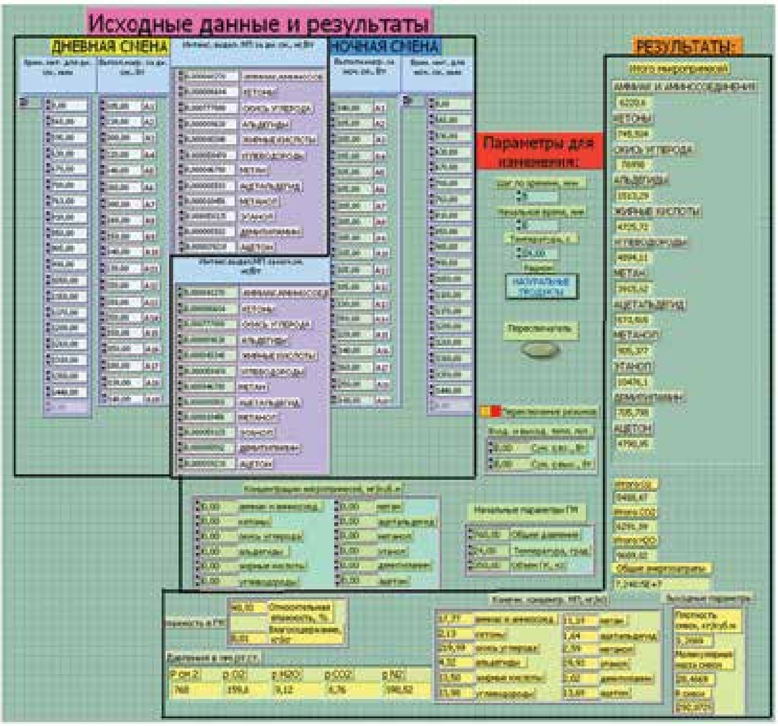

Исходные данные и результаты
ДНЕВНАЯ СМЕНА
НОЧНАЯ СМЕНА
РЕЗУЛЬТАТЫ:
Параметры для изменения:

Fig. 9. Control panel of 'Crew' simulation model.

The main routine is interrelated via 'DAY/NIGHT' subroutine which simulates the crew's energy expenditure in the day and night shifts depending on the activity/rest regimens and according to 'Mass Balance' subroutine operation.

The 'Crew' simulation model is experimentally verified based on a correlation between computational experiment resulted (Figure 10) and published specification.

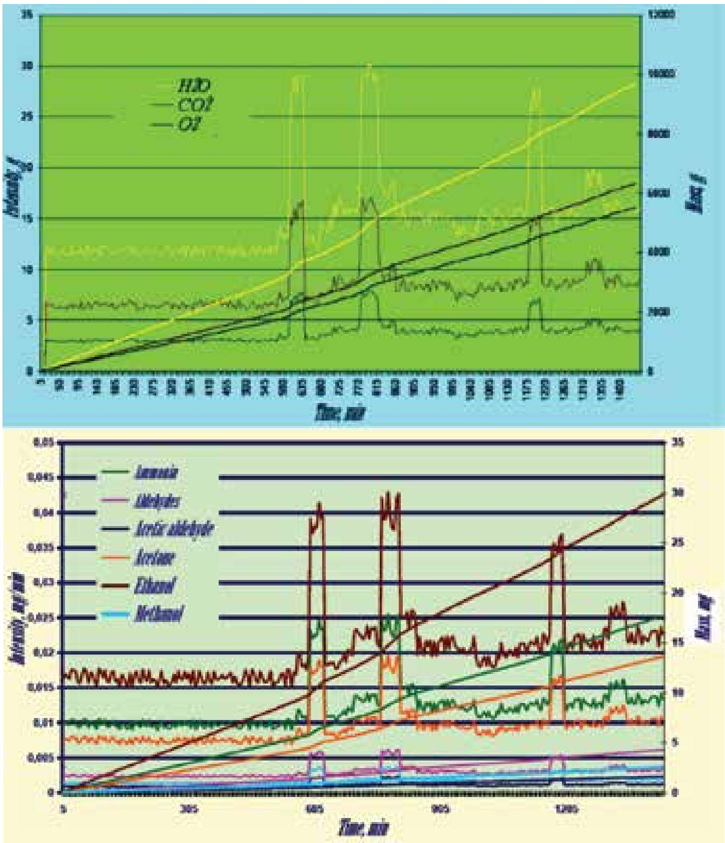


Fig. 10. Typical computational experiment results on 'Crew' simulation model.

4.3 Virtual simulator program implementation

When program implementing virtual simulators of some IRLSS systems its performance data governed by technologies applied they are based on and principles of design execution are adopted as a baseline.

The Air Revitalization and Monitoring Systems (ARMS) are designed to obviate the need to replace units and/or components in prolonged operation. If a unit or a component fails 'cold' or 'hot' redundancy is used to ensure system functioning. Thus, the ONS may be localized without unit replacement.

The systems such as WRS-AC based on sorption/catalytic processes and modular construction require the replacement of some units run out of their lives.

As an example, OGS program implementation as part of the ARMS is considered below (Figure 11).

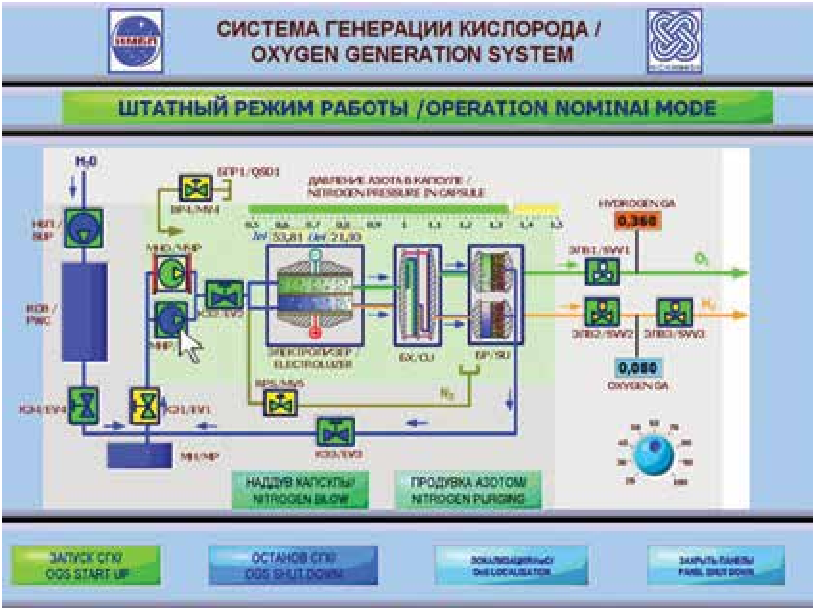

СИСТЕМА ГЕНЕРАЦИИ КИСЛОРОДА /
OXYGEN GENERATION SYSTEM
ШТАТНЫЙ РЕЖИМ РАБОТЫ /OPERATION NOMINAI MODE
H₂0
НБП / SUP
КОВ / PWC
КЭ4/EV4
КЭ1/EV1
МН/MP
КЭ3/EV3
БПР1/QSD1
ВР4/MV4
МНО/MMP
МНР
КЭ2/EV2
ДАВЛЕНИЕ АЗОТА В КАПСУЛЕ /
NITROGEN PRESSURE IN CAPSULE
0,5
0,6
0,7
0,8
0,9
1
1,1
1,2
1,3
1,4
1,5
53,81
21,30
ЭЛЕКТРОЛИЗЕР /
ELECTROLIZER
БХ/CU
БР/SU
ВР5/MV5
N₂
HYDROGEN GA
0,360
ЭЛВ1/SWV1
O₂
ЭЛВ2/SWV2
ЭЛВ3/SWV3
H₂
0,080
OXYGEN GA
НАДДУВ КАПСУЛЫ/
NITROGEN BLOW
ПРОДУВКА АЗОТОМ/
NITROGEN PURGING
ЗАПУСК СГК/
OGS START UP
ОСТАНОВ СГК/
OGS SHUT DOWN
OGS LOCALISATION
PANEL SHUT DOWN

Fig. 11. Front panel of the 'OGS' virtual simulator subroutine in '*OPERATION NOMINAL MODE*' (НБП/SUP=supply unit pump; КОВ/PWC=pre-purification water container; КЭI/EVI=electromagnetic valve; МНО/MMP and МНР/RMP=main micro-pump and reserve micro-pump; БХ/CU=cooling unit; БР/SU=separator unit; GA=gas analyzer).

A possible off-normal situation generated at the IWP is illustrated in Figure 12.

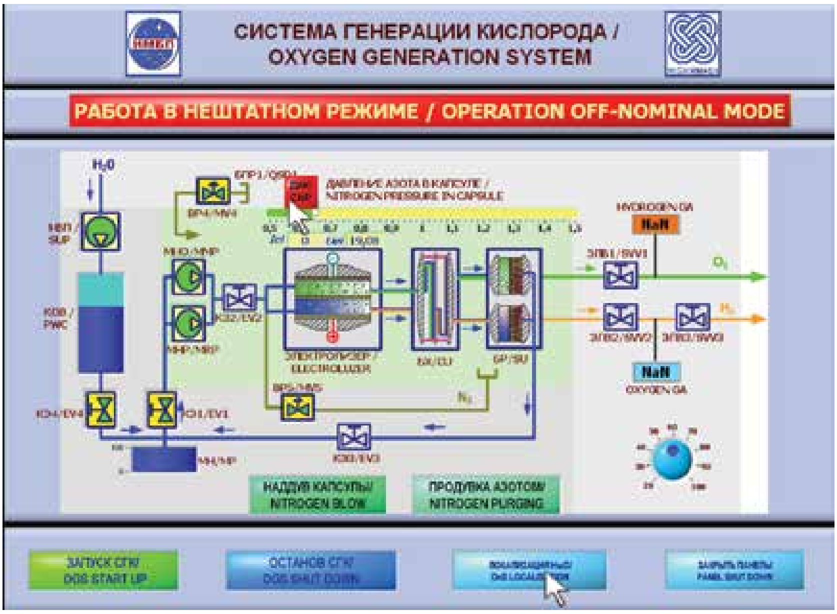

СИСТЕМА ГЕНЕРАЦИИ КИСЛОРОДА /
OXYGEN GENERATION SYSTEM
РАБОТА В НЕШТАТНОМ РЕЖИМЕ / OPERATION OFF-NOMINAL MODE
H₂0
НБП / SUP
КОВ / PWC
КЭ4/EV4
КЭ1/EV1
МН/MP
КЭ3/EV3
БПР1/QSD1
ВР4/MV4
МНО/MMP
МНР/MRP
КЭ2/EV2
ДАВЛЕНИЕ АЗОТА В КАПСУЛЕ /
NITROGEN PRESSURE IN CAPSULE
19,08
ЭЛЕКТРОЛИЗЕР /
ELECTROLIZER
БХ/CU
БР/SU
ВР5/MV5
N₂
HYDROGEN GA
NaN
ЭЛВ1/SWV1
O₂
ЭЛВ2/SWV2
ЭЛВ3/SWV3
H₂
NaN
OXYGEN GA
НАДДУВ КАПСУЛЫ/
NITROGEN BLOW
ПРОДУВКА АЗОТОМ/
NITROGEN PURGING
ЗАПУСК СГК/
OGS START UP
ОСТАНОВ СГК/
OGS SHUT DOWN
PANEL SHUT DOWN

Fig. 12. Front panel of the 'OGS' virtual simulator subroutine in '*OPERATION OFF-NOMINAL MODE*' (Pressure in the canister is below norm.).

The line '*CURRENT STATUS*' displays '*NORMAL OPERATION*' or '*ONS*' inscriptions in the upper part of the control panel. The subroutine generates the following signals:

- current system status;
- a combined signal indicating the necessity of maintenance or ONS localization.

When the '*OFF-NORMAL SITUATION*' signal is displayed on the control panel of the 'PMM atmosphere' main routine the operator shall switch to the OGS subroutine control panel by pressing the key and jump to ONS localization operations.

As an example of OGS shutdown the ONS by the signal '*PRESSURE IN THE CANISTER BELOW NORM* (*CNP*)' is considered. The '*OFF-NORMAL OPERATION*' inscription and '*CNP*' inscription on the OGS control panel light up. To localize the ONS the operator shall press the '*ONS LOCALIZATION*' key on the OGS control panel.

The canister pressurization panel (Figure 13) opens and the operator carries out all the required operations to pressurize the canister.

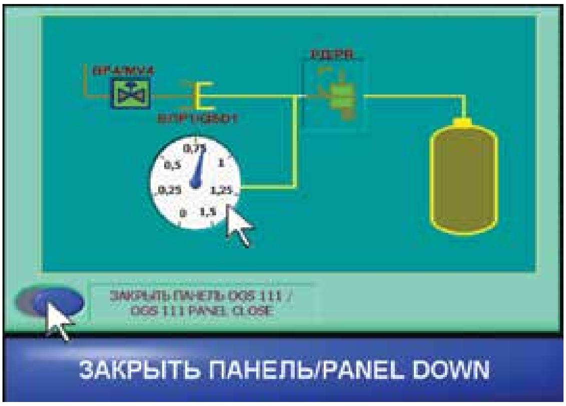

РДPR
0,75
0,5
1
0,25
1,25
0
1,5
ЗАКРЫТЬ ПАНЕЛЬ OGS 111 /
OGS 111 PANEL CLOSE
ЗАКРЫТЬ ПАНЕЛЬ/PANEL DOWN

Fig. 13. Control panel of the subroutine 'Canister pressurization'.

Then the operator shall return to the subroutine by pressing '*PANEL DOWN*' key. After ONS has been localized the operator shall put the system in operation by pressing the '*SYSTEM START-UP'* key and check the startup.

The program implementation of on-board the PMM and TCS syste ms virtual simulators is accomplished similar to that of the IRLSS system virtual simulators (Figure 14 and Figure 15).

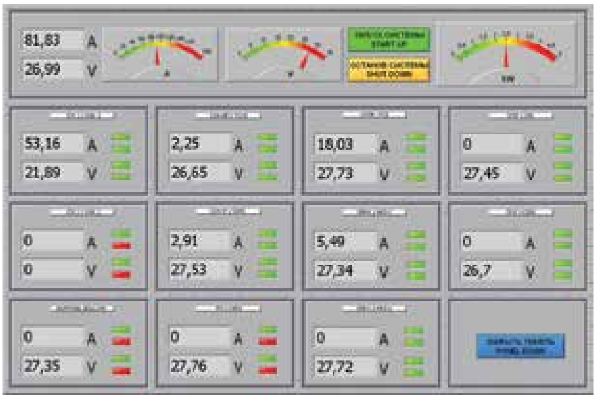

81,83 A
26,99 V
START UP
SHUT DOWN
53,16 A
21,89 V
2,25 A
26,65 V
18,03 A
27,73 V
0 A
27,45 V
0 A
0 V
2,91 A
27,53 V
5,49 A
27,34 V
0 A
26,7 V
0 A
27,35 V
0 A
27,76 V
0 A
27,72 V

Fig. 14. Control panel of the PSS VS.

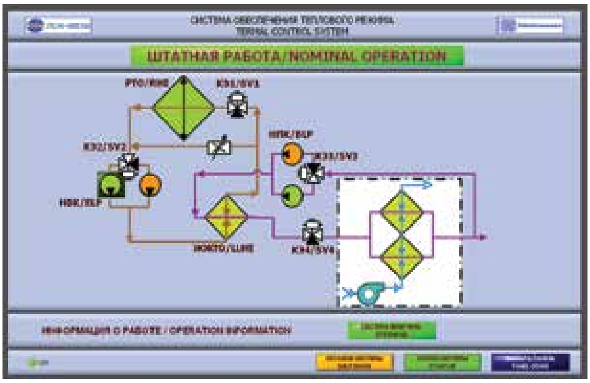

СИСТЕМА ОБЕСПЕЧЕНИЯ ТЕПЛОВОГО РЕЖИМА
TERMAL CONTROL SYSTEM
ШТАТНАЯ РАБОТА/NOMINAL OPERATION
РТО/RHE
К31/SV1
К32/SV2
НПК/BLP
К33/SV3
ЖЖТО/LLHE
К34/SV4
ИНФОРМАЦИЯ О РАБОТЕ / OPERATION INFORMATION

Fig. 15. Control panel of the TCS VS.

OGS-2 electrical trainer (ET) program implementation is for inculcating in crewmembers the practical skills in start-up, normal functioning, and shutdown, and in case of off-normal situations.

The ET (Figure 16) consists: an electrical operational breadboard of the liquid unit (LU); a post-purification unit; a signal and command synchronization unit (SCSU); a commutation unit; an Electron-VM monitoring and control unit implemented by an individual subroutine integrated in the HSCSOS software architecture.

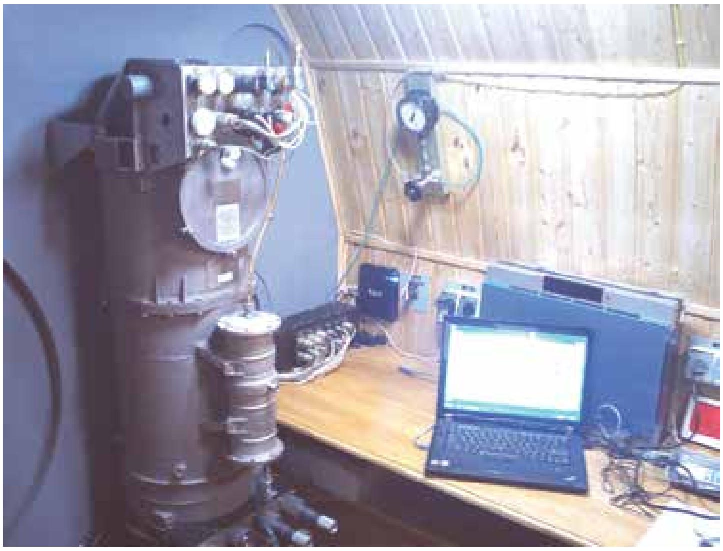


Fig. 16. Electrical trainer of the OGS-2 system based on the Electron-VM standard system.

The ET architecture is based on a combination of the simulation model of functioning realized at the IWP, and standard system hardware. A set of existing units incorporated in the ET architecture is used due to availability, and necessity of carrying out manual operations on the standard system hardware. The signals generated by the sensors of the LU electrical breadboard, as well as the signals and commands are simulated on the IWP computer according to a control algorithm, then converted in the communication unit and enter the SCSU unit to be executed by the LU components.

5. Measurement and monitoring of crewmember's psycho-physiological parameters

A medico-engineering system Biomouse (BMEA) is incorporated in the HSCSOS hardware architecture to perform psycho-physiological tests (Figure 17).

The applied procedure for assessing the functional organism state is based on use of calculation analysis of cardiac rhythm parameters. This procedure enables rapid assessment of the influence of cardiovascular system on the basis of cardiac rhythm parameters. The functionality of cardiovascular system and the excitation degree of the vegetative nervous system are calculated (Baevsky, R.M. at al., 1998).

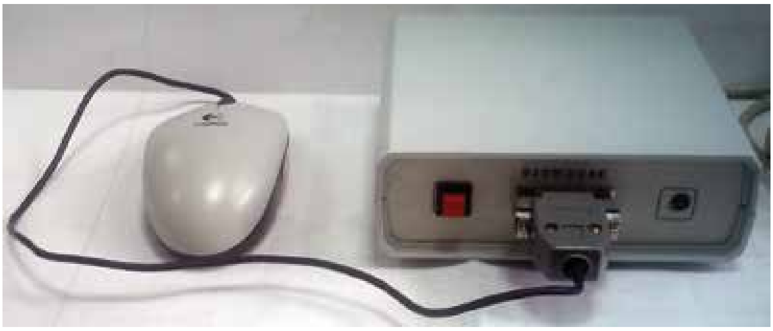


Fig. 17. A general view of the Biomouse medico-engineering system.

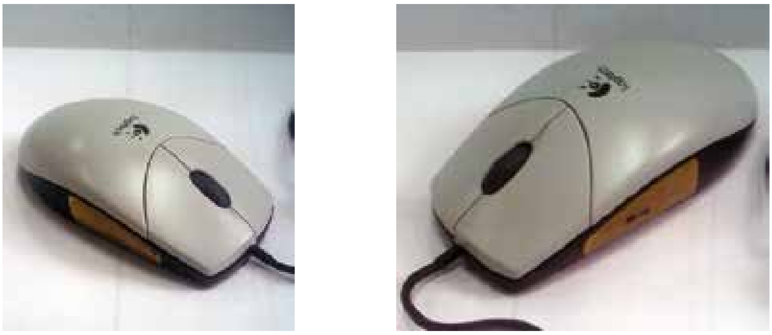


Fig. 18. Manipulator of mouse type with the built-in combined sensors: a view from the PPG sensor is shown in the left and a view from GSR and ECR sensors is shown in the right.

The BMEA system consists of a measurement unit and manipulators of mouse type installed on the OT1 and OT2 with the built-in combined sensors allowing simultaneously and continuously to register three physiological parameters (Figure 18): photo-plethysmogram (PPG), galvanic skin reflex (GSR) and electro-coetaneous resistance (ECR).

The estimation of a crewmember's psycho-physiological condition prior to the beginning of activities and on the termination of activities on localization ONS was conducted in tests mentioned below (ZAO "Neurolab", 2008):

- *Variation hronokardiometriya (WRC)* is method of rapid assessment of the cardio-intervalgrams cardiovascular system regulatory mechanisms. Calculated level of the cardiovascular system functionality and autonomic homeostasis, in addition, recorded: maximum, minimum and average values, as well as fashion, mode amplitude, standard deviation and magnitude of the sequence of cardio. Primary information is photoplethysmogram (a signal from an optical infrared sensor in a digital form). From this signal is allocated an array of cardio intervals, which is subjected to statistical processing.

- *Complex visual-motor reaction (CVMR)* is designed to study the functional state of central nervous system and elements of the operator's attention to human efficiency. The test is based on a study of the statistical characteristics of distribution of the set reaction time. On the screen appear consistently distinguished by the color of light stimuli - circles red or green, the test subject must quickly put out by pressing the right or left mouse button. Recorded response time and response error (omission, premature depression, abnormal response). The following parameters are calculated: the average response time, standard deviation, the number of errors of each type. Based on the statistical parameters of the algorithm on the attached class state of the operator's central nervous system for two-dimensional scale is calculated.
- *Reaction to a moving object (MOR)* is the test to evaluate balance of excitation and inhibition in the nervous system, as well as functional changes under the influence of the load. In this test, the test subject must stop the moving hour hand as close to 12 o'clock by using the Space key.
- *Mirror coordinograph (MC)* is designed to determine the level of stress stability of the test subject. In this test, the operator must use mouse to quickly pass a curved path on the screen without touching its edges. Time and the fact of the contour, the number of touches and time are recorded. These parameters determine the quality of the operator's actions. In addition, before and after the passage of the contour levels are recorded and the mean pacing heart rate. Changes of these parameters are interpreted as the 'value' of the operator. Performance assessment is based on two criteria: 'value' and 'quality'. 'Quality' is composed of indicators such as time of the circuit, the number of touches, while touches.

The BHEA software integrated in the HSCSOS server section software at the IWP is presented by specific database, which executes the following functions:

- storage of a database surveyed crewmembers;
- formation of a set of tests for examination;
- processing results of examination and storage of results in an archival file.

6. HSCSOS operational use in 105-day experiment under Mars-500 project

The main purpose of research is an estimation of efficiency of servicing by the crew of the IRLSS systems. The following problems should be solved for achievement of the given purpose:

- estimation of sufficiency of the controllable parameters list for the analysis of functioning and servicing on the basis of use of the HSCSOS and an electrical trainer of the Electron-VM integrated in complex in conditions of long autonomous mission;
- estimation of efficiency of activity of the operator on localization of the off-nominal situations arising at functioning of systems and/or deviations of the environment controllable parameter values from prescribed values;
- estimation of efficiency of acceptance of independent solutions by crewmembers in the ONS localization;

- estimation of efficiency of ways of display of the information on values of the environment controllable parameters in analysis of functioning and servicing with the use of virtual simulators of systems;
- estimation of the ONS localization influence on the mental and physiological state of crewmembers in conditions of long-term autonomous mission.

6.1 Technique of an experiment

In realization of initial phase, being the final stage of the crew training makes tentative estimation of the crewmember action efficiency in the ONS localization. Formation of the particular situation arising in the IRLSS specific system operation is made on the basis of random sample by the Instructor.

Localization of the arisen ONS is made by each crewmember with the use of the on-board instruction from the operator's workplace disposed in the Main Control Board. In this case both a rigid copy of the on-board instruction, and its electronic version which is available in a format *.PDF on Operator's Terminal 1 and Operator's Terminal 2, can be used without Mission Control Center recommendations.

Generated off-normal situations are characterized by different degrees of complexity in their localization:

- simple (service of a complex, localization of some ONS, not demanding replacements of units, etc.);
- average complexity (the most part of ONS entered demanded the replacement of units);
- complicated (actuate crewmember's activity with the Electron-VM electrical trainer, and also ONS, demanding a long-period operation on elimination or monitoring (imitation of a fire or leakages).

For assessment of the crewmember gained efficiency the following is considered:

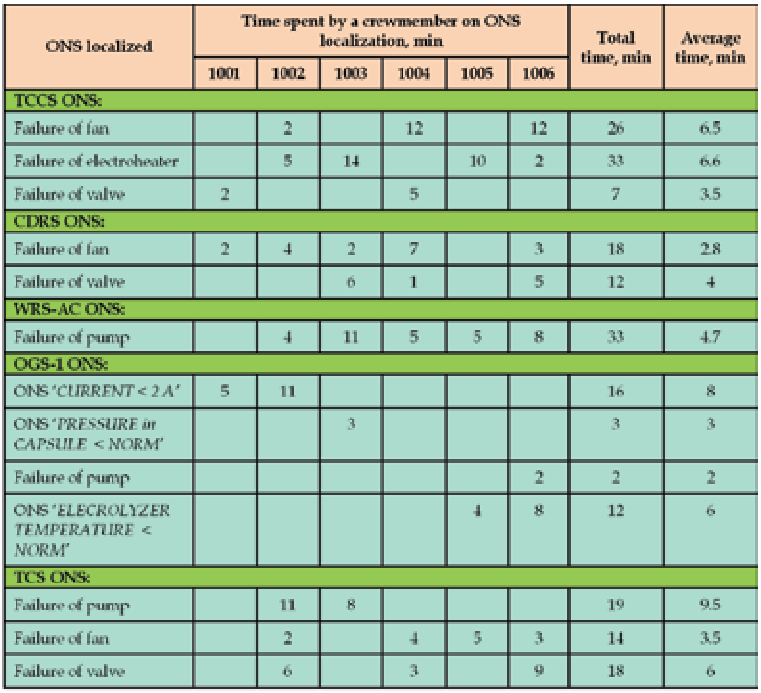
- complexity of necessary camera skills depending on the solved problem of ONS localization;
- time of reaction on ONS being in parameter time-dependent day of the ONS occurrence and from congestion of crew other problems;
- attentiveness of crew during ONS localization;
- time spent for ONS localization;
- total amount of the activities considering total of solved tasks including monitoring of system operation, activity with an electrical simulator and realization of maintenance;
- dynamics of activity formation.

The increased duration of localization of failure indicates absence of attention concentration in some crewmembers that is connected or with realization of some additional activities, either with fatigue, or with presence of distracting factors.

6.2 Results of experiment

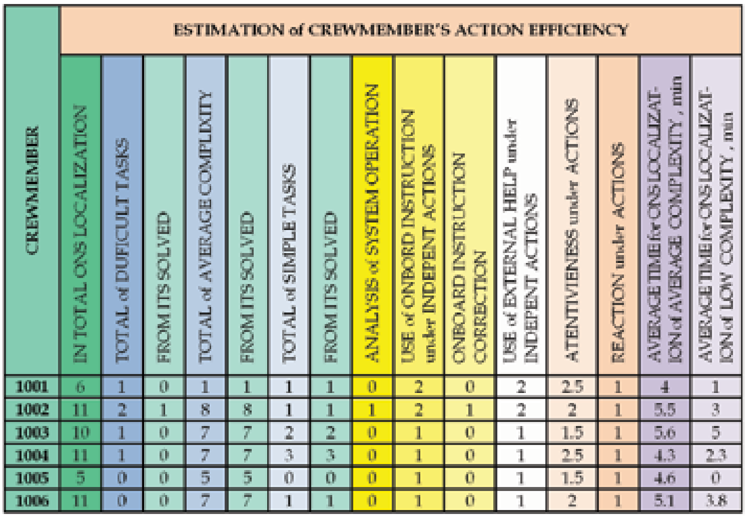
During experiment 52 tasks in total (including activity with the Electron-VM electrical trainer) are generated. 51 tasks are successfully solved. Views generated of the ONS and times expended on its localization are given in Table 2. Results of the crewmember action efficiency estimation in the ONS localization are presented in Table 3.

The typical results of experiments are shown on Figure 19 and Figure 20.



ONS localized	Time spent by a crewmember on ONS localization, min						Total time, min	Average time, min
	1001	1002	1003	1004	1005	1006		
TCCS ONS:								
Failure of fan		2		12		12	26	6.5
Failure of electroheater		5	14		10	2	33	6.6
Failure of valve	2			5			7	3.5
CDRS ONS:								
Failure of fan	2	4	2	7		3	18	2.8
Failure of valve			6	1		5	12	4
WRS-AC ONS:								
Failure of pump		4	11	5	5	8	33	4.7
OGS-1 ONS:								
ONS '*CURRENT < 2 A*'	5	11					16	8
ONS '*PRESSURE in CAPSULE < NORM*'			3				3	3
Failure of pump						2	2	2
ONS '*ELECROLYZER TEMPERATURE < NORM*'					4	8	12	6
TCS ONS:								
Failure of pump		11	8				19	9.5
Failure of fan		2		4	5	3	14	3.5
Failure of valve		6		3		9	18	6

Table 2. ONS types generated and times expended on its localization.



CREWMEMBER	ESTIMATION of CREWMEMBER'S ACTION EFFICIENCY														
	IN TOTAL ONS LOCALIZATION	TOTAL of DUFICULT TASKS	FROM ITS SOLVED	TOTAL of AVERAGE COMPLIXITY	FROM ITS SOLVED	TOTAL of SIMPLE TASKS	FROM ITS SOLVED	ANALYSIS of SYSTEM OPERATION	USE of ONBORD INSTRUCTION under INDEPENT ACTIONS	ONBOARD INSTRUCTION CORRECTION	USE of EXTERNAL HELP under INDEPENT ACTIONS	ATENTIVIENESS under ACTIONS	REACTION under ACTIONS	AVERAGE TIME for ONS LOCALIZATION of AVERAGE COMPLEXITY , min	AVERAGE TIME for ONS LOCALIZATION of LOW COMPLEXITY , min
1001	6	1	0	1	1	1	1	0	2	0	2	2.5	1	4	1
1002	11	2	1	8	8	1	1	1	2	1	2	2	1	5.5	3
1003	10	1	0	7	7	2	2	0	1	0	1	1.5	1	5.6	5
1004	11	1	0	7	7	3	3	0	1	0	1	2.5	1	4.3	2.3
1005	5	0	0	5	5	0	0	0	1	0	1	1.5	1	4.6	0
1006	11	0	0	7	7	1	1	0	1	0	1	2	1	5.1	3.8

Table 3. Results of the crewmember action efficiency estimation at ONS localization.

With a task of elimination of leakages in the PMM illuminators conducted from 06.29.09 to 07.08.09, the crew has failed. The reason is easing of attentiveness and fatigue of crew at the final stage of the experiment.

Commissioning of Electron-VM electrical trainer has been successfully conducted by the 06.25.09, operator 1002.

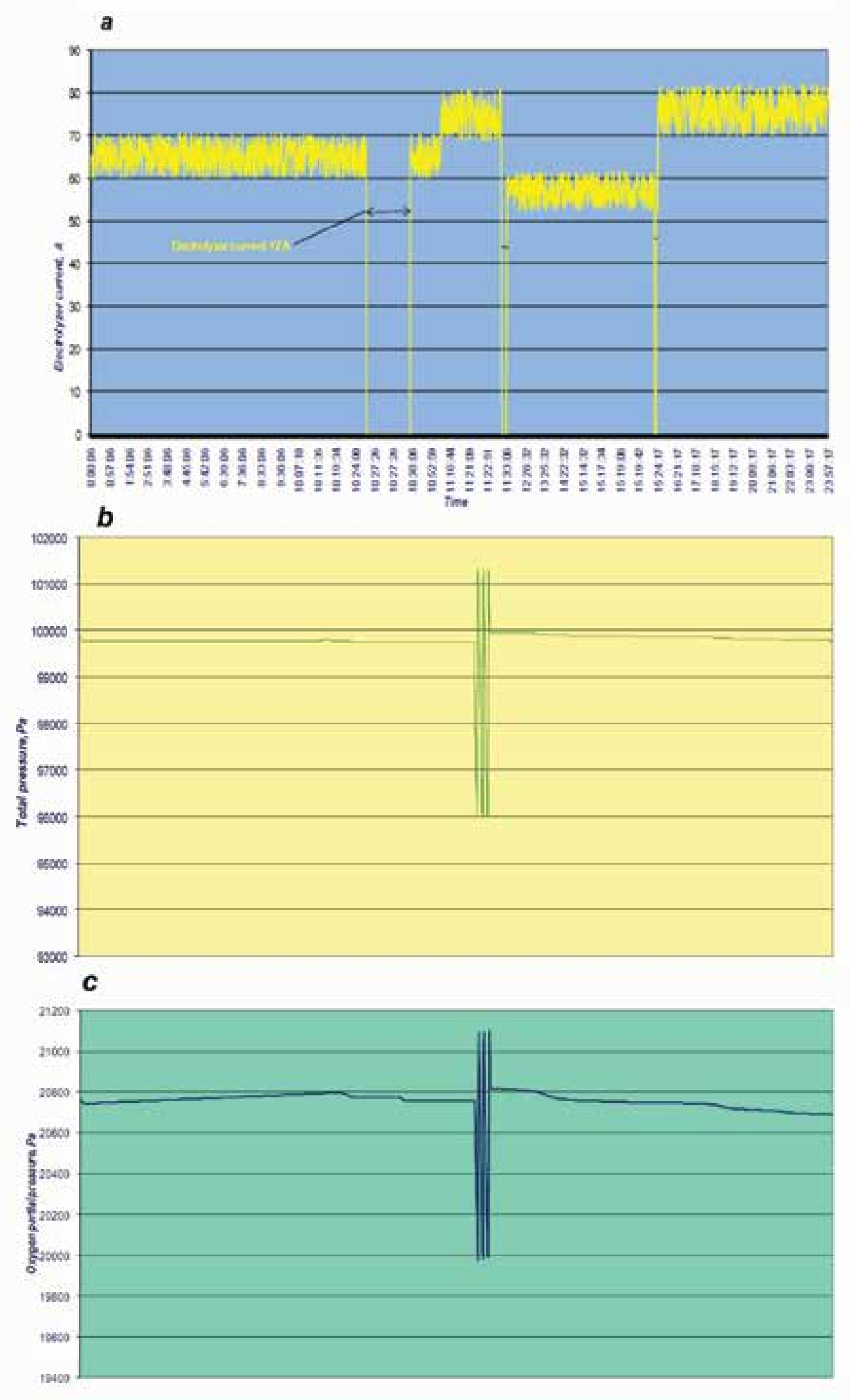


Fig. 19. Localization of the ONS *'ELECTROLYZER CURRENT < 2 A'* (Crewmember 1002): a - electrolyzer current change; b - total pressure change; c - oxygen partial pressure change.

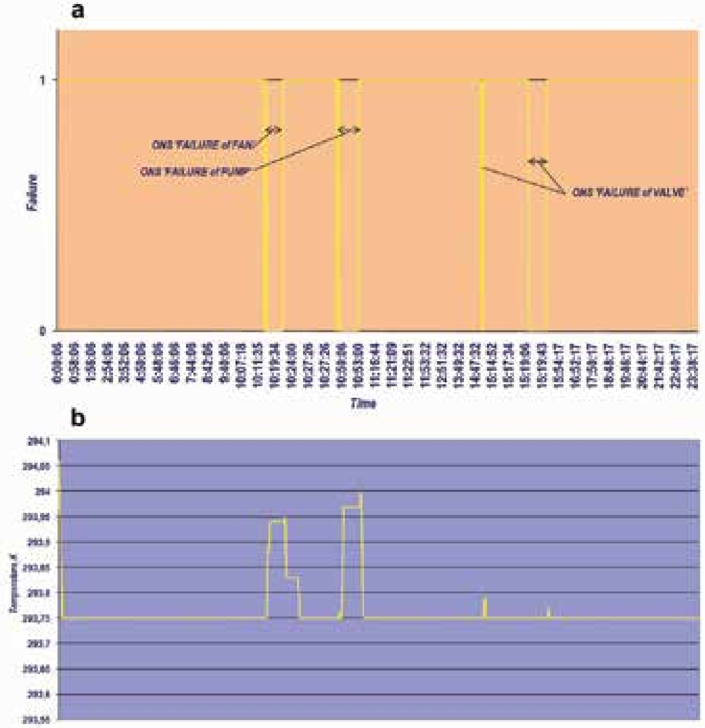


Fig. 20. Localization of the ONS in TCS (Crewmembers: 1005, 1006, 1001): a – failures; b – change of temperature in the PPM.

7. Conclusion

In the chapter the approach to analysis of interaction in the Man-IRLSS system relating to a complex man-machine system, in which human-operator interacts with a technical device during production of environment components, management, processing information, etc is considered.

The offered approach is advanced for the ground medical/engineering experiment imitating interplanetary flight to Mars at which use of standard aboard regeneration life support systems is complicated and is economically unprofitable.

The HSCSOs application has allowed solving the following primary problems:

- to conduct an estimation of crewmember action efficiency at service of the IRLSS and at ONS localization caused by probable failures in its functioning and deviations of the environment controllable parameters from preset values in view of a degree of readiness of crewmembers and conditions of long isolation;
- to research interactions in the IRLSS-Crew system in real time with the purpose of the medical/engineering and ergonomic requirements to IRLSS systems.

Including of standard systems in the HSCSOS architecture is effected by replacement of corresponding virtual simulators with standard systems. As this takes place information channels to and from systems are locked in corresponding logic devices controlling directly actual hardware. Data exchange procedure between specific virtual systems corresponds to a logic structure of flow exchange between actual systems therefore replacement of virtual systems with actual ones will not be problematic.

The considered approaches to research of the IRLSS virtual simulators can be used at development of 'man – machine' for other particular mission.

8. Acknowledgment

Authors express thanks to the colleagues from Joint-Stock Company 'NIIchimmash' Lev Gavrilov, Aleksey Kochetkov, Victor Andreev and Roman Sachkov, as well as the our colleagues from SRC RF – IMBP RAN Juriy Sinyak and Vladimir Trikolkin in many respects promoting fulfillment of the given charter.
Authors are grateful to the reviewer of this chapter Aleksandar Lazinica from CEO for valuable comments made at its preparation.

9. References

Adamovich, B. A., & Gorshenin, V. B. (1997). *Life outside of the Earth.* Moscow, Science.

Baevsky, R.M., Kirillov, O.I., & Kletskin, S.Z. (1984). *The mathematical analysis of changes of a cordial rhythm at stress.* Moscow, Science.

Kurmazenko, E.A., Samsonov, N.M., Farafonov N.S., & Dokunin I.V (1997). S*imulation Subsystem Models for Analysis of the Integrated Regenerative Life Support Systems Functioning. Proceeding of 6th European Symposium on Space Environmental Systems,* Noordwijk, the Netherlands.

Kurmazenko, E.A., Gavrilov, L.I., Samsonov, N.M. Dokunin, I.V., Romanov S.Ju., & Ryabkin, A.M. (1998) *A Man-Made Gas Atmosphere Simulation Model of International Space Station's Russian Segment. SAE* Technical Paper Series #981718.

Kurmazenko, E.A., Samsonov, N.M., Farafonov, N.S., Gavrilov, L.I., Dokunin I.V., & Kotelnikova, M.A. (2000). *Crew software simulation for Integrated Life Support System Operation Study.* SAE Technical Paper Series #2000-2120.

Kurmazenko, E.A., Gavrilov, L.I., Kochetkov, A.A., Khabarovskiy, N.N. Demin, E.P., Grigorjev A.I, & Baranov V.N. (2008). *Life Support System Virtual Simulators for Mars-500 Ground-Based Experiment.* Proceeding of 59th International Astronautical Congress, Glasgow, Scotland.

Kurmazenko, E.A., Gavrilov, L.I., Kochetkov, A.A., & Khabarovskiy, N.N. (2009). *Space Ecological/Engineering System for the Manned Interplanetary Vehicles Crew: Status and Key Technologies for its Development.* Proceeding of the 60th International Austronautical Congress, Daejeon, Republic of Korea.

Kurmazenko, E.A., Kochetkov, A.A., Gavrilov, L.I., Khabarovskiy, N.N., Kamaletdinova, G.R., Morukov,B.V., Demin, E.P., & Trikolkin, V.I (2010). *Crew's Service of the Virtual Integrated Regenerative Life-Support System Operation: 105-Day Experiment Results under Mars-500 Project.* AIAA Technical Paper Series #2010-6168.

Pravetskiy, V.N., Samsonov, V.M., Utyamyshev, R.I., & Kurmazenko, E.A. (1981). *Separate Problems of the Life-Support and Safety System Development for a Spaceflight Vehicle Crew. Scientific Reading on Aircraft and Astronautics 1980.* Moscow, Science.

Savina, V.P., & Kuznetsova, T.I. (1980) *Trace contaminant sources and their toxicological evaluation.* Space Biology, *v. 42,* Moscow, Science.

User guide complex BioMouse CPP and CPP-01-01b (options "Professional" and "Research "), introducing the principles of operation of the product with the functions of the software, and implemented methods (2008). Moscow, ZAO "Neurolab", www.neurolab.ru

23

Educational Opportunities in BME Specialization - Tradition, Culture and Perspectives

Wasilewska-Radwanska Marta[1], Augustyniak Ewa[2],
Tadeusiewicz Ryszard[1,3] and Augustyniak Piotr[1,3]
[1]*Multidisciplinary School of Engineering in Biomedicine,*
AGH-University of Science and Technology,
[2]*Faculty of Humanities, AGH-University of Science and Technology,*
[3]*Faculty of Electrical Engineering, Automatics, Computer Science and Electronics,*
AGH-University of Science and Technology,
Poland

1. Introduction

1.1 The traditions of biomedical physics and engineering in Poland

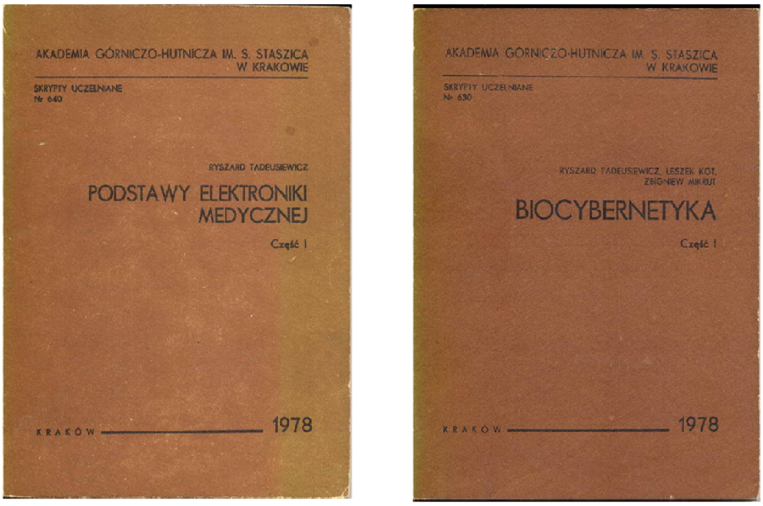
Medical physics and engineering education in Poland started in the 1930s with the foundation of the Radium Institute in Warsaw by Maria Sklodowska-Curie. Prof. Cezary Pawlowski, one of the assistants and then collaborators of Mme Curie (fig. 1), organized the first courses in medical physics and biomedical engineering at the Physics Department of the Radium Institute.

The first course in medical engineering started at the Faculty of Electrical Engineering of the Warsaw University of Technology in the 1950s. Then, at the Faculty of Electrical Engineering, Automatics, Computer Science and Electronics of the AGH University of Science and Technology (former University of Mining and Metallurgy) in Krakow, Prof. Ryszard Tadeusiewicz organized the first courses in biomedical engineering in the 1970s. Fig. 2 shows the first Polish textbooks in Medical Electronics and in Biocybernetics. Note the year of the issue of both books, 1978.

Until the academic year 2005/2006, education in biomedical engineering was offered only as a specialization in other fields of studies, e.g. mechanics, automatics & robotics and electronics. The development of new technologies in medical diagnosis and therapy required a new approach to biomedical engineering education. Therefore, a consortium was set up of six technical universities (in alphabetical order): The AGH University of Science and Technology (Krakow), The Gdansk University of Technology (Gdansk), The Silesian University of Technology (Gliwice), The Technical University of Lodz (Lodz), The Warsaw University of Technology (Warsaw) and The Wroclaw University of Technology (Wroclaw). The consortium developed a new programme of education and then applied to the Ministry of Science and Higher Education for an official permit to create a new field of studies referred to as "Biomedical Engineering" (BME). In June 2006, the Ministry gave its consent to this proposal. The AGH University of Science and Technology was first in Poland to enroll students in BME in the academic year 2006-2007. In 2007-2008, all the members of the consortium had students in BME. In the academic year 2010-2011, BME education is being offered by 16 technical universities in Poland (Table 1).

Fig. 1. Dated 1911. Probably one of the earliest photographs to show Maria Sklodowska Curie (first from the left), a double Nobel Prize laureate and professor at the Sorbonne (Paris), and Walery Goetel (first from the right), professor and future rector of the Mining Academy (now the AGH-UST), on a mountain trip which could have led to Maria Sklodowska Curie's becoming three years later Director of the Red Cross Radiology Service during World War I. She is also known to have converted many ordinary cars into ambulances equipped with mobile radiology units. These cars, called "petite Curie," transported X-ray apparatus to the wounded at the battle front, thus saving the lives of many French soldiers.

Medical physics education in Poland started in 1950 with the technical physics specialization created by Prof. Cezary Pawlowski at the Warsaw University of Technology, and at the AGH University of Science and Technology (former University of Mining and Metallurgy) in Krakow by Prof. Marian Miesowicz. In the 1970s, a medical physics programme was initiated at Warsaw University and at the Jagiellonian University in Krakow. In 1990, a specialization in Radiation Physics and Dosimetry was started at the AGH University of Science and Technology in Krakow, which since 1991-1992 has been run as Medical Physics and Dosimetry in close cooperation with the Collegium Medicum (Faculty of Medicine) of the Jagiellonian University. In the academic year 2010-2011, about 15 universities and technical universities offered courses for students in medical physics (Table 2).



AKADEMIA GÓRNICZO-HUTNICZA IM. S. STASZICA
W KRAKOWIE

SKRYPTY UCZELNIANE
Nr 640

RYSZARD TADEUSIEWICZ

PODSTAWY ELEKTRONIKI
MEDYCZNEJ

Część I

KRAKÓW 1978

AKADEMIA GÓRNICZO-HUTNICZA IM. S. STASZICA
W KRAKOWIE

SKRYPTY UCZELNIANE
Nr 630

RYSZARD TADEUSIEWICZ, LESZEK KOT,
ZBIGNIEW MIKRUT

BIOCYBERNETYKA

Część I

KRAKÓW 1978

Fig. 2. The first Polish textbooks for biomedical engineering students issued by AGH-UST.

2. A multidisciplinary school – organizational background and curricula

As teaching Biomedical Engineering requires specialists representing many different areas of research and different competences, it was impossible to select a staff on the basis of the individual personnel resources of one particular faculty. Therefore to teach biomedical engineering students, a Multidisciplinary School of Engineering in Biomedicine (MSIB – the acronym used in all Polish documents describing the aims and the scope of the school) was founded (fig. 3). Below you will find the general outline of the structure and main guidelines for the School (Wasilewska-Radwanska & Augustyniak, 2009).

Fig. 3. Rector of the AGH-UST, Professor Ryszard Tadeusiewicz, signs the MSIB foundation act (2005).

2.1 The external situation

The external situation in Poland in 2005, when the Multidisciplinary School of Engineering in Biomedicine was founded, provided several opportunities and challenges (Augustyniak 2008). These initial external conditions can be classified into three groups:

The first group referred to Polish medical technology-related enterprises and/institutions:

- Local industry was rather undeveloped; we estimate the number of local medical technology-related enterprises to be about 100, but most of them (40%) were very small businesses, so-called micro-enterprises, having 1-5 employees, or small enterprises (30%) with 6-50 employees. Bigger enterprises were usually sales- or service representatives of international corporations, without independent human resources management (fig. 4);
- The relation between research and industry was weak; the way from technical innovation to marketing of a final product was very formalized;
- The results of research done by technical universities were financially unattractive and did not match industry needs; industry management preferred independent research rather than cooperation with universities;
- The average technological level of the health care was low, with notable exceptions in some selected centers.

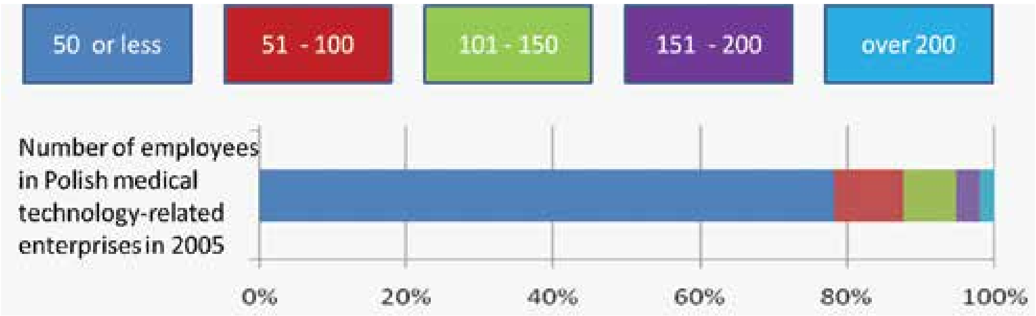


Fig. 4. The employment structure in Polish medical technology-related enterprises in 2005 (ROTMED Consortium 2006).

The second group was related to Polish medical technology needs based on social demands:

- There was an urgent and important need for the development of medical technology because of the poorer quality of social health services in Poland compared with those in highly developed EU countries (fig. 5)

The third group was related to Polish university traditions and traditional models of teaching:

- The experience with two-tier structure of degree courses/university studies (Bachelor's and Master's) was very inadequate; there were no clear guidelines for curricula, syllabuses and examinations, nor for assessment of the teaching quality; the existing government regulations were insufficient;
- There was no experience of teaching in English, and professional bibliography in Polish was very limited;
- The organizational chart of a multidisciplinary school was innovative and rarely implemented by universities, the university funds' distribution mechanisms being inadequate;

- Biomedicine-related research was carried out in several faculties within the framework of other disciplines such as computer or material sciences, electrical or mechanical engineering; there was an adequate number of professors and assistant-professors representing high-performance field-oriented output;
- The principles of school organization enabled quality-based staff selection unlimited by state employment regulations;
- There was a growing interest in medical technology from good candidates; and
- Some recent governmental regulations, aiming at improving the quality of health care, facilitated the employment of clinical engineers.

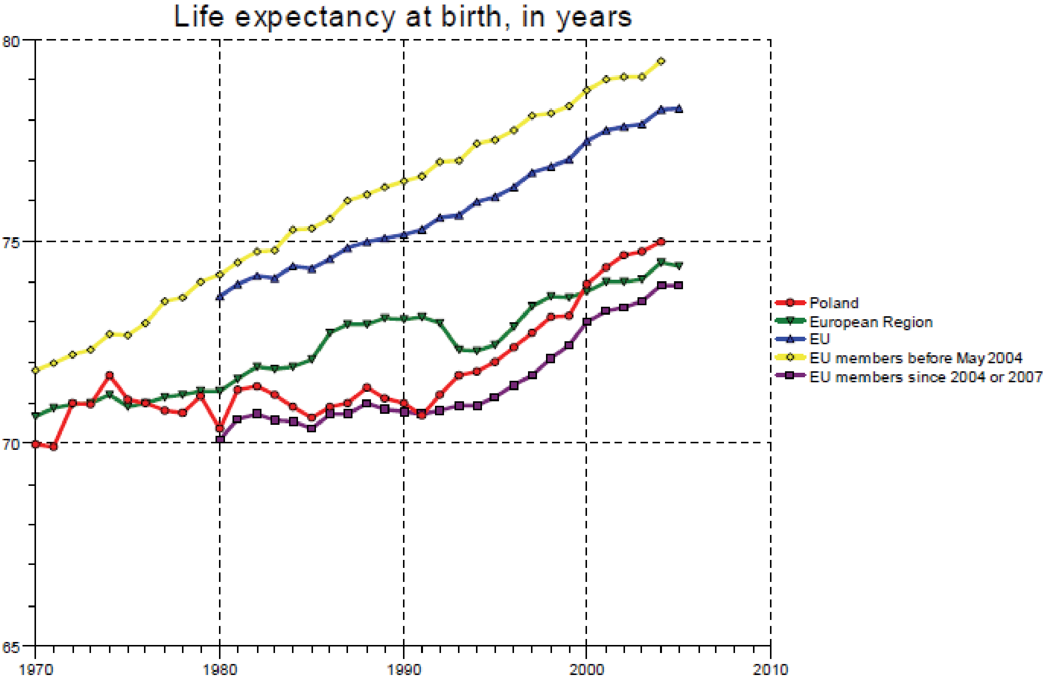


Fig. 5. A comparative plot showing the situation in social health services in Poland vs. the developed EU countries (ROTMED Consortium 2006).

All the above issues and challenges have led to ever stronger inter-university cooperation and integration into society. Representatives of eight Polish universities made every effort to establish educational standards in biomedical engineering as a separate field of study currently offered in Poland for about 1000 candidates each year.

2.2 The organizational scheme and place of MSIB in university structure

The MSIB is located at the AGH University of Science and Technology and has been in operation since the academic year 2005-2006. Although MSIB has been formed on the basis of the staff formally belonging to five faculties, it is treated as a separate part of the AGH University and has its own students. Formally, MSIB's structure is similar to that of other faculties. It is governed by a Board of 18 persons. This Board, approved by the University Senate, is made up of professors with not less than a DSc degree who are teaching at MSIB, as well as of an adequate representation of students. At present, the professors represent five faculties:

- Faculty of Electrical Engineering, Automatics, Computer Science and Electronics,
- Faculty of Materials Science and Ceramics,
- Faculty of Mechanical Engineering and Robotics,
- Faculty of Metals Engineering and Industrial Computer Science, and
- Faculty of Physics and Applied Computer Science.

One of the Board’s tasks is to recommend to the Rector appointments for the Head and the Deputy Head of the School. The appointed Head is also President of the Board. The main responsibility of the Board is to supervise the education process, assure its highest quality, verify and, if necessary, correct academic curricula, prepare staff assignments and implement other objectives of the School. The Head also represents the MSIB in the University Board on par with deans of other faculties.

From the student's viewpoint, there is no organizational difference between the faculty and the Multidisciplinary School. Both have a Dean's Office, a staff of qualified teachers, a social support system and a student board. As far as education is concerned, the rights and responsibilities of the Head of the School are identical to those of a Dean, the only difference being that research is carried out in laboratories in various faculties run by individual professors rather than in the organizational framework of MSIB.

Since medical sciences are not represented in the AGH-University of Science and Technology, six medicine-oriented lectures (e.g. anatomy, physiology, medical deontology, history of medicine) are given by professors of the Collegium Medicum (Medical College) of the Jagiellonian University. The agreement between the universities gives students the opportunity to attend lectures and to participate in laboratory exercises in the Faculty of Medicine. This cooperation is mutually beneficial since it provides an alternative, i.e. medicine-based viewpoint, for our medical colleagues and medicine students. Unfortunately, current medical curricula in Poland do not include engineering aspects in medicine, however some lecturers from AGH-UST are among those who take part in postgraduate studies and technology-oriented teaching projects for medicine students or medical doctors.

2.3 General layout of curricula

The BME teaching programs in the Multidisciplinary School of Engineering in Biomedicine AGH University of Science and Technology follow all the Polish legal regulations, including the national standards for academic teaching set out by the Ministry of Science and Higher Education (Ministry of Science and Higher Education 2007), and the guidelines of the Bologna Process (including the Educational Credits Transfer System and Accumulation System-ECTS). The current program presented in the block diagram in Fig. 6 consists of:

- A single 7-semester track leading to the First (Undergraduate) Degree (Bachelor’s/Engineer’s);
- Five domain-oriented 3 or 4-semester tracks leading to the Second (Graduate) Degree (Master’s);
- A single 8-semester track leading to the Third Degree (Doctor’s).

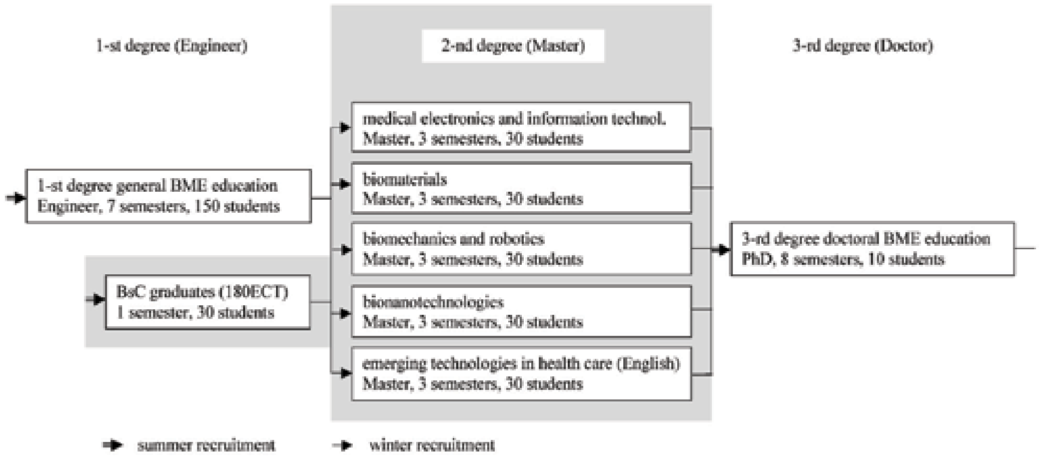

1-st degree (Engineer)
2-nd degree (Master)
3-rd degree (Doctor)
medical electronics and information technol.
Master, 3 semesters, 30 students
1-st degree general BME education
Engineer, 7 semesters, 150 students
biomaterials
Master, 3 semesters, 30 students
biomechanics and robotics
Master, 3 semesters, 30 students
3-rd degree doctoral BME education
PhD, 8 semesters, 10 students
BsC graduates (180ECT)
1 semester, 30 students
bionanotechnologies
Master, 3 semesters, 30 students
emerging technologies in health care (English)
Master, 3 semesters, 30 students
summer recruitment
winter recruitment

Fig. 6. An education track block diagram for biomedical engineering at MSIB AGH-UST.

After a careful review of the needs from prospective employers' point of view, the availability of existing infrastructure and resources and a detailed study of reports from more experienced colleagues, we have decided to formulate and put into practice several rules and mechanisms that provide a broad basic education in all possible BME domains and fast adaptation of the program to the variability of an unstable, constantly developing local employment market.

A complete description of the offer consisting of 189 lectures [64 for the Bachelor's (First) degree , 104 in five domains for the Master's (Second) degree and 21 for the Doctor's (Third) degree)] is available from the MSIB web service www.biomed.agh.edu.pl (see Appendix, table 3). By way of example, we have selected two curricula for the specialization called "Emerging Health Care Technologies" taught in English for the Master's Degree and for the Doctor's Degree (PhD) studies (see Appendix, table 4) and three course sheets -syllabuses (see Appendix, table 5). Some selected themes for PhD theses may also be worth mentioning:

- 3D Reconstruction of Brain Glial Cells (2010),
- Automatic Facial Action Recognition in Face Images and the Analysis of Images in the
- Human-Machine Interaction Context (2009),
- 3D Segmentation of Medical Data from Computed Tomography and Endoscopic Video Records (2008).

Full texts of these theses are available on-line from the Biomedical Engineering PhD students' web service www.embio.agh.edu.pl.

2.4 Program adaptation mechanisms

In addition to the Bachelor's Degree program of a single track (for 150 students each year), various measures were implemented in order to increase the adaptation range of the proposed tracks. A high degree of flexibility, and a wide offer of elective courses involve students as partners in the educational process. This method of program organization with the essential participation of the students has an additional advantage: it helps to develop the students' responsibility and flexibility that is necessary in the difficult workplaces employing biomedical engineers.

The Second Degree (Master's) program (also proposed for 150 students per year) requires that students select one of five offered parallel tracks. Four of them are taught in Polish and oriented towards the main branches of biomedical engineering:

- Medical electronics and information technologies,
- Biomaterials,
- Biomechanics and robotics,
- Bionanotechnologies,
- Emerging health care technologies.

Fortunately, the formula of the multidisciplinary school that is based on human resources and the infrastructure of five faculties takes advantage of the support thus provided and is sufficient to comply with the requirements of high-quality teaching.

This variety of human and technological resources in the absence of institutional limitations and work contracts makes it possible to have a free hand in creating teaching syllabuses depending on the current market needs and students' interests. The creation of learning programs (new syllabuses and specialties) which comes under the authority of the Head of the faculty is also supplemented by three types of individual lines of study:

- Implementation of some selected additional subjects of study in addition to the standard syllabus;
- Individual schedule of study. i.e. implementation of the standard program following the accelerated of decelerated schedule, or
- Individual course of study making possible the selection of subjects outside the standard syllabus to replace elective subjects.

The creation of the course of study is based on quality indexes systematically collected from students, lecturers and employers[1]. The assessment of the quality of the learning process is regularly made by:

- Estimation of current learning progress made by lecturers;
- Estimation of learning results by the Head of the Faculty based on statistical evaluation;
- Questionnaires filled in by students about their lecturers (carried out by Department Deans and made public in a adapted from to the Head of the MSIB);
- Questionnaires filled out by students on the course of their studies;
- Questionnaires on the professional abilities of students carried out in institutions responsible for the students' field work.

All the quality indexes are analyzed by the Board and are used in making the Head responsible for implementing decisions such as a face-to-face talk with the lecturer, change of the lecturer, modification in the sequence of subjects or a change in the syllabus.

3. Specific characteristics of BME-related corporate culture

Although every profession has its own best business practices, biomedical engineering has some specific characteristics. This relatively new discipline bridges the gap between medicine and technology, and, by applying various novel methods and techniques, directly influences the duration and quality of human life (Schwartz 1988). Consequently, a single medical procedure involves multiple actors, many of whom, mainly engineers, usually remain behind the scene. The cooperation between engineers and members of the medical profession can succeed only if they all share some common understanding about their roles and mutually dependencies. Moreover, a medical procedure requires, among other things, technical excellence and high ethical awareness of the participating engineers. Finally, the clinical engineer frequently works shoulder to shoulder with medics in an emergency and under time pressure, burdened by stress and responsibility, where common feelings, ideas and behavior are a decisive factor in the final result. Therefore three key aspects may be identified as those related to BME professional culture:

---

[1] In future

- Understanding and good cooperation within a multidisciplinary team;
- Striving for the technical excellence and
- Human-centric ethical background.

As such circumstances are specific to this profession, the need for the organizational culture is fairly high. In case of a biomedical engineer, however, the relations are far more complex than the schemes sketched by corporate culture promoters for a typical customer service (Denison & Mishra, 1995). The above-mentioned elements of the BME-related corporate culture cannot be formed only as spontaneous forms of behavior of clinical engineers. They must be prepared by special methods and techniques of teaching and learning to help young students become highly qualified and well-prepared specialists also in the ethical dimension of their work. This process of ethical formation must be closely combined with professional skill formation, because ethics is not simply an additional competence of the future biomedical engineer, but rather a crucial element determining the use of basic knowledge in all practical applications. Therefore teachers who teach biomedical engineering students must take this additional demand into account (Augustyniak E. & Augustyniak P. 2010).

3.1 Mutual understanding and good cooperation within a multidisciplinary team

Biomedical engineering integrates various technical disciplines, but it also assumes the presence of an engineer in every situation where medicine is being practiced. Therefore, one of the principal requirements is the ability to work in a multidisciplinary team, in which common ideas, values and behavioral patterns have a very practical implication influencing the emergency response, adaptation to conditions of any healthcare mission and a sense of intellectual curiosity about technological development. In such teams, various professional learning-related particular interests are present, and mutual understanding instead of competition is the key to success.

The organizational culture of the School created by all workers from various departments of the School fosters an emphasis on teamwork and intercommunication, especially in the group of specialists from various fields. Common understanding, personal involvement and dedication constitute an essential pattern of a learning organization. In a team which includes a physician, nurse, pharmacist and an engineer, the effective use of knowledge and skill by separate persons is dependent on the mutual understanding within each person's field of competence. In addition, the cooperating specialists motivate each other and thus create novel solutions resulting from the synergy of various experiences.

3.2 Striving for technical excellence

Working in a team with a doctor, a nurse and a pharmacist, the biomedical engineer is responsible for all technology-related ideas, and is perceived as someone capable of solving all problems in his or her area of responsibility; all of this combined may become a source of abnormal stress. Moreover, the support of life, as no other application of technology, has a direct relation to the human being, his or her health and happiness and to fundamental values. Consequently, an assumption of excellence distinguishes a biomedical engineer from engineers of other specialties. Furthermore, since technology is currently responsible for the ever greater efficiency and effectiveness of medicine, it can also be blamed for any adverse effects of medical procedures.

Striving for excellence at all levels of the implementation of detailed goals of achieving cooperation and suggesting technical solutions to problems is a typical characteristic of the organizational culture. In addition to tasks involving the implementation of educational programs (curricula and syllabuses), one of the School's main objectives is to transmit standards. Thus, when the School becomes an archetype or model of the future workplace and the lecturers employed by the School are expected to set an example of how excellence should be achieved, the care taken to provide a proper form of the organizational culture is a *sine qua non* educational requirement (O'Reilly & Chatman, 1996). If the future graduate student acquires adequate competences during the course of his or her studies, their position will be determined in a multidisciplinary team in real situations in health services. Organizational rules and regulations that are in force in the School, which employs lecturers from various faculties, make the presentation of the lecturers' various roles in the learning process easier by taking into account the person's specific knowledge, skills and attitudes. This approach has been used to make an analogy pointing to the technical excellence of an engineer as a basic argument supporting his or her key role in the multidisciplinary team employed in health services.

Fig. 7. Students discussing laboratory projects (2008).

3.3 The ethical background

Engineers who are specialists in other fields of technology than medical or biological engineering are rarely confronted with human life in its various forms and stages, such as birth, adolescence, disablement, disease, pain, death, etc. Therefore, special forms and patterns of education should be dedicated to attitudes about human nature in general. Since the engineering process in biomedicine is always part of a medical procedure, the ethical aspects of medicine should be a mandatory element of lectures and practical work in a BME-teaching institution.

The attitudes, norms and values promoted in the organizational culture of the School have an effect on the way how the future graduates will think and act in their professional life. The sense of responsibility and reliability combined with a sense of humanitarianism, curiosity and creativity are among the most desirable attributes of a biomedical engineer, while also held in high value by employers in other fields. The above values are not easily transferable to traditional university teaching methods such as lectures, problem-solving or experiments, therefore elements of the organizational culture may be of some help in this context.

The most important elements of this culture are:

- Forms of a specific tradition established within a given group to strengthen the feeling of identifying with the group's aims;
- Undertaking novel professional activities such as writing textbooks, promoting the new profession in mass media or encouraging volunteer work by both lecturers and students;
- Permanent evaluation of the School's activity and the teaching and learning programs based on external assessment and the opinions provided by lecturers and students.

4. Opportunities and perspectives of BME-teaching institutions

The MSIB authorities have designed a unique, well-balanced corporate culture program and continue to develop it with success. The MSIB itself as a teaching institution was started from scratch, so from the beginning it was based on the general principles of a learning organization.

4.1 A blueprint for an organizational culture

The establishment of a new institution created an opportunity for devising a profession-specific, independent system of shared values and beliefs implying and reinforcing a collective identity centered on the mission of the biomedical engineer. Our aims exceeded "standard faculty" programs, providing students with government-designed minimum curricula. Our main objective is for MSIB to become a leading research and educational center (Augustyniak et al. 2010). Despite a relatively short history of only five years, the School has undoubtedly earned high marks in Poland and has become well-recognized abroad. This was possible due to:

- Employment of the best professionals in the field as teachers;
- Active creation of cooperation links with other universities in Poland and abroad;
- Publication of a BME-student quarterly - a periodical promoting the profession and reinforcing the students' intellectual confidence and emotional group commitment;

- Leadership and promotion of common activities, exchanges, encouragement of staff and students' mobility, additional lectures etc.;
- Dissemination of knowledge and participation of staff members in international conferences and their activity in commissions and other opinion-forming bodies.

Fig. 8. Professor Zbigniew Kakol, Rector of AGH-UST, confers an Engineering Diploma (2010).

The organizational culture of the MSIB is based on a deeply motivated concept of a high-quality teaching and learning institution, aiming at the education of marketable BME professionals (Gordon & DiTomaso, 1992). The School is student-friendly and supports students' educational initiatives in various fields of activity. What is expected and/or requested from the staff and students are well-defined goals to achieve. In this spirit, thanks to the common sense of responsibility, the internal atmosphere in the School supports cooperation rather than competition.

To make common attitudes and values more popular and generally recognized, the organizational culture of the MSIB offers the following:

- Ceremonies (matriculation, the Dean's address at the New Year, summarizing achievements challenges, and future plans, etc.);
- Corporate design (a well-recognized logo, a precisely defined color, organizational badges for Board members, corporate T-shirts for students);
- Outstanding personalities - undoubtedly this honor goes to Professor Ryszard Tadeusiewicz, the founder of the School and former Rector of AGH-UST, a pioneer in biomedicine and author of several BME-related books since the mid-70s;
- Common recreational activities (mountain trekking, sailing, skiing, student sports teams in various disciplines, volunteer-based events in hospitals and hospices, initiatives on behalf of the handicapped, extramural students' organizations and forums).

The corporate culture is built on selected values and behavioral norms that influence both the employees and students through the internal and external institutional policy. The commonly accepted standards constitute the internal ambience and friendliness of the School, but as a long-term investment become everyday habits. We easily observe teachers transferring these standards to their own departments, which justifies the belief that in the future the graduates will also promote such standards in their workplaces.

Fig. 9. Informal mountain trekking of MSIB students and the Dean (first from the left).

4.2 The employment perspectives of graduates

The high quality of educational outcomes is of vital interest for students, their prospective employers and university staff. All three groups of partners aim for a common goal using different approaches and capabilities. Cultivating the understanding of the School as a common value and sharing the collective identity and commitment reinforces compatibility and multi-professional team building. This training field for interpersonal learning reveals three enhancing feedback loops:

- Attractive employers increase students' motivation and good graduates extend the development perspectives of their employers;
- Innovative employers offer career opportunities for good researchers in applied sciences, and more effective professors provide high-quality knowledge to industry which in turn enhances the university's reputation;
- Good candidates are inclined to look for good lectures and challenging projects, and become a source of professional satisfaction to professors who make efforts to improve their lectures.

These relations, demonstrated above in a simplified form, are fundamental for the team-building spirit, which is based on the commitment of students, tutors and professionals to a quality-focused efficient education. In the MSIB, the healthcare professionals participate in the preparation of curricula, provide opportunities for one-month internships and formulate challenges for young teams of scientists. The university staff cultivates the students' creativity by supervising laboratory exercises and interim projects which constitute a demonstration of the graduates' skills, and can be partly applied in healthcare and industry. The excellence of the university staff in science is also a good example attracting best students who often try to follow or to participate in research themselves.

The most highly appreciated educational initiatives are those focused on education-oriented basics, behavioral norms and values including self-teaching. Equally important to scope and breadth of knowledge are technical skills and human-centric attitudes. The basics of human sciences, ethics and law are indispensable elements of education and practice of the School. Two other important features are:

- Flexibility which makes it possible for the School to adapt itself to changing trends of the employment market and which enables graduates to keep pace with quickly-developing technical sciences; and
- A set of relations between the teacher and/or the dean and the student which focuses on communication and exchange of information. This not only makes relations within the School administration easier but also establishes a pattern for future relations in a multi-professional environment.

Current employment perspectives mostly result from the growing importance of extensive use of technology in health care. Clinical engineer, manufacturing engineer, researcher and sales representative are the four main types of specializations in terms of future professional careers. Nevertheless, local employment markets are weak, and the effort to raise the awareness of a new profession to a higher level is among the principal concerns of the MSIB Management. General statistics are favorable for the AGH-UST graduates: 75% of them find their first job within one month after graduation, and 95% within three months. According to other surveys, in 500 of Poland's biggest enterprises, AGH-UST graduates constitute the second largest percentage among senior management staff.

5. Concluding remarks

The special demands imposed on BME education require that we should seek unusual organizational solutions and non-standard teaching and learning methods, related to the specific and unique educational challenges posed by BME. We hope that these theoretical and practical methods and techniques developed at the MSIB and presented above may be of interest to other educational institutions in BME field.

6. Acknowledgments

This work was supported by the Rector of the AGH University of Science and Technology in Krakow, Poland. The authors want to express their gratitude to the Walery Goetel family for their consent to publish the archival photo of Maria Sklodowska Curie.

7. Appendix

Institutions	Faculty/ Department	Organisation (current state)/specializations
AGH Univ. of Science and Technology (Krakow)	Electrical Eng., Automatics, Comp. Sc.& Electronics	Multidisciplinary School of engineering In Biomedicine (MSIB)/Med. Electronics&Inform.; Biomat.,; Biomech.&Rob.;Bionanot; ; Emerging Techn.Health Car
Bialystok Univ. of Techn.(Bialystok)	Mechanical Engineering	/Biomechanics; Rehabilitation Engineering; Medical Materials
Czestochowa Univ. of Techn. (Czestochowa)	Mechanical Engineering & Comp. Sc.	Rehabiltation Engineering; Medical Informatics.
Gdansk Univ. of Technology (Gdansk)	Electronics, Telecomunications &Informatics	Inter-Faculty full-time studies /Inf.Med.; ElectrM.; ChemM; PhysM
Koszalin Univ. of Techn. (Koszalin)	Institute of Mechatronics, Nanotechnology& Vac. Techn.	/Implants; Medical Instrumentation; Manipulators
Krakow Univ. of Technology (Krakow)	Mechanical Engineering	/Clinical Eng. Dental Biomechanics; Biomechanics of Injuries
Lublin Univ. of Techn. (Lublin)	Mechanical Engineering	In cooperation with Lublin Medical Univ./CT; MR; Aut&Rob; Biomat Art Organs; M Electr ;Infor in Medicine
Lodz Univ. of Techn. (Lodz)	El., Electronic, Comp.& Control Eng.	Inter-Faculty Education Centre with the Fac. of El., Electronic, Comp.& Control Eng./Biomats; Biocorrosion; Biomeasurements
Poznan Univ. of Techn. (Poznan)	Mechanical Engineering&Management	Mechatronics/ Biomedical Engineering
Silesian Univ. of Technology (Gliwice)	Biomedical Engineering	/InformaticsM; M Electr.;Sensors & Biomed;; Implants Reahabilitation
Silesian Univ. (Katowice)	Computing Sc. and Material Sc.	/ Biomed.Inform;; Med. Imaging; Telemed&Clinical Inform. Systems;
Warsaw Univ. of Technology (Warsaw)	Mechatronics/ I. of Metrology&Biomed. Engineering and Electronics and Information Technology	Inter-Faculty studies:e F. of Mechatronics& F. of Electronics&Inf. Techn./Biomech.; Biomat.s; Biomed. Sensors; Biomed.Imaging; Rehab.Eng; ;Prostheses&Art Organs; Clin.Eng. ; Biotechnology; Medical Informatics
Wroclaw Univ. of Technology (Wroclaw)	Fund. Problems of Technology	/Electromed Equip; Smart Transducers; ; Fiber Optics&d Lasers in Med.Equip.; Computer MDiagn.
Zielona Gora University (Zielona Gora)	Mechanics; Electrical Engineering, Comp.Sc.&Telecommunic.; Biological Sc.	Inter-Faculty studies. /Medical Informatics; Medical Electronics; Biomechanics; Biomaterias
Bydgoszcz Univ. of Tech. (Bydgoszcz)	Mechanics	/Technical Medical Consulting; Medical Informatics
West Pomeranian Univ. of Techn. (Szczecin)	Electrical Engineering	/Biomedical and Acoustic Engineering

Table 1. A list of institutions which provide education in Biomedical Engineering (BME) in Poland (since academic year 2010-2011).

Institutions	Faculty/ Department	Discipline	Specialization
Univ. of Bialystok (Bialystok) www.uwb.edu.pl	Physics	P*	MP
Univ.of Gdansk (Gdansk) www.ug.gda.pl	Math, Physics &Informatics in cooperation with Gdansk Medical Univ.	MP*	MP
Univ.of Silesia (Katowice) www.us.edu.pl	Math., Phys. and Chem.	MP	Clinical Dosimetry; Optics in Medicine; Electroradiology
Jan Kochanowski Univ. of Humanities and Sc. (Kielce) www.ujk.edu.pl	Math.&Sc.	P	MP
AGH Univ. of Sc .&Techn. (Krakow) www.agh.edu.pl	Phys.& Appl. Comp. Sc.	MP	Dosimetry and Electronics in Medicine; Imaging and Biometrics
M. Curie-Sklodowska Univ. (Lublin) http://mfi.umcs.lublin.pl	Math., Phys. Comp. Sc.	P	Molecular and medical biophysics
Un. of Lodz (Lodz) www.uni.lodz.pl	Phys& Chem.	P	MP
Techn.Univ.Lodz (Lodz) www.p.lodz.pl	TPhys., Comp Sc.& Appl.Math.	TP*	MP
Univ.of Opole (Opole) www.wmfi.uni.opole.pl	Math., Phys & Chem.	MP	MP
Adam Mickiewicz Univ. (Poznan) http://amu.edu.pl	Physcis	B****	MP
Ignacy Lukasiewicz Techn.Univ. (Rzeszow) http://portal.prz.edu.pl	Math.& Appl. Physics	TP	Physics and Informatics in Medicine
Univ.of Szczecin (Szczecin) www.us.szc.pl	Math.& Physics	P	Biomedical Physics
Nicolaus Copernicus Univ. (Torun) www.umk.pl	Phys., Astr& Informatics	MP	MP and Comp. Appl.
Warsaw Univ. (Warsaw) www.uw.edu.pl	Physics/Inst. of Exp.Phys.	APBM*****	MP
Warsaw Univ. of Techn. (Warsaw) www.pw.edu.pl	Physics	TP	MP
Wroclaw Univ. of Techn. (Wroclaw) www.pwr.wroc.pl	Fund. Problems of Techn.	TP	Nano-engineering
Univ. of Wroclaw (Wroclaw) www.uni.wroc.pl	Physics and Astr.	TP	MP

*)MP=Medical Physics ; **) P=Physics ;***)TP=Technical Physics; ****B=Biophysics; *****APBM=Applications of Physics in Medicine and Biology

Table 2. A list of institutions which provide education in Medical Physics (MP) in Poland (since academic year 2010-2011).

semester	First Degree(Bachelor's)
1.	• Information technologies • Mathematics • Physics • General chemistry • Biocybernetics • Biology and genetics • Propedeutics of medical sciences
2.	• Mathematics • Physics • Statistics and probability theory • Principles of electrical engineering • Principles of electronics • Organic chemistry
3.	• Sport • Foreign languages • Physics laboratory • Materials sciences • Principles of metrology • Mechanics and strength of materials • Fundamental anatomy • Principles of physiology • Computer programming
4.	• Sport • Foreign language • Computer Aided Design *-or-* Design with Finite Elements Method • Elements of Biochemistry • Biophysics • Biomaterials • Sensors and non-electrical measurements *-or-* Integrated measurement systems • Digital signal processing

semester	First Degree (continued)
5.	• Foreign languages • Medical Physics Biomechanics • Computer graphics Fundamentals of graphical programming languages • Programming languages *-or-* Object programming • Automatics and Robotics
6.	• Foreign languages • Biomechanics - project • Implants and Artificial Organs • Electronic Medical Instrumentation • Medical Imaging Technology • History of medicine • Elective 1 - Cryptography and data ciphering systems - Chemometry - Ergonomics and occupational medicine • Elective 2 - Principles of management in biotechnical systems - Microcontroller programming in C/C++ - Globalization and modernization problems - Introduction to environmental philosophy - Glass- and glass-ceramic materials in medicine
7.	• Introduction to diagnostics with ionizing radiation • Medical deontology • Introduction to philosophy • Legal and ethic issues in biomedical engineering • Diploma seminar • Engineering project and examination for Bachelor's degree • Elective 3 - Biomineral science - Practical electronics - Programming of control and measurement systems

Table 3. (a) Syllabus of the First Degree (Bachelor's) studies in biomedical engineering

semester	**Second Degree (Master's)**		
	medical electronics and information technologies	biomaterials	biomechanics and robotics
1.	• Identification and modelling of biological structures and processes • Tissue and genetic engineering • Fundamentals of telemedicine • Neural networks • Electronics Systems for Clinical Applications • Information systems in health care • Picture archiving and communication systems • Elective 1 - Design of VLSI circuits - Multimedia systems in medicine - Advanced methods for programming of multithreaded applications	• Clinical trials • Ceramic Biomaterials • Polymer Biomaterials • Identification and modeling of biological structures and processes • Tissue and genetic engineering • Information systems in health care • Implantation technologies • Elective 1 - Electron microscopy in biomedical engineering - Neurochemistry and neuropharmacology	• Biomedical signals processing • Identification and modelling of biological structures and processes • Tissue and genetic engineering • Rehabilitation Technology • Biomechanical designs • Servomechanisms and advanced control systems • Control systems in medical devices • Visual surgery support techniques
2.	• Research of biomaterials and tissues • Dedicated medical diagnostics algorithms • Multimodal interfaces • Rehabilitation Technology • Medical imaging systems • Telesurgery and medical robots • Individual project • Elective 2 - Algorithms for medical image analysis and processing - Cognitive informatics	• Research of biomaterials and tissues • Composite biomaterials • Metallic biomaterials Telesurgery and medical robots • Rehabilitation Technology • Fundamentals of applied crystalography • Individual project	• Research of biomaterials and tissues • Ergonomics • Intelligent materials and structures • Acoustical diagnosis • Information systems in health care • Telesurgery and medical robots • Image processing for surgery support • Individual project
3.	• Voice computer communication • Computer support for acoustic diagnostics • Fundamentals of embedded systems design • Diploma seminar • **Master thesis and examination**	• Fundamentals of regenerative medicine • Diploma seminar • **Master thesis and examination** • Elective 3 - Selected problems of neurobiology - Surface engineering	• Bionics • Pharmaceutical industry equipment • Diploma seminar • **Master thesis and examination** • Elective 3 - Computer aided of engineering - EPLAN - Pharmaceutical industry materials and designs - Nanotechnology

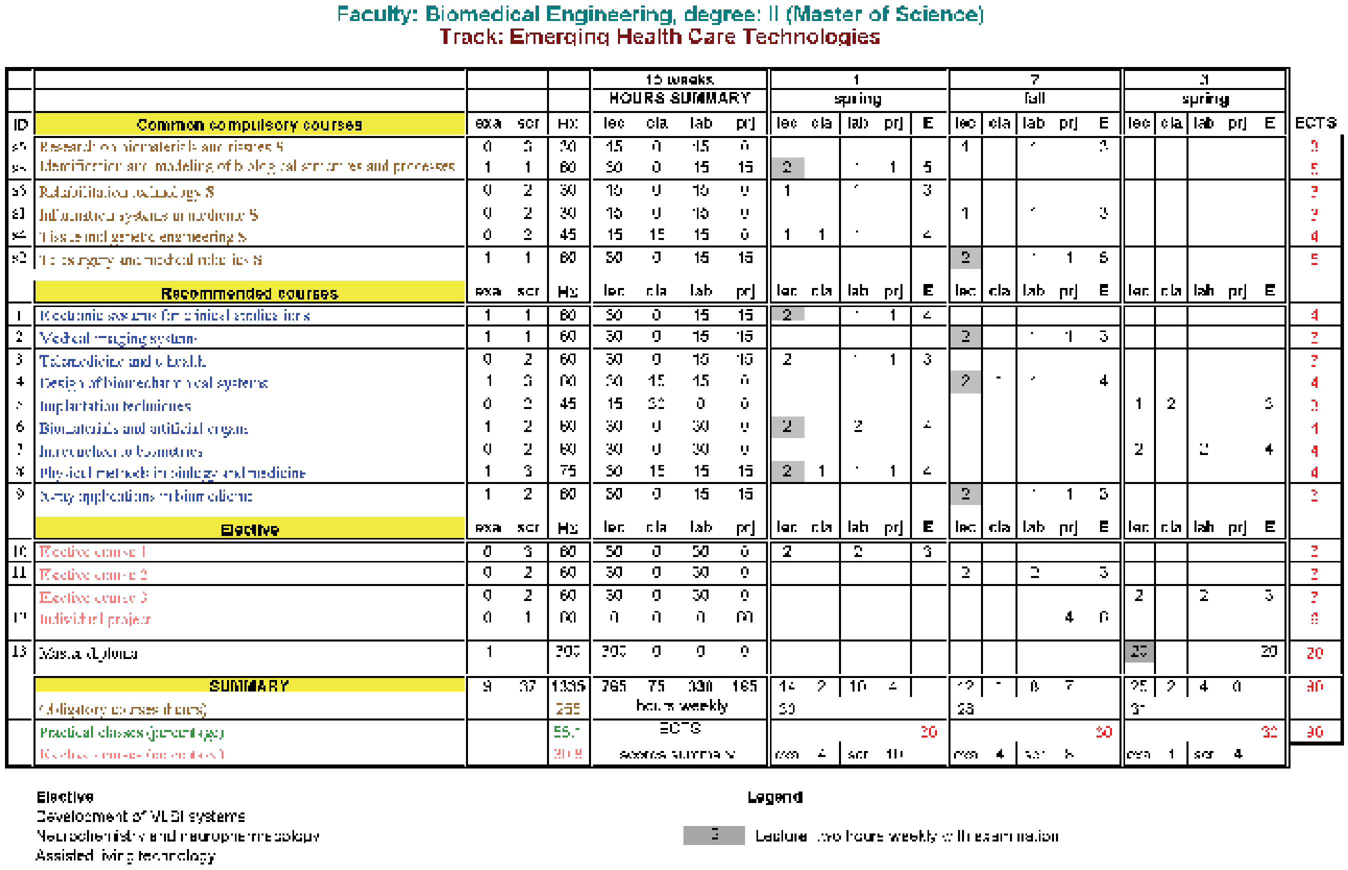
Table 3. (b) Syllabus of the Second Degree (Master's) studies in biomedical engineering.

semester	**Second Degree (Master's)**	
	bionanotechnologies	emerging health care technologies
1.	• Symmetries and structures, solid body and molecules • Soft tissue physics • Polymers • Physics of thin film surfaces • Identification and modeling of biological structures and processes • Tissue and genetic engineering • Physical methods in biology and medicine • Information systems in health care • Elective 1 - Biotechnological challenges in biophysics - Instrumental analysis methods - Fundamentals of cell and tissue engineering - Structural backgrounds of cell biology	• Biomaterials and artificial organs • Electronics Systems for Clinical Applications • Information systems in health care • Physical methods in biology and medicine • Telemedicine and e-health • Tissue and genetic engineering • Assisted Living Technologies
2.	• Research of biomaterials and tissues • Magnetic nanomaterials • X-ray applications in biomedicine • Applications of magnetic resonance in biomedical research • Telesurgery and medical robots • Rehabilitation Technology • Introduction to radiobiology • Individual project	• Design of biomechatronical systems • Identification and modeling of biological structures and processes • Medical imaging systems • Research of biomaterials and tissues • Rehabilitation Technology • Telesurgery and medical robots • X-ray applications in biomedicine • Development of VLSI systems • Individual project
3.	• Optical method of matter investigation • Neuroelectronics • Diploma seminar • **Master thesis and examination** • Elective 3 - General and molecular genetics - Protein engineering - Leukocyte and cancer cells transportation - Fluorescent and confocal - microscopy - Molecular modeling of bioparticles - Photobiology and photomedicine	• Implantation Techniques • Introduction to Biometrics • Neurochemistry and neuropharmacology • Diploma seminar • **Master thesis and examination**

Table 3. (c) Syllabus of the Second Degree (Master's) studies in biomedical engineering.

semester	Third Degree (Doctor's)
1.	• Graph theory • Methods of systems optimization • Biocybernetics • Medical sensors and measurements
2.	• Graph theory • Methods of systems optimization • Medical imaging in clinical practice • Biometry and medical statistics
3.	• Information systems in telemedicine • Biomechanics i acoustics • Biomedical digital signal processing • Biomaterials and artificial organs
4.	• Electronic medical instrumentation • Medical physics
5.	• Medical image analysis • Modeling of biological systems
6.	• Electives (2 of 6) - Dedicated algorithms for biosignal interpretation - Integrated systems SoC in medical diagnostics and therapy - Intelligent sensor arrays - Biophysics - Biological interfaces - Advanced equipment in medicine and rehabilitation
7.	• Philosophy / Economy
8.	• individual research

Table 3. (d) Syllabus of the Third Degree (Doctor's) studies in biomedical engineering.



Faculty: Biomedical Engineering, degree: II (Master of Science)
Track: Emerging Health Care Technologies

					15 weeks				1					2					3					
					HOURS SUMMARY				spring					fall					spring					
ID	**Common compulsory courses**	exa	scr	Hx	lec	cla	lab	prj	lec	cla	lab	prj	E	lec	cla	lab	prj	E	lec	cla	lab	prj	E	ECTS
s5	Research on biomaterials and tissues S	0	3	30	15	0	15	0						1		1		3						3
s6	Identification and modeling of biological structures and processes	1	1	60	30	0	15	15	2		1	1	5											5
s3	Rehabilitation technology S	0	2	30	15	0	15	0	1		1		3											3
s1	Information systems in medicine S	0	2	30	15	0	15	0						1		1		3						3
s4	Tissue and genetic engineering S	0	2	45	15	15	15	0	1	1	1		4											4
s2	Telesurgery and medical robotics S	1	1	60	30	0	15	15						2		1	1	5						5
	Recommended courses	exa	scr	Hx	lec	cla	lab	prj	lec	cla	lab	prj	E	lec	cla	lab	prj	E	lec	cla	lab	prj	E	
1	Electronic systems for clinical applications	1	1	60	30	0	15	15	2		1	1	4											4
2	Medical imaging systems	1	1	60	30	0	15	15						2		1	1	3						3
3	Telemedicine and e-health	0	2	60	30	0	15	15	2		1	1	3											3
4	Design of biomechanical systems	1	3	60	30	15	15	0						2	1	1		4						4
5	Implantation techniques	0	2	45	15	30	0	0											1	2			3	3
6	Biomaterials and artificial organs	1	2	60	30	0	30	0	2		2		4											4
7	Introduction to biometrics	0	2	60	30	0	30	0											2		2		4	4
8	Physical methods in biology and medicine	1	3	75	30	15	15	15	2	1	1	1	4											4
9	X-ray applications in biomedicine	1	2	60	30	0	15	15						2		1	1	3						3
	Elective	exa	scr	Hx	lec	cla	lab	prj	lec	cla	lab	prj	E	lec	cla	lab	prj	E	lec	cla	lab	prj	E	
10	Elective course 1	0	3	60	30	0	30	0	2		2		3											3
11	Elective course 2	0	2	60	30	0	30	0						2		2		3						3
12	Elective course 3	0	2	60	30	0	30	0											2		2		3	3
	Individual project	0	1	60	0	0	0	60									4	6						6
13	Master diploma	1		300	300	0	0	0											20				20	20
	SUMMARY	9	37	1335	765	75	330	165	14	2	10	4		12	1	8	7		25	2	4	0		30
	Obligatory courses (hours)			255	hours weekly				30					23					31					
	Practical classes (percentage)			55.7	ECTS								30					30					32	30
	Elective courses (percentage)			20.9	exams summary				exa	4	scr	10		exa	4	scr	6		exa	1	scr	4		

Elective
Development of VLSI systems
Neurochemistry and neuropharmacology
Assisted living technology

Legend

2	Lecture two hours weekly with examination

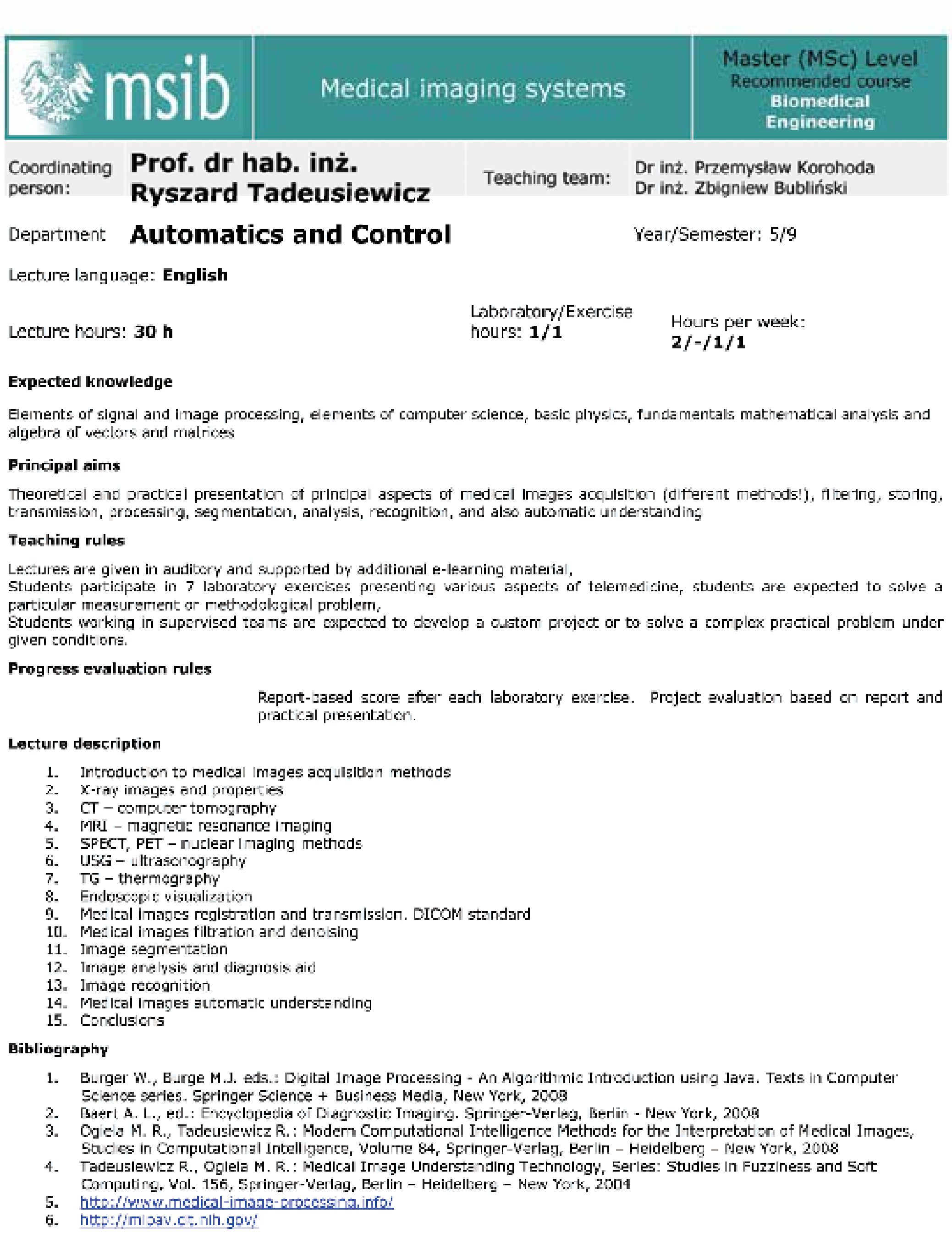
Table 4. (a) Syllabus of the Second Degree (Master's) studies in biomedical engineering, specialization Emerging health care technologies.

Postgraduate BME studies AGH – TRACK TABLE

discipline "biocybernetics and biomedical engineering"

Id	Couse name	Professor	year I		year II		year III		year IV		remarks
		(confidential)	Sem.1	Sem.2	Sem.3	Sem.4	Sem. 5	Sem. 6	Sem. 7	Sem. 8	
Common courses											
1	Graph theory		30	30							
3	Methods of systems optimization		30	30							
5	Philosophy / Economy								30		
Discipline-oriented courses											
6	Biocybernetics		30								
7	Medical imaging in clinical practice			30							
8	Medical sensors and measurements		30								
9	Biometry and medical statistics			30							
3bi1	Information systems in telemedicine				30						4 obligatory courses
3bi2	Biomechanics and acoustics				30						
3bi3	Biomedical digital signal processing				30						
3bi4	Biomaterials and artificial organs				30						
4bi1	Electronic medical instrumentation					30					2 obligatory courses
4bi2	Medical physics					30					
5bi1	Medical image analysis						30				2 obligatory courses
5bi2	Modeling of biological systems						30				
6bi1	Dedicated algorithms for biosignal interpretation										2 obligatory courses to select out of 6 total duration 60 h
6bi2	Integrated systems SoC in medical diagnostics and therapy							30			
6bi3	Intelligent sensor arrays										
6bi4	Biophysics							30			
6bi5	Biological interfaces										
6bi6	Advanced equipment in medicine and rehabilitation										

Table 4. (b) Syllabus of Third Degree (Doctor's) studies in biomedical engineering.



msib

Medical imaging systems

Master (MSc) Level
Recommended course
Biomedical Engineering

Coordinating person:	**Prof. dr hab. inż. Ryszard Tadeusiewicz**	Teaching team:	Dr inż. Przemysław Korohoda Dr inż. Zbigniew Bubliński
Department	**Automatics and Control**		Year/Semester: 5/9

Lecture language: **English**

Lecture hours: **30 h**	Laboratory/Exercise hours: **1/1**	Hours per week: **2/-/1/1**

Expected knowledge

Elements of signal and image processing, elements of computer science, basic physics, fundamentals mathematical analysis and algebra of vectors and matrices

Principal aims

Theoretical and practical presentation of principal aspects of medical images acquisition (different methods!), filtering, storing, transmission, processing, segmentation, analysis, recognition, and also automatic understanding

Teaching rules

Lectures are given in auditory and supported by additional e-learning material,
Students participate in 7 laboratory exercises presenting various aspects of telemedicine, students are expected to solve a particular measurement or methodological problem,
Students working in supervised teams are expected to develop a custom project or to solve a complex practical problem under given conditions.

Progress evaluation rules

Report-based score after each laboratory exercise. Project evaluation based on report and practical presentation.

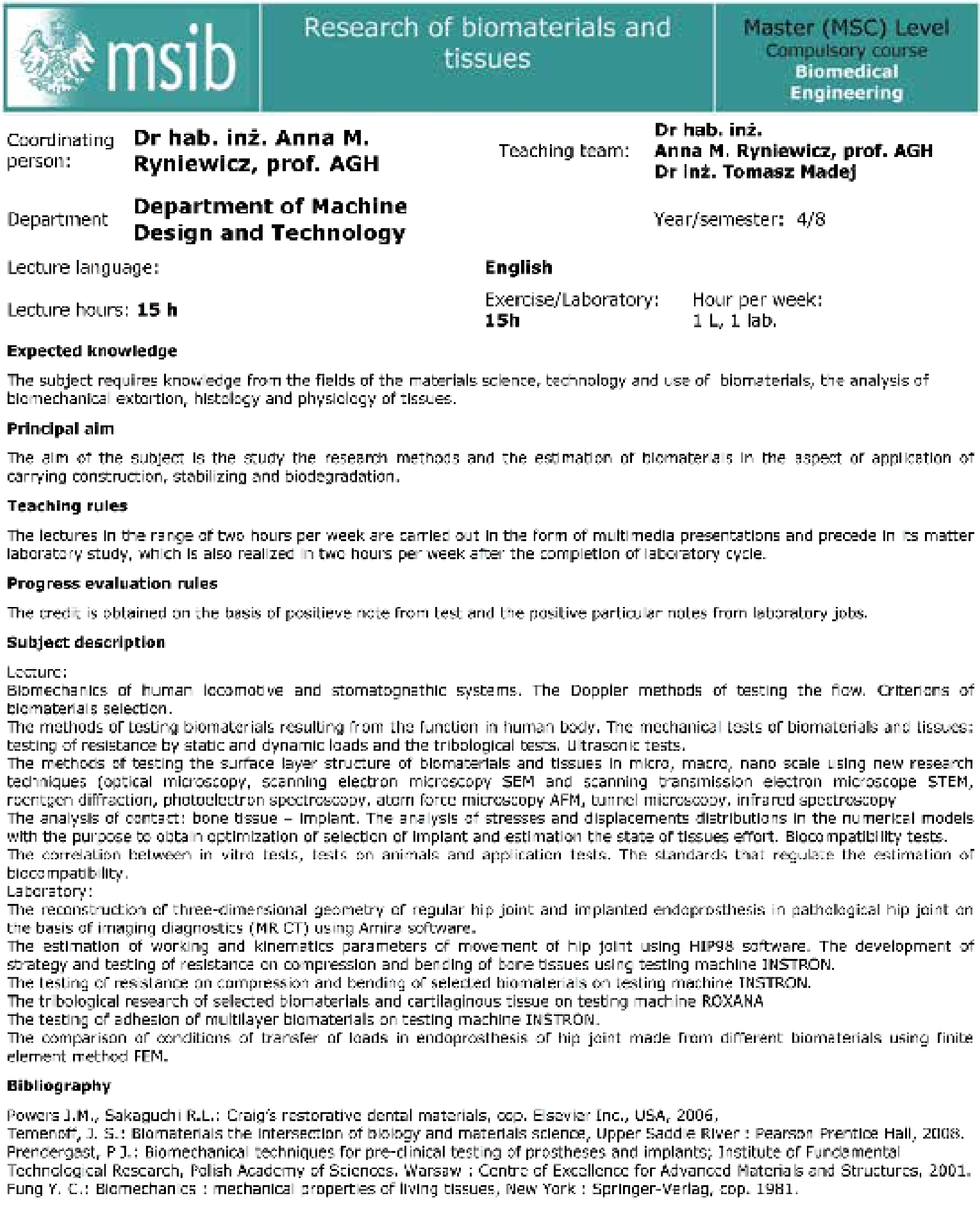
Lecture description

1. Introduction to medical images acquisition methods
2. X-ray images and properties
3. CT – computer tomography
4. MRI – magnetic resonance imaging
5. SPECT, PET – nuclear imaging methods
6. USG – ultrasonography
7. TG – thermography
8. Endoscopic visualization
9. Medical images registration and transmission. DICOM standard
10. Medical images filtration and denoising
11. Image segmentation
12. Image analysis and diagnosis aid
13. Image recognition
14. Medical images automatic understanding
15. Conclusions

Bibliography

1. Burger W., Burge M.J. eds.: Digital Image Processing - An Algorithmic Introduction using Java. Texts in Computer Science series. Springer Science + Business Media, New York, 2008
2. Baert A. L., ed.: Encyclopedia of Diagnostic Imaging. Springer-Verlag, Berlin - New York, 2008
3. Ogiela M. R., Tadeusiewicz R.: Modern Computational Intelligence Methods for the Interpretation of Medical Images, Studies in Computational Intelligence, Volume 84, Springer-Verlag, Berlin – Heidelberg – New York, 2008
4. Tadeusiewicz R., Ogiela M. R.: Medical Image Understanding Technology, Series: Studies in Fuzziness and Soft Computing, Vol. 156, Springer-Verlag, Berlin – Heidelberg – New York, 2004
5. http://www.medical-image-processing.info/
6. http://mipav.cit.nih.gov/

Table 5. (a) Syllabus of the course "Medical imaging systems" (Second Degree-Master's).



msib	Research of biomaterials and tissues	Master (MSC) Level Compulsory course **Biomedical Engineering**

Coordinating person:	**Dr hab. inż. Anna M. Rynlewicz, prof. AGH**	Teaching team:	**Dr hab. inż. Anna M. Ryniewicz, prof. AGH Dr inż. Tomasz Madej**
Department	**Department of Machine Design and Technology**		Year/semester: 4/8
Lecture language:		**English**	
Lecture hours: **15 h**		Exercise/Laboratory: **15h**	Hour per week: 1 L, 1 lab.

Expected knowledge

The subject requires knowledge from the fields of the materials science, technology and use of biomaterials, the analysis of biomechanical extortion, histology and physiology of tissues.

Principal aim

The aim of the subject is the study the research methods and the estimation of biomaterials in the aspect of application of carrying construction, stabilizing and biodegradation.

Teaching rules

The lectures in the range of two hours per week are carried out in the form of multimedia presentations and precede in its matter laboratory study, which is also realized in two hours per week after the completion of laboratory cycle.

Progress evaluation rules

The credit is obtained on the basis of positieve note from test and the positive particular notes from laboratory jobs.

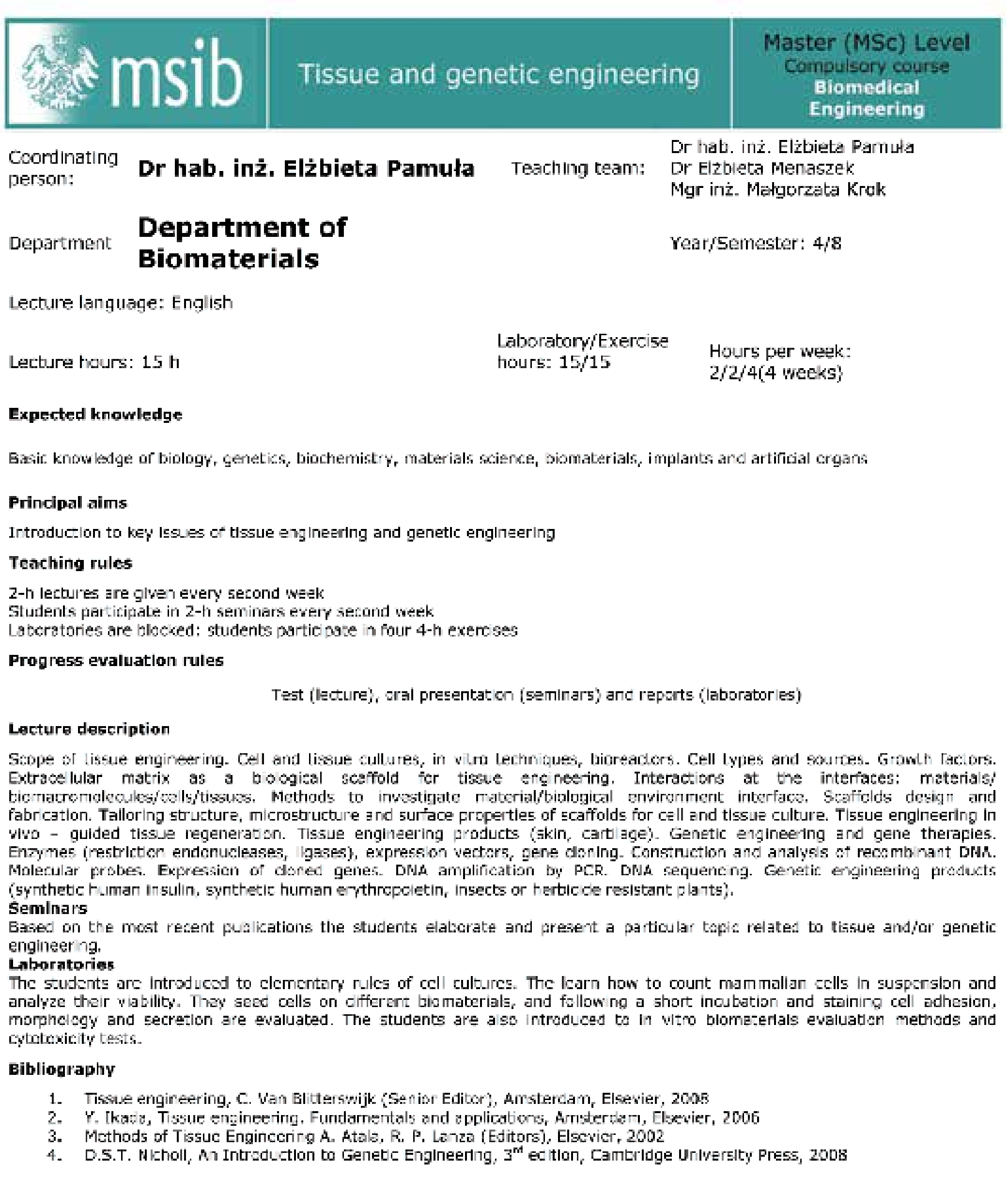
Subject description

Lecture:
Biomechanics of human locomotive and stomatognathic systems. The Doppler methods of testing the flow. Criterions of biomaterials selection.
The methods of testing biomaterials resulting from the function in human body. The mechanical tests of biomaterials and tissues: testing of resistance by static and dynamic loads and the tribological tests. Ultrasonic tests.
The methods of testing the surface layer structure of biomaterials and tissues in micro, macro, nano scale using new research techniques (optical microscopy, scanning electron microscopy SEM and scanning transmission electron microscope STEM, roentgen diffraction, photoelectron spectroscopy, atom force microscopy AFM, tunnel microscopy, infrared spectroscopy
The analysis of contact: bone tissue – implant. The analysis of stresses and displacements distributions in the numerical models with the purpose to obtain optimization of selection of implant and estimation the state of tissues effort. Biocompatibility tests.
The correlation between in vitro tests, tests on animals and application tests. The standards that regulate the estimation of biocompatibility.
Laboratory:
The reconstruction of three-dimensional geometry of regular hip joint and implanted endoprosthesis in pathological hip joint on the basis of imaging diagnostics (MR CT) using Amira software.
The estimation of working and kinematics parameters of movement of hip joint using HIP98 software. The development of strategy and testing of resistance on compression and bending of bone tissues using testing machine INSTRON.
The testing of resistance on compression and bending of selected biomaterials on testing machine INSTRON.
The tribological research of selected biomaterials and cartilaginous tissue on testing machine ROXANA
The testing of adhesion of multilayer biomaterials on testing machine INSTRON.
The comparison of conditions of transfer of loads in endoprosthesis of hip joint made from different biomaterials using finite element method FEM.

Bibliography

Powers J.M., Sakaguchi R.L.: Craig's restorative dental materials, cop. Elsevier Inc., USA, 2006.
Temenoff, J. S.: Biomaterials the intersection of biology and materials science, Upper Saddle River : Pearson Prentice Hall, 2008.
Prendergast, P J.: Biomechanical techniques for pre-clinical testing of prostheses and implants; Institute of Fundamental Technological Research, Polish Academy of Sciences. Warsaw : Centre of Excellence for Advanced Materials and Structures, 2001.
Fung Y. C.: Biomechanics : mechanical properties of living tissues, New York : Springer-Verlag, cop. 1981.

Table 5. (b) Syllabus of the course "Research of biomaterials and tissues" (Second Degree-Master's).



msib

Tissue and genetic engineering

Master (MSc) Level
Compulsory course
Biomedical Engineering

Coordinating person: **Dr hab. inż. Elżbieta Pamuła**

Teaching team:
Dr hab. inż. Elżbieta Pamuła
Dr Elżbieta Menaszek
Mgr inż. Małgorzata Krok

Department **Department of Biomaterials**

Year/Semester: 4/8

Lecture language: English

Lecture hours: 15 h

Laboratory/Exercise hours: 15/15

Hours per week: 2/2/4(4 weeks)

Expected knowledge

Basic knowledge of biology, genetics, biochemistry, materials science, biomaterials, implants and artificial organs

Principal aims

Introduction to key issues of tissue engineering and genetic engineering

Teaching rules

2-h lectures are given every second week
Students participate in 2-h seminars every second week
Laboratories are blocked: students participate in four 4-h exercises

Progress evaluation rules

Test (lecture), oral presentation (seminars) and reports (laboratories)

Lecture description

Scope of tissue engineering. Cell and tissue cultures, in vitro techniques, bioreactors. Cell types and sources. Growth factors. Extracellular matrix as a biological scaffold for tissue engineering. Interactions at the interfaces: materials/biomacromolecules/cells/tissues. Methods to investigate material/biological environment interface. Scaffolds design and fabrication. Tailoring structure, microstructure and surface properties of scaffolds for cell and tissue culture. Tissue engineering in vivo – guided tissue regeneration. Tissue engineering products (skin, cartilage). Genetic engineering and gene therapies. Enzymes (restriction endonucleases, ligases), expression vectors, gene cloning. Construction and analysis of recombinant DNA. Molecular probes. Expression of cloned genes. DNA amplification by PCR. DNA sequencing. Genetic engineering products (synthetic human insulin, synthetic human erythropoietin, insects or herbicide resistant plants).

Seminars

Based on the most recent publications the students elaborate and present a particular topic related to tissue and/or genetic engineering.

Laboratories

The students are introduced to elementary rules of cell cultures. The learn how to count mammalian cells in suspension and analyze their viability. They seed cells on different biomaterials, and following a short incubation and staining cell adhesion, morphology and secretion are evaluated. The students are also introduced to in vitro biomaterials evaluation methods and cytotoxicity tests.

Bibliography

1. Tissue engineering, C. Van Blitterswijk (Senior Editor), Amsterdam, Elsevier, 2008
2. Y. Ikada, Tissue engineering. Fundamentals and applications, Amsterdam, Elsevier, 2006
3. Methods of Tissue Engineering A. Atala, R. P. Lanza (Editors), Elsevier, 2002
4. D.S.T. Nicholl, An Introduction to Genetic Engineering, 3rd edition, Cambridge University Press, 2008

Table 5. (c) Syllabus of the course "Tissue and genetic engineering" (Second Degree-Master's).

8. References

Augustyniak P. (2008) Proceedings of the First National Conference on Biomedical Engineering Education (OKIBEdu). *Acta Bio-optica et Informatica Medica - Biomedical Engineering*. Vol 3?

Augustyniak E, Augustyniak P. (2010) From the Foundation Act to the Corporate Culture of a BME Teaching Institute, *Proceedings of the 32nd annual international conference of the IEEE Engineering in Medicine and Biology Society*, 2010, pp. 319–322

Augustyniak P, Tadeusiewicz R & Wasilewska-Radwańska M. (2010) BME Education Program Following the Expectations from the Industry, Health Care and Science, In: *Medicon 2010, XII Mediterranean Conference on Medical and Biological Engineering and Computing*, Bamidis PD, Pallikarakis N. (Eds.) 2010 Springer, (IFMBE Proceedings vol. 29), pp. 945–948.

Denison D. & Mishra A. (1995) Toward a theory of organisational culture and effectiveness, *Organisation Science*, Vol. 6 No.2, 1995, pp. 204-23.

Gordon G. & DiTomaso N. (1992) Predicting corporate performance from organization culture, *Journal of Management Studies*, Vol. 29 No.6, 1992 pp. 783-98

Ministry of Science and Higher Education (2007) Educational Standards for Higher Education, No 49 Biomedical Engineering (in Polish)

Monzon E. (2005) The Challenges of Biomedical Engineering Education in Latin America. *Proceedings of the 27th annual international conference of the IEEE Engineering in Medicine and Biology Society*, 2005: 2403-2405

O'Reilly C. & Chatman J. (1996) Culture as social control: corporations, cults and commitment, *Research in Organisational Behaviour*, Vol. 18, 1996 pp. 157-200.

ROTMED Consortium (2006) Rapport from foresight examinations entitled: Analysis of the state of Polish medical technologies sector, IBIB PAN, Warsaw 2006 (in Polish)

Schwartz MD. (1988) Biomedical Engineering Education, In: *Encyclopedia of medical Devices and Instrumentation*, Webster JG. (Ed.), Wiley, New York, 1988: 392- 403

Wasilewska-Radwanska M. & Augustyniak P. (2009) Multidisciplinary School as a BME Teaching Option. *IFMBE Proceedings*, Vol. 25, 2009, pp. 200–203

Part 4

Biotechnology

24

Poly (L-glutamic acid)-Paclitaxel Conjugates for Cancer Treatment

Shuang-Qing Zhang
National Center for Safety Evaluation of Drugs,
National Institutes for Food and Drug Control,
China

1. Introduction

One of the effective approaches to develop new anticancer drugs is to prepare polymer-anticancer drug conjugates. The polymer-anticancer drug conjugates include polymer-protein conjugates, polymer-drug conjugates and supramolecular drug-delivery systems. In 1975, the concept of a polymer-drug conjugate was first proposed by Ringsdorf (Ringsdorf, 1975). In his model, a bioactive anticancer agent was attached to a suitable polymeric carrier, directly or through a biodegradable linker. Currently developed delivery systems for anticancer agents encompass colloidal systems (liposomes, emulsions, nanoparticles and micelles), polymer implants and polymer conjugates. These delivery systems are able to provide enhanced therapeutic efficacy and reduce toxicity of anticancer agents mainly by altering the pharmacokinetics and biodistribution of the drugs (Kim & Lim, 2002). The idea is attractive and could form the basis of a new generation of anticancer agents (Sugahara et al., 2007).

Many polymers have been investigated as carriers for conjugates, including poly(glutamic acid), poly(L-lysine), poly(malic acid), poly(aspartamides), poly((N-hydroxyethyl)-L glutamine), poly(ethylene glycol), poly(styrene-co-maleic acid/anhydride), poly(N-(2-hydroxypropyl) methacrylamide) copolymer, poly(ethyleneimine), poly(acroloylmorpholine), poly(vinylpyrrolidone), poly(vinylalcohol), poly(amidoamines), divinylethermaleic anhydride/acid copolymer, dextran, pullulan, mannan, dextrin, chitosan, hyaluronic acid and proteins *etc.*. Coupling low molecular weight anticancer drugs to high or low molecular weight polymers is an effective method for improving the therapeutic index of clinically used agents.

Several candidates have been evaluated in clinical trials, such as N-(2-hydroxypropyl) methacrylamide conjugates of doxorubicin, camptothecin, paclitaxel, and platinum (II) complexes (Haag & Kratz, 2006). The conjugation of cytotoxic agents to a hydrophilic polymer may convey several advantages, (1) increased water solubility; (2) protection from hydrolysis and proteolysis; (3) prolonged half-life and enhanced bioavailability of drug; (4) reduction of toxicity, immunogenicity and antigenicity; (5) controlled release or specific targeting through an enhanced permeability and retention (EPR) effect. In the chapter, we will focus on poly (L-glutamic acid)-paclitaxel (PG-PTX), which can improve the therapeutic index, pharmacokinetic properties, safety and efficacy of paclitaxel (PTX).

2. The anticancer agent PTX

PTX, an anticancer agent isolated from the trunk bark of the Pacific Yew tree, Taxus brevifolia, shows a wide spectrum of anticancer activity for a variety of human cancers, including breast, ovarian, non-small-cell lung, prostate, head and neck, colon cancers and so on (Rowinsky, 1997). PTX can induce mitotic arrest and apoptosis in proliferating cells by targeting tubulin, which is a component of the mitotic spindle. PTX binds to the N-terminal 31 amino acids of the β-tubulin subunit and prevents depolymerization. As a result, the mitotic spindle is disabled, cell division can not be completed, and the cell replication in the late G2 or M phase of the cell cycle is inhibited. The cancer cells are killed by disrupting the dynamics necessary for cell division (Bhalla, 2003). The anticancer mechanism of PTX is shown in Fig. 1.

The clinical use of PTX is limited by its high hydrophobicity, low solubility, high systemic exposure, poor pharmacokinetic characteristics, and the lack of selective tumor uptake (Parveen & Sahoo, 2008; ten Tije et al., 2003). The clinical use of PTX also leads to many side effects. Side effects of PTX include nausea, vomiting, diarrhea, mucositis, myelosuppression, cardiotoxicity, neurotoxicity and hypersensitivity reactions, and the latter two are mainly owing to polyoxyethylated castor oil (Cremophor® EL) and ethanol used for solubilizing PTX (Rogers, 1993; Sugahara et al., 2007). PTX for injection is supplied in 50% Cremophor® EL and 50% dehydrated ethanol.

Despite premedication with corticosteroids and antihistamines, PTX still induces minor reactions (e.g., flushing and rash) in approximately 40% of patients and major potentially life-threatening reactions in 1.5%−3% of patients (Lemieux et al., 2008; Price & Castells, 2002). Hydrophobicity of PTX is also associated with unfavorable kinetics, high levels of protein binding, and high volumes of distribution often greatly exceeding total body water. Together, all these factors have a negative impact on the therapeutic index because only a small proportion of the drug administered actually reaches the tumor site (Singer, 2005). As mentioned above, the efficacy and tolerability of PTX are limited by its low solubility, high systemic exposure, and the lack of selective tumor uptake.

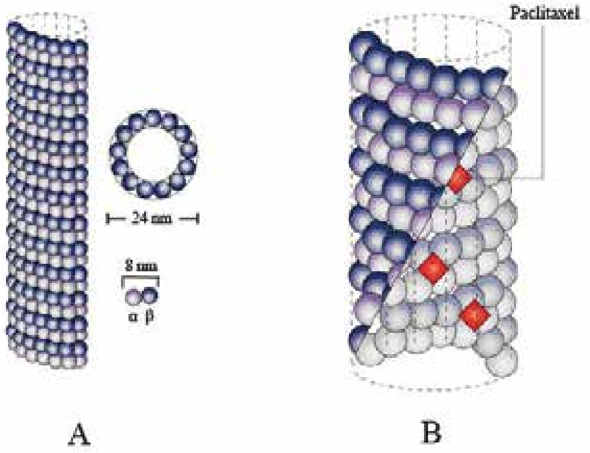


Fig. 1. Schematic illustration of PTX anticancer mode: heterdimers of α- and β-tubulin assemble to form a highly dynamic microtubule which plays an extremely important role in the process of mitosis (A); microtubule-targeted paclitaxel binds along the interior surface of the microtubule, suppressing its dynamics mitosis (B).

3. The EPR effect and endocytosis of polymer-drug conjugates

Anticancer polymer-drug conjugates can be divided into two targeting modalities: passive and active. While clinical anticancer activity has been achieved by passive macromolecular drug delivery systems, further selectivity is possible by active targeting (Luo & Prestwich, 2002). Polymer-drug conjugates can promote passive tumor targeting by EPR effect and allow for lysosomotropic drug delivery following endocytic capture of the drugs (Greco & Vicent, 2008).

The view that the polymer-drug conjugates passively accumulate in tumor tissues because EPR effect is clearly supported by the electron microscopic observation that the peripheral tumor vascular endothelium has quantitatively more fenestrations and open junctions than normal vessels (Li, 2002; Roberts & Palade, 1997). Tumor vasculature is more permeable to macromolecules than normal vasculature because the structures between the neovasculature in tumors and the mature vasculature in normal organs are different (Roberts & Palade, 1997; Singer, 2005). The paucity of lymphatic vessels in tumor tissues allows the retention of these macromolecules in the interstitial space, which leads to 10 to 100-fold increase in intratumoral drug concentrations for a prolonged time when compared with an equivalent dose of the anticancer drug given according to the conventional methods.

The phenomenon of EPR effect is applicable for almost all rapidly growing solid tumors and it has been widely used in cancer-targeting drug design (Iyer et al., 2007; Maeda et al., 2000; Reddy, 2005). The EPR effect is molecular weight (MW) - and size-dependent and is most effective with agents whose MWs are 50 000 or greater, which is above the threshold for renal excretion. Due to the different pathways to enter the cells between the small molecule drugs and the macromolecule drugs, multi-drug resistance (MDR) can be minimized at the same time (Boddy et al., 2005; Greish et al., 2003; Shaffer et al., 2007). Because of the stronger metabolic activity of the cancer cells, in addition to the EPR effect, cancer cells show a higher degree of uptake of macromolecules by endocytosis than normal cells (Li, 2002). The process of EPR and endocytosis is illustrated in Fig. 2.

4. PG-PTX conjugate

4.1 Modified formulations of PTX

Because of the unfavorable properties of PTX, it is urgent to develop a more effective strategy to improve its water solubility and selectivity towards tumor tissues (Maeda et al., 2009). Several approaches have been utilized to increase the therapeutic index of PTX. More water-soluble formulations of PTX have been investigated, including a nanoparticulate formulation (ABI-007), a polymeric micellar formulation (Genexol-PM), and a liposomal formulation and covalent linkage to macromolecule polymers that alter the pharmacokinetics of the parent drug (Ibrahim et al., 2002; Kim et al., 2004; Soepenberg et al., 2004).

4.2 The formation and the anticancer mechanism of PG-PTX

PG-PTX (paclitaxel poliglumex, CT-2103, Xyotax®, Opaxio®) is a water-soluble macromolecular conjugate that links PTX with PG (Li et al., 1996). PTX is conjugated by ester linkage to the γ-carboxylic acid side chains, leading to a relatively stable conjugate (Li et al., 1998a). Because the conjugation site is the 2 hydroxyl group of PTX, which is a crucial site for tubulin binding, the conjugate does not interact with β-tubulin and is inactive (Li, 2002; Rogers, 1993). The median MW of PG-PTX is 38.5 kDa, with a PTX content of approximately 36% on a w/w basis, equivalent to about one PTX ester linkage per 11 PG units of the polymer (Fig. 3) (Bonomi, 2007; Rogers, 1993).

Morphological analysis and biochemical characterizations demonstrate that both PTX and PG-PTX are able to induce apoptosis in cells expressing wild-type p53 or mutant p53, to arrest cells in the G2/M phase of the cell cycle, and to down-regulate HER-2/neu expression. Furthermore, when PG-PTX is compared with other water-soluble derivatives of PTX, including small-molecular-weight sodium pentetic acid-PTX and polyethylene glycol-PTX conjugate (MW 5 kDa), they all show the same effects on telomeric association, mitotic index chromatin condensation, and formation of apoptotic bodies (Multani et al., 1999). These results indicate that PG-PTX has the same mechanisms of action as PTX.

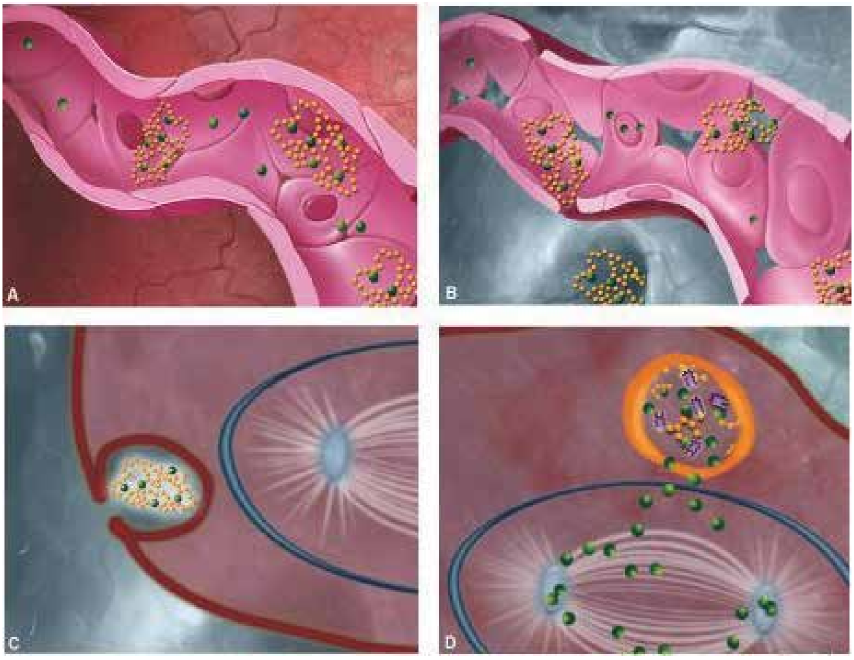
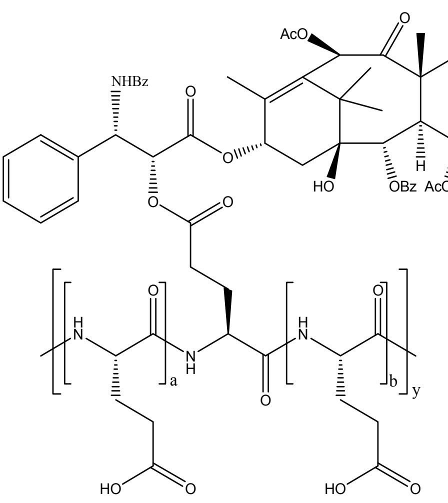


Fig. 2. Illustration of EPR effect and endocytosis. Different from blood vessels in normal tissue (A), those in tumor tissue (B) have porous openings, through which large-size conjugate leaks and is preferentially trapped and distributed to the tumor tissue. Once in the tumor tissue, the conjugate is taken up by the tumor cells through a cellular process called endocytosis (C). The conjugate releases active agent (D) via metabolism by lysosomal enzymes inside the lysosome of the tumor cell.

4.3 PG as the carrier of PG-PTX

Compared with other synthetic polymers that have been tested in clinical studies, PG is unique because it is composed of natural L-glutamic acid linked together through amide bonds rather than the nondegradable C-C backbone. The free γ-carboxyl group in each repeating unit of L-glutamic acid is negatively charged under a neutral pH condition, which makes the polymer water-soluble. The carboxyl groups can also provide functionality for drug attachment. PG is not only water-soluble and biodegradable, it is also nontoxic. All these characteristics make PG a unique candidate as the carrier of polymer-drug conjugates for selective delivery of chemotherapeutic agents, especially for PTX (Parveen & Sahoo, 2008).



a + b averages 10

a and b can = 0

Fig. 3. Schematic representation of PG-PTX structure. The structure shown is illustration of a fragment of the molecule. On average there are approximately 10 non-conjugated monomer glutamic acid units (a + b) for every molecule conjugated to a PTX molecule (y).

4.4 The metabolism of PG-PTX

PG-PTX conjugate is stable in the systemic circulation but can be broken down by intracellular lysosomal enzymes to release the drug after entering cells by endocytosis. The proposed mechanism by which PG-PTX is metabolized includes endocytosis of the polymer-drug conjugate followed by intracellular release of active PTX by proteolytic activity of the lysosomal enzyme cathepsin B, an exocarboxydipeptidase, and diffusion of PTX into the nucleus (Turk et al., 2001). The process is presented in Fig. 4. This finding may have biological relevance as expression of cathepsin B is upregulated in malignant cells, particularly during tumor progression period (Podgorski & Sloane, 2003). These data support a model in which PG-PTX accumulates in tumor tissues through the EPR effect, followed by the cathepsin B-mediated release of PTX. The kinetics of intracellular formation of several PG-PTX metabolites have been quantified in vitro and have been found to be largely dependent on cathepsin B. Metabolites that have been detected in vivo include diglutamyl-PTX and monoglutamyl-PTX. Monoglutamyl-PTX is an unstable compound that can be nonenzymatically degraded to release free PTX. Specific enzyme inhibitors such as CA-074 methyl ester, a cell-permeable irreversible inhibitor of cathepsin B, and EST, a cell-permeable irreversible inhibitor of cysteine proteases, dramatically decrease the formation of monoglutamate PTX and unconjugated PTX in tumor cells that have been incubated with PG-PTX (Rogers, 1993).

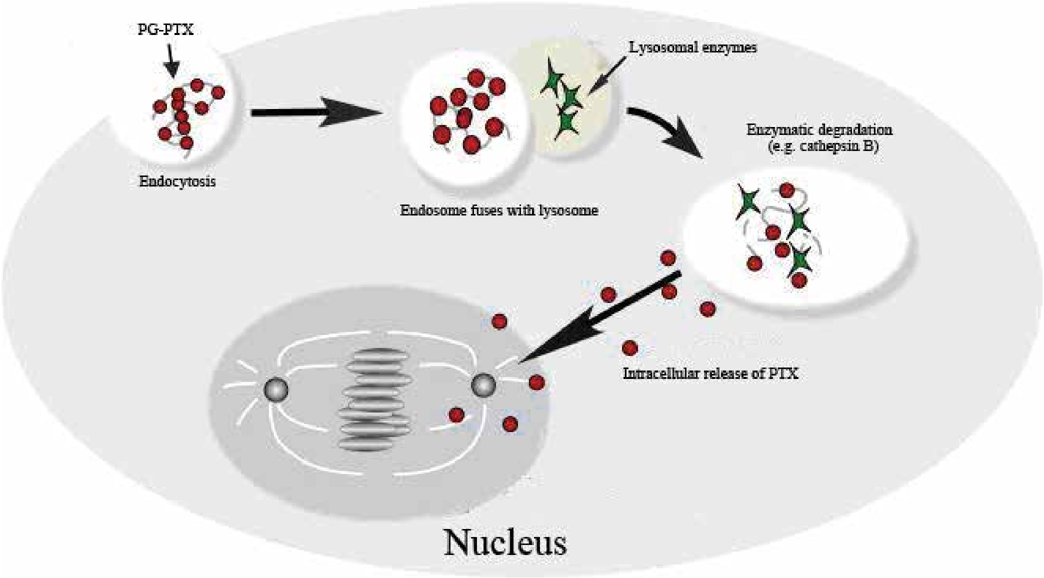


Fig. 4. Metabolism of PG-PTX by endocytosis and enzymatic degradation.

4.5 The advantages of PG-PTX

4.5.1 PG-PTX reduces the side effects of PTX in clinical application

Compared with PTX, the solubility, uptake, tumor retention, and anticancer efficacy of PG-PTX are increased. Many clinical studies so far have confirmed several advantages of PG-PTX in the treatment of cancer patients. This macromolecular conjugate PG-PTX eliminates the need for Cremophor® EL and, therefore, decreases infusion time and the risk of hypersensitivity. Compared to standard taxanes, PG-PTX in phase I and phase II studies shows encouraging outcomes with reduced neutropenia and alopecia, and allows a more convenient administration schedule without the need for routine premedications (Rogers, 1993). PG-PTX induces hypersensitivity reactions in less than 1% of patients, without premedication, and only rare severe reactions have been reported. Furthermore, patients undergoing PG-PTX therapy have a better quality of life, because there is no significant hair loss, nerve damage, or neutropenia at the current dose. PG-PTX is water-soluble and can be administrated rapidly as a 10 -20 min infusion rather than hours when administrating PTX. The recommended phase II dose of PG-PTX is 235 mg/m^2 every 3-week administered over a 10 min infusion without premedication (Sabbatini et al., 2004). Twenty-six patients were treated with PG-PTX in the Phase I study of PG-PTX administered weekly for patients with advanced solid malignancies. The recommended dose of PG-PTX for subsequent disease-directed studies is 70 mg/m^2 weekly. Most patients experienced at least one drug-related adverse event during the study (Table 1) (Mita et al., 2009). Ninety-nine patients were treated with PG-PTX in a multi-center phase II study of PG-PTX as an intravenous (i.v.) infusion (approximately 10 min) at a dose of 175 mg/m^2 on day 1 of each 3-week cycle. And the treatment-related adverse events including non-laboratory-based and laboratory-based maximum CTC toxicities in all cycles are listed in Table 2.

	Grade 1	Grade 2	Grade 3	Total
Adverse event -70 mg/m2				
Neutropenia		1(13%)	1(13%)	2(25%)
Anemia	1(13%)	1(13%)		2(25%)
Peripheral neuropathy	2(25%)	2(25%)		2(50%)
Anorexia	1(13%)			1(13%)
Diarrhea			1(13%)	1(13%)
Nausea	1(13%)			1(13%)
Pain in limb	1(13%)			1(13%)
Fatigue		1(13%)		1(13%)
Adverse event-all cohorts				
Neutropenia		2(8%)	3(12%)	5(19%)
Anemia	1(4%)	1(4%)	2(8%)	4(15%)
Peripheral neuropathy	5(16%)	4(8%)	1(4%)	10(38%)
Anorexia	1(4%)	1(4%)		2(8%)
Diarrhea	3(12%)	1(4%)	1(4%)	5(19%)
Nausea	3(12%)	1(4%)		4(15%)
Rash	1(4%)	2(8%)		3(12%)
Myalgia	3(12%)	2(8%)		5(19%)
Pain in limb	1(4%)	1(4%)		2(8%)
Fatigue	4(15%)	3(12%)		7(27%)
Hypersensitivity	1(4%)	1(4%)		2(8%)

Table 1. Adverse events related to PG-PTX administration (*n* = 26).

	Grade 1	Grade 2	Grade 3	Grade 4
Non-laboratory-based maximum CTC toxicity, all cycles				
Neuropathy	20	15	15	0
Fatigue	10	19	5	0
Musculoskeletal	11	10	1	0
Myalgia/arthralgia	7	5	1	0
Nausea	21	4	1	0
Stomatitis/mucositis	7	1	1	0
Hoarseness	0	0	1	0
Drug hypersensitivity	5	8	1	0
Laboratory-based maximum CTC toxicity, all cycles				
Neutropenia	20	14	15	9
Leukopenia	13	18	18	0
Anemia	46	18	6	0

Abbreviation: CTC, National Cancer Institute Common Toxicity Criteria.

Table 2. Treatment-related adverse events (*n* = 99).

4.5.2 PG-PTX promotes anticancer efficacy and reduces toxicity by prolonging tumor exposure and minimizing systemic exposure to active drug

A single i.v. injection of PG-PTX at its maximum tolerated dose (MTD) equivalent to 60 mg of PTX/kg and at a lower dose equivalent to 40 mg of PTX/kg results in the disappearance of an established implanted 13762F mammary adenocarcinoma (mean size, 2000 mm^3) in rats. Similarly, mice bearing syngeneic OCa-1 ovarian carcinoma (mean size, 500 mm^3) are tumor-free within 2 weeks after a single i.v. injection of PG-PTX at a dose equivalent to 160 mg of PTX/kg (Li et al., 1998b). MTD of PTX in rats and mice are 20 mg/kg and 60 mg/kg, respectively. In contrast, MTD of a single i.v. injection of PG-PTX in rats and mice are 60 mg/kg and 160 mg/kg, respectively. MTD of PG-PTX was approximately 160 mg/kg to 200 mg/kg in immunocompetent mice and 120 mg/kg to 150 mg/kg in immunodeficient animals. At their respective MTDs, single-dose PG-PTX is more efficacious than PTX in Cremophor® EL/ethanol (Li et al., 1999). PG-PTX has shown anticancer activity in preclinical studies with human tumor xenografts and in early phase I trials. MTD of PG-PTX as a single agent, based on the first cycle toxicity of patients, is 235 mg/m^2 (Verschraegen et al., 2009). Biodistribution in mice bearing OCa-1 tumor treated with i.v. injections of tritium-labeled PG-[^{3}H] PTX shows a five times greater distribution of PTX to tumor tissues than those treated with PTX (Li et al., 2000c), which was demonstrated by whole-body autoradiograph (Fig. 5).

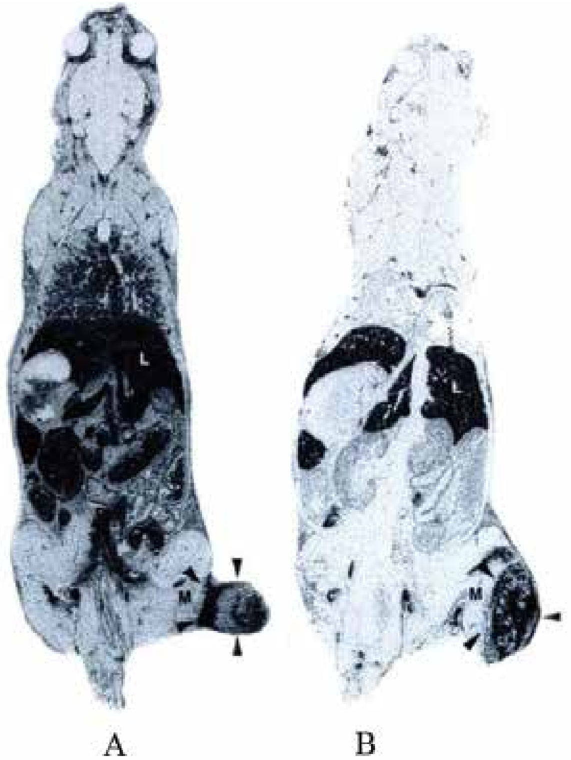


Fig. 5. Whole-body autoradiographs of mice killed 1 day (A) and 6 days (B) after tail vein injection of PG-[^{3}H]PTX. Most radioactivity was localized to tumor periphery at 1 day after injection, but by day 6, radioactivity had diffused into the center of the tumor. L: liver; M: muscle; Arrow head: tumor.

Preclinical studies in animal tumor models demonstrate that PG-PTX is more effective than PTX and it is associated with prolonged tumor exposure but minimized systemic exposure to the active drug. The slow release of the active drug from a well-designed polymer carrier results in sustained high intratumoral drug levels and lower plasma concentrations of the active drug. To accomplish this, the polymer conjugate should release the active drug in tumor tissues rather than in the plasma during circulation. As a result, exposure of normal tissues will be limited, which is potentially associated with a more favorable toxicity profile (Li et al., 2000c). Thus, enhanced tumor uptake and sustained release of PTX from PG-PTX in tumor tissues are major factors contributing to its markedly improved in vivo anticancer activities.

4.5.3 The favorable pharmacokinetic properties of PG-PTX

The superior anticancer activity of PG-PTX in preclinical studies suggests that PG-PTX might have favorable pharmacokinetic properties. Many studies suggest that PG-PTX exerts the anticancer activity by the continuous release of free PTX, and that the favorable pharmacokinetics of PG-PTX conjugate in vivo is likely the main cause contributing to its advanced anticancer activity (Oldham et al., 2000). Female mice with subcutaneous B16 murine melanomas are given PG-[^{3}H] PTX at the equivalent dose 40 mg/kg of [^{3}H] PTX i.v. infusion. Tumor samples are collected at regular intervals up to 144 h after infusion, and the concentrations of PG-PTX and PTX are determined by LC/MS analysis. Tumor exposure to total taxanes is increased by a factor of 3 (C_{max}) or a factor of 12 (AUC) in mice treated with PG-[^{3}H] PTX compared with the mice treated with [^{3}H] PTX (Table 3) (Chipman et al., 2006). PG-[^{3}H] PTX has a much longer half-life in plasma than [^{3}H] PTX (Fig. 6).

Whereas PTX has an extremely short half-life in the plasma of mice ($t_{1/2}$ = 29 min), the apparent half-life of PG-PTX is prolonged ($t_{1/2}$ = 317 min) (Li et al., 1998b). In clinical trials, PG-PTX is given as a 30-min infusion every 3 weeks. Patients were treated at dose levels ranging from 30 mg/m^2 to 720 mg/m^2. PG-PTX is detectable in plasma of all patients and has a long plasma half-life of up to 185 h, and the results are consistent with preclinical findings. Furthermore, concentrations of free PTX released from PG-[^{3}H] PTX remain relatively constant up to 6 d after infusing. Moreover, peak plasma concentrations of free PTX are less than 0.1 μM 24 h after PG-PTX administration at doses up to 480 mg/m^2 (176 mg/m^2 PTX equivalents) (Boddy et al., 2001).

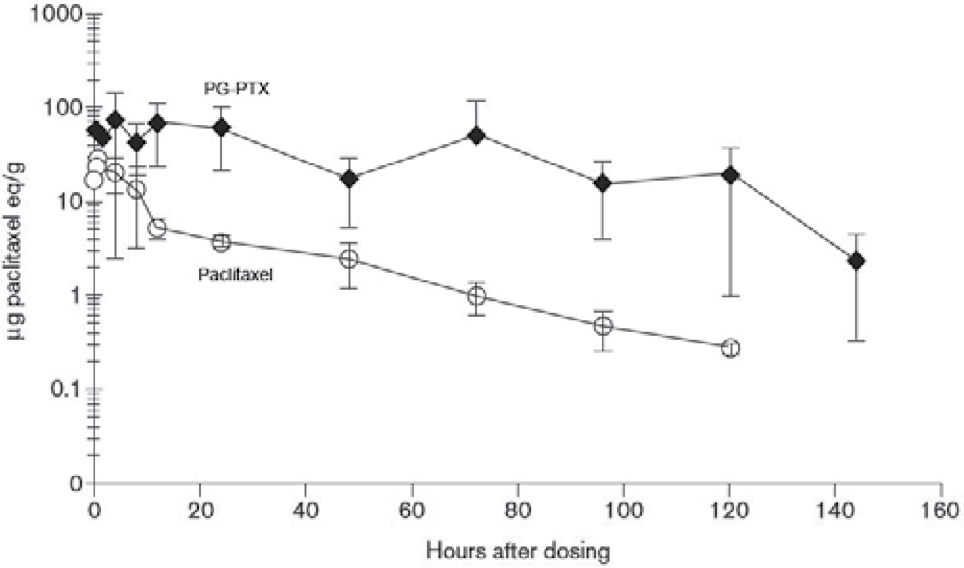


Fig. 6. Tumor concentration of [^{3}H]paclitaxel after treatment with PG-[^{3}H] PTX and [^{3}H] paclitaxel in female mice with s.c. B16 melanomas at a dose of 40 mg paclitaxel.

In another phase I study, PG-PTX is administrated at 70 mg/m^2. The mean maximal concentration (C_{max}) is 41.2 ± 8.60 µg/mL and the C_{max} is reached right after the end of the infusion. The plasma concentration declines with a mean terminal half-life of 15.7 ± 3.17 h. The mean AUC at the MTD is 455 ± 112 µg/h/mL and the mean average systemic plasma clearance is 0.16 ± 0.04 L/h/m^2 (Sabbatini et al., 2004). At the MTD, the mean volume of distribution at steady state and during the terminal phase are 1.41 ± 0.28 L/m^2 and 3.62 ± 1.13 L/m^2, the mean C_{max} of unconjugated PTX is 0.21 ± 0.07 µg/mL, the mean T_{max} is 0.56 ± 0.18 h, the mean terminal half life is 16.6 ± 7.85 h, and the mean AUC is 3.15 ± 1.16 µg/h/mL. The ratio of the free PTX AUC to the conjugated PTX AUC is 0.7%. The plasma concentration of PTX released from PG-PTX increases largely in proportion to the dose and remains similar after repeated administration (Sabbatini et al., 2004).

	C_{max} (µg/g)	T_{max} (h)	AUC (µg/h/g)	MRT (h)
PG-[^{3}H] PTX				
Total taxanes	72.0	4	4547	51
PTX	4.0	72	345	66
[^{3}H] PTX				
Total taxanes	26.7	1.5	384	23
PTX	22.4	1.5	261	17

Table 3. Preclinical tumor pharmacokinetics.

4.5.4 PG-PTX facilitates the radiotherapy and chemotherapy

Preclinical studies in animal tumor models demonstrate the enhanced safety and efficacy of PG-PTX relative to PTX when administered as a single agent or in conjunction with radiation. Studies show that PG-PTX given 24 h before or after radiotherapy enhances tumor growth delay significantly more than PTX. PG-PTX dramatically potentiates tumor radiocurability after single-dose or fractionated irradiation without affecting acute normal tissue injury. PG-PTX increases the therapeutic ratio of radiotherapy more than that previously reported for other taxanes (Milas et al., 2003). PG-PTX not only can produce a much stronger radiopotentiating effect than PTX, the kinetics of its radiopotentiating effect is also different from that of PTX. Delays in the growth of syngeneic murine ovarian OCa-1 tumors grown intramuscularly in C3Hf/Kam mice are used as the treatment end point.

PG-PTX given 24 h before tumor irradiation increases the efficacy of tumor radiation by a factor of more than 4 (Li et al., 2000a). Furthermore, the combination of radiation and PG-PTX can produce a significantly greater tumor growth delay than treatment with radiation and PTX when both drugs are given at the same equivalent PTX dose of 60 mg/kg 24 h after tumor irradiation (enhancement factors, 4.44 versus 1.50) (Li et al., 2000b). When the treatment end point is tumor cure, the enhancement factors are 8.4 and 7.2 of fractionated and single dose radiation, respectively. These values are greater than those produced by other taxanes or by any other chemotherapeutic drugs or radiosensitizer tested so far. PG-PTX may exert its radiopotentiation activity through increased tumor uptake of PG-PTX and sustained release of PTX in the tumor (Li et al., 2000a). To determine whether prior irradiation affects tumor uptake of PG-PTX, PG-[^{3}H] PTX is injected into mice with OCa-1 tumors 24 h after 15 Gy local irradiation (Li et al., 2000b). The uptake of PG-[^{3}H] PTX in irradiated tumors is 28%−38% higher than that in nonirradiated tumors at different times after PG-[^{3}H] PTX injection, indicating that tumor irradiation can increase the accumulation of PG-PTX in the tumors (Fig. 7).

Anticancer activity in patients who have failed previous chemotherapy, including PTX treatment, is observed with PG-PTX (Sabbatini et al., 2004). Ninety-nine patients in a multi-center phase II study treated with PG-PTX as an i.v. infusion approximately 10 min at a dose of 175 mg/m^2 on day 1 of each 3-week cycle have received at least one cycle of treatment. Response rates categorized by platinum sensitivity are shown in Table 4.

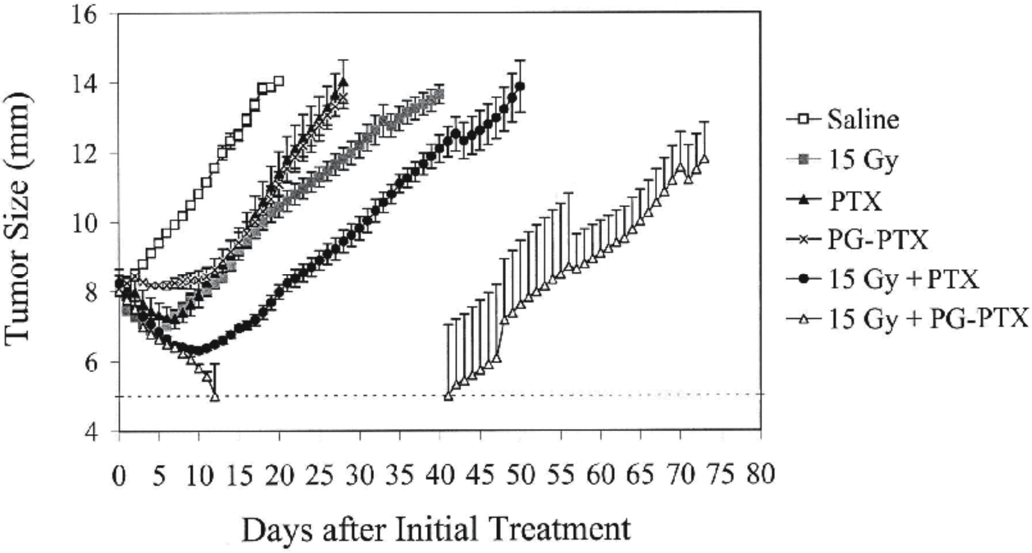


Fig. 7. Effects of combined radiation and PG-PTX and radiation and PTX on the growth of OCa-1 tumors in mice.

No. of Regimens	Response					
	PR		SD		PD	
	No. of Patients	(%)	No. of Patients	(%)	No. of Patients	(%)
Platinum-sensitive, $n = 42$						
1 or 2 prior regimens	5/18	28	6/18	33	7/18	39
≥3 prior regimens	1/24	4	11/24	46	12/24	50
Total	6/42	14	17/42	40	19/42	45
Platinum-resistant/refractory, $n = 57$						
1 or 2 prior regimens	2/21	10	4/21	19	15/21	71
≥3 prior regimens	2/36	6	11/36	31	23/36	64
Total	4/57	7	15/57	26	38/57	67

Abbreviations: PR, partial response; SD, stable disease; PD, progressive disease.

Table 4. Response by Platinum Sensitivity.

5. Summary

Clinical proof of concept for polymer conjugates has already been achieved over the last three decades, with a family of polymer-protein conjugates reaching the market and a growing list of polymer-drug conjugates currently in clinical studies. The application of polymer-anticancer drug conjugates in anticancer treatment is a promising field with growing opportunities to achieve medical treatments with highly improved therapeutic value (Vicent et al., 2008). PG-PTX conjugate can improve the anticancer activity, enhance the safety and efficacy, ameliorate the pharmacokinetic properties and so on. Therefore, the application of PG-PTX facilitates the clinical therapy of a variety of human cancers.

However, many challenges still exist, providing opportunities to improve this platform technology further. The clinical development of anticancer agents utilizing various delivery systems is actively undergoing. New technologies and multidisciplinary approaches to develop advanced drug delivery systems, applicable to a wide range of anticancer agents, may eventually lead to an effective cancer therapy in the future.

6. Acknowledgement

This work was supported by the National Key Technologies R & D Program for Drug Safety Evaluation (Grant No. 2008ZX09305-002) and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

7. References

Bhalla, K. N. (2003). Microtubule-targeted anticancer agents and apoptosis. *Oncogene*, Vol. 22, No. 56, (December 2003), pp. 9075-9086, ISSN 0950-9232

Boddy, A. V.; Griffin M. J.; Sludden J.; Thomas H. D.; Fishwick K.; Wright J. G.; Plumner E. R.; Highley M. & Calvert A. H. (2001). Pharmacological study of paclitaxel duration of infusion combined with GFR-based carboplatin in the treatment of ovarian cancer. *Cancer chemotherapy and pharmacology*, Vol. 48, No. 1, (July 2001), pp. 15-21, ISSN 0344-5704

Boddy, A. V.; Plummer E. R.; Todd R.; Sludden J.; Griffin M.; Robson L.; Cassidy J.; Bissett D.; Bernareggi A.; Verrill M. W. and others. (2005). A phase I and pharmacokinetic study of paclitaxel poliglumex (XYOTAX), investigating both 3-weekly and 2-weekly schedules. *Clinical Cancer Research*, Vol. 11, No. 21, (November 2005), pp. 7834-7840, ISSN 1078-0432

Bonomi, P. (2007). Paclitaxel poliglumex (PPX, CT-2103): macromolecular medicine for advanced non-small-cell lung cancer. *Expert Review of Anticancer Therapy*, Vol. 7, No. 4, (April 2007), pp. 415-422, ISSN 1744-8328

Chipman, S. D.; Oldham F. B.; Pezzoni G. & Singer J. W. (2006). Biological and clinical characterization of paclitaxel poliglumex (PPX, CT-2103), a macromolecular polymer-drug conjugate. *International Journal of Nanomedicine*, Vol. 1, No. 4, 2006), pp. 375-383, ISSN 1176-9114

Greco, F. & Vicent M. J. (2008). Polymer-drug conjugates: current status and future trends. *Frontiers in Bioscience*, Vol. 13, No., 2008), pp. 2744-2756, ISSN 1093-4715

Greish, K.; Fang J.; Inutsuka T.; Nagamitsu A. & Maeda H. (2003). Macromolecular therapeutics: advantages and prospects with special emphasis on solid tumour targeting. *Clinical pharmacokinetics*, Vol. 42, No. 13, 2003), pp. 1089-1105, ISSN 0312-5963

Haag, R. & Kratz F. (2006). Polymer therapeutics: concepts and applications. *Angewandte Chemie*, Vol. 45, No. 8, (February 2006), pp. 1198-1215, ISSN 1433-7851

Ibrahim, N. K.; Desai N.; Legha S.; Soon-Shiong P.; Theriault R. L.; Rivera E.; Esmaeli B.; Ring S. E.; Bedikian A.; Hortobagyi G. N. and others. (2002). Phase I and pharmacokinetic study of ABI-007, a Cremophor-free, protein-stabilized, nanoparticle formulation of paclitaxel. *Clinical Cancer Research*, Vol. 8, No. 5, (May 2002), pp. 1038-1044, ISSN 1078-0432

Iyer, A. K.; Greish K.; Seki T.; Okazaki S.; Fang J.; Takeshita K. & Maeda H. (2007). Polymeric micelles of zinc protoporphyrin for tumor targeted delivery based on EPR effect and singlet oxygen generation. *Journal of drug targeting*, Vol. 15, No. 7-8, (August-September 2007), pp. 496-506, ISSN 1061-186X

Kim, C. K. & Lim S. J. (2002). Recent progress in drug delivery systems for anticancer agents. *Archives of pharmacal research*, Vol. 25, No. 3, (June 2002), pp. 229-239, ISSN 0253-6269

Kim, T. Y.; Kim D. W.; Chung J. Y.; Shin S. G.; Kim S. C.; Heo D. S.; Kim N. K. & Bang Y. J. (2004). Phase I and pharmacokinetic study of Genexol-PM, a cremophor-free, polymeric micelle-formulated paclitaxel, in patients with advanced malignancies. *Clinical Cancer Research*, Vol. 10, No. 11, (June 2004), pp. 3708-3716, ISSN 1078-0432

Lemieux, J.; Maunsell E. & Provencher L. (2008). Chemotherapy-induced alopecia and effects on quality of life among women with breast cancer: a literature review. *Psychooncology*, Vol. 17, No. 4, (April 2008), pp. 317-328, ISSN 1099-1611

Li, C. (2002). Poly(L-glutamic acid)--anticancer drug conjugates. *Advanced Drug Delivery Reviews*, Vol. 54, No. 5, (September 2002), pp. 695-713, ISSN 0169-409X

Li, C.; Ke S.; Wu Q. P.; Tansey W.; Hunter N.; Buchmiller L. M.; Milas L.; Charnsangavej C. & Wallace S. (2000a). Potentiation of ovarian OCa-1 tumor radioresponse by poly (L-glutamic acid)-paclitaxel conjugate. *International Journal of Radiation Oncology, Biology, Physics*, Vol. 48, No. 4, (November 2000a), pp. 1119-1126, ISSN 0360-3016

Li, C.; Ke S.; Wu Q. P.; Tansey W.; Hunter N.; Buchmiller L. M.; Milas L.; Charnsangavej C. & Wallace S. (2000b). Tumor irradiation enhances the tumor-specific distribution of poly(L-glutamic acid)-conjugated paclitaxel and its antitumor efficacy. *Clinical Cancer Research*, Vol. 6, No. 7, (July 2000b), pp. 2829-2834, ISSN 1078-0432

Li, C.; Newman R. A.; Wu Q. P.; Ke S.; Chen W.; Hutto T.; Kan Z.; Brannan M. D.; Charnsangavej C. & Wallace S. (2000c). Biodistribution of paclitaxel and poly(L-glutamic acid)-paclitaxel conjugate in mice with ovarian OCa-1 tumor. *Cancer chemotherapy and pharmacology*, Vol. 46, No. 5, 2000c), pp. 416-422, ISSN 0344-5704

Li, C.; Price J. E.; Milas L.; Hunter N. R.; Ke S.; Yu D. F.; Charnsangavej C. & Wallace S. (1999). Antitumor activity of poly(L-glutamic acid)-paclitaxel on syngeneic and xenografted tumors. *Clinical Cancer Research*, Vol. 5, No. 4, (April 1999), pp. 891-897, ISSN 1078-0432

Li, C.; Yu D.; Inoue T.; Yang D. J.; Milas L.; Hunter N. R.; Kim E. E. & Wallace S. (1996). Synthesis and evaluation of water-soluble polyethylene glycol-paclitaxel conjugate as a paclitaxel prodrug. *Anticancer Drugs*, Vol. 7, No. 6, (August 1996), pp. 642-648, ISSN 0959-4973

Li, C.; Yu D. F.; Newman R. A.; Cabral F.; Stephens L. C.; Hunter N.; Milas L. & Wallace S. (1998a). Complete regression of well-established tumors using a novel water-soluble poly(L-glutamic acid)-paclitaxel conjugate. *Cancer Res*, Vol. 58, No. 11, (Jun 1998a), pp. 2404-2409, ISSN 0008-5472

Li, C.; Yu D. F.; Newman R. A.; Cabral F.; Stephens L. C.; Hunter N.; Milas L. & Wallace S. (1998b). Complete regression of well-established tumors using a novel water-soluble poly(L-glutamic acid)-paclitaxel conjugate. *Cancer Research*, Vol. 58, No. 11, (June 1998b), pp. 2404-2409, ISSN 0008-5472

Luo, Y. & Prestwich G. D. (2002). Cancer-targeted polymeric drugs. *Current Cancer Drug Targets*, Vol. 2, No. 3, (September 2002), pp. 209-226, ISSN 1568-0096

Maeda, H.; Bharate G. Y. & Daruwalla J. (2009). Polymeric drugs for efficient tumor-targeted drug delivery based on EPR-effect. *European Journal of Pharmaceutics and Biopharmaceutics:*, Vol. 71, No. 3, (March 2009), pp. 409-419, ISSN 1873-3441

Maeda, H.; Wu J.; Sawa T.; Matsumura Y. & Hori K. (2000). Tumor vascular permeability and the EPR effect in macromolecular therapeutics: a review. *Journal of Controlled Release*, Vol. 65, No. 1-2, (March 2000), pp. 271-284, ISSN 0168-3659

Milas, L.; Mason K. A.; Hunter N.; Li C. & Wallace S. (2003). Poly(L-glutamic acid)-paclitaxel conjugate is a potent enhancer of tumor radiocurability. *International Journal of Radiation Oncology, Biology, Physics*, Vol. 55, No. 3, (March 2003), pp. 707-712, ISSN 0360-3016

Mita, M.; Mita A.; Sarantopoulos J.; Takimoto C. H.; Rowinsky E. K.; Romero O.; Angiuli P.; Allievi C.; Eisenfeld A. & Verschraegen C. F. (2009). Phase I study of paclitaxel poliglumex administered weekly for patients with advanced solid malignancies. *Cancer chemotherapy and pharmacology*, Vol. 64, No. 2, (July 2009), pp. 287-295, ISSN 1432-0843

Multani, A. S.; Li C.; Ozen M.; Imam A. S.; Wallace S. & Pathak S. (1999). Cell-killing by paclitaxel in a metastatic murine melanoma cell line is mediated by extensive telomere erosion with no decrease in telomerase activity. *Oncology Reports*, Vol. 6, No. 1, (January-February 1999), pp. 39-44, ISSN 1021-335X

Oldham, E. A.; Li C.; Ke S.; Wallace S. & Huang P. (2000). Comparison of action of paclitaxel and poly(L-glutamic acid)-paclitaxel conjugate in human breast cancer cells. *International Journal of Oncology*, Vol. 16, No. 1, (January 2000), pp. 125-132, ISSN 1019-6439

Parveen, S. & Sahoo S. K. (2008). Polymeric nanoparticles for cancer therapy. *Journal of drug targeting*, Vol. 16, No. 2, (February 2008), pp. 108-123, ISSN 1061-186X

Podgorski, I. & Sloane B. F. (2003). Cathepsin B and its role(s) in cancer progression. *Biochemical Society Symposium*, Vol. No. 70, 2003), pp. 263-276, ISSN 0067-8694

Price, K. S. & Castells M. C. (2002). Taxol reactions. *Allergy and Asthma Proceedings*, Vol. 23, No. 3, (May-June 2002), pp. 205-208, ISSN 1088-5412

Reddy, L. H. (2005). Drug delivery to tumours: recent strategies. *Journal of Pharmacy and Pharmacology*, Vol. 57, No. 10, (October 2005), pp. 1231-1242, ISSN 0022-3573

Ringsdorf, H. (1975). Structure and properties of pharmacologically active polymers. *Journal of Polymer Science: Polymer Symposia*, Vol. 51, No. 1, 1975), pp. 135-153,

Roberts, W. G. & Palade G. E. (1997). Neovasculature induced by vascular endothelial growth factor is fenestrated. *Cancer Research*, Vol. 57, No. 4, (February 1997), pp. 765-772, ISSN 0008-5472

Rogers, B. B. (1993). Taxol: a promising new drug of the '90s. *Oncology nursing forum*, Vol. 20, No. 10, (November-December 1993), pp. 1483-1489, ISSN 0190-535X

Rowinsky, E. K. (1997). The development and clinical utility of the taxane class of antimicrotubule chemotherapy agents. *Annual review of medicine*, Vol. 48, No., 1997), pp. 353-374, ISSN 0066-4219

Sabbatini, P.; Aghajanian C.; Dizon D.; Anderson S.; Dupont J.; Brown J. V.; Peters W. A.; Jacobs A.; Mehdi A.; Rivkin S. and others. (2004). Phase II study of CT-2103 in patients with recurrent epithelial ovarian, fallopian tube, or primary peritoneal carcinoma. *Journal of Clinical Oncology*, Vol. 22, No. 22, (November 2004), pp. 4523-4531, ISSN 0732-183X

Shaffer, S. A.; Baker-Lee C.; Kennedy J.; Lai M. S.; de Vries P.; Buhler K. & Singer J. W. (2007). In vitro and in vivo metabolism of paclitaxel poliglumex: identification of metabolites and active proteases. *Cancer chemotherapy and pharmacology*, Vol. 59, No. 4, (March 2007), pp. 537-548, ISSN 0344-5704

Singer, J. W. (2005). Paclitaxel poliglumex (XYOTAX, CT-2103): a macromolecular taxane. *Journal of Controlled Release*, Vol. 109, No. 1-3, (December 2005), pp. 120-126, ISSN 0168-3659

Soepenberg, O.; Sparreboom A.; de Jonge M. J.; Planting A. S.; de Heus G.; Loos W. J.; Hartman C. M.; Bowden C. & Verweij J. (2004). Real-time pharmacokinetics guiding clinical decisions; phase I study of a weekly schedule of liposome encapsulated paclitaxel in patients with solid tumours. *European Journal of Cancer*, Vol. 40, No. 5, (March 2004), pp. 681-688, ISSN 0959-8049

Sugahara, S.; Kajiki M.; Kuriyama H. & Kobayashi T. R. (2007). Complete regression of xenografted human carcinomas by a paclitaxel-carboxymethyl dextran conjugate (AZ10992). *Journal of Controlled Release*, Vol. 117, No. 1, (January 2007), pp. 40-50, ISSN 0168-3659

ten Tije, A. J.; Verweij J.; Loos W. J. & Sparreboom A. (2003). Pharmacological effects of formulation vehicles: implications for cancer chemotherapy. *Clinical pharmacokinetics*, Vol. 42, No. 7, 2003), pp. 665-685, ISSN 0312-5963

Turk, V.; Turk B. & Turk D. (2001). Lysosomal cysteine proteases: facts and opportunities. *EMBO Journal*, Vol. 20, No. 17, (September 2001), pp. 4629-4633, ISSN 0261-4189

Verschraegen, C. F.; Skubitz K.; Daud A.; Kudelka A. P.; Rabinowitz I.; Allievi C.; Eisenfeld A.; Singer J. W. & Oldham F. B. (2009). A phase I and pharmacokinetic study of paclitaxel poliglumex and cisplatin in patients with advanced solid tumors. *Cancer* *chemotherapy and pharmacology*, Vol. 63, No. 5, (April 2009), pp. 903-910, ISSN 1432-0843

Vicent, M. J.; Dieudonne L.; Carbajo R. J. & Pineda-Lucena A. (2008). Polymer conjugates as therapeutics: future trends, challenges and opportunities. *Expert Opinion on Drug Delivery*, Vol. 5, No. 5, (May 2008), pp. 593-614, ISSN 1742-5247

25

Hydrophobic Interaction Chromatography: Fundamentals and Applications in Biomedical Engineering

Andrea Mahn
Universidad de Santiago de Chile,
Chile

1. Introduction

Hydrophobic interaction chromatography (HIC) a powerful technique used for separation and purification of biomolecules. It was described for the first time by Shepard & Tiselius (1949), using the term "salting-out chromatography". Later, Shaltiel & Er-el (1973) introduced the term "hydrophobic chromatography". Finally, Hjerten (1973) described this technique as "hydrophobic interaction chromatography", based on the retention of proteins on weakly hydrophobic matrices in presence of salt. Owing of its high versatility and efficiency, HIC is widely used for the separation and purification of proteins in their native state (Porath et al., 1973), as well as for isolating protein complexes (Chaturvedi et al., 2000) and studying protein folding and unfolding (Bai et al., 1997).

HIC has been applied in separating homologous proteins (Fausnaugh & Regnier, 1986), receptors (Zhang et al., 2008), antibodies (Kostareva et al., 2008), recombinant proteins (Lienqueo et al., 2003) and nucleic acids (Savard & Schneider, 2007). HIC shows similar capacity to ion exchange chromatography (IEC) and a high level of resolution. Since it exploits a different principle than IEC and other separation techniques it can be used as an orthogonal method to achieve the purification of complex protein mixtures (Haimer et al., 2007). In this chapter, the theoretical principles underlying macromolecule retention in HIC are reviewed and discussed in sight of their application for predicting macromolecule behavior in HIC. Besides, novel applications of HIC are discussed regarding their suitability on Biomedical Engineering.

2. Theoretical principles underlying macromolecule retention in Hydrophobic Interaction Chromatography

2.1 Thermodynamics fundamentals

Hydrophobicity can be defined, in general terms, as the repulsion between a non-polar molecule and a polar environment, such as that conferred by water, methanol, and other polar solvents. Two hydrophobic molecules (non-polar) located in a polar environment will show a trend to minimize the contact with the polar solvent. This is accomplished by coming in contact with each other thus minimizing the molecular surface exposed to the solvent. This phenomenon is known as "hydrophobic interaction". Hydrophobic interaction is the most common macromolecular interaction in biological systems. It is also the driving force of several biological and physicochemical processes, such as protein folding, antigen-antibody recognition, stabilization of enzyme-substrate complexes, among others.

From a thermodynamic point of view, the interaction between hydrophobic molecules is an entropy-driven process, based on the second law of Thermodynamics and considering that temperature (T) and pressure (P) remain constant during the process, in this case, the hydrophobic interaction between two biological molecules. Considering equation (1), when a non-polar molecule enters in contact with a polar solvent (usually water), an increase in the degree of order of the solvent molecules that surround the hydrophobic molecule is observed, producing a decrease in entropy ($\Delta S < 0$). Given that enthalpy (ΔH) does not suffer a significant increase in this kind of processes (constant temperature) in comparison with $T\Delta S$, an overall positive change in the Gibbs energy ($\Delta G > 0$) is produced. Hence, the dissolution of a non-polar molecule in a polar solvent does not occur spontaneously, since it is thermodynamically unfavorable.

$$\Delta G = \Delta H - T\Delta S \tag{1}$$

The thermodynamics situation changes when two or more non-polar molecules are located in a polar environment. In this case, the hydrophobic molecules spontaneously aggregate because of hydrophobic interaction, and in this way the hydrophobic surfaces of the macromolecules become hidden from the polar surrounding. Entropy increases ($\Delta S > 0$) owing to a displacement of the highly structured solvent molecules surrounding the exposed surface of the hydrophobic molecules towards the solvent bulk consisting of less structured molecules. As a consequence, the Gibbs energy decreases ($\Delta G < 0$), and therefore, hydrophobic interaction becomes a thermodynamically favorable process. In conclusion, the hydrophobic interaction between two or more non-polar molecules in a polar solvent solution is a spontaneous process governed by a change in entropy. Accordingly, hydrophobic interactions can be weakened by raising temperature or by modifying the solvent polarity through the addition of another solute.

2.2 Retention mechanisms in Hydrophobic Interaction Chromatography

Macromolecule retention in HIC occurs due to hydrophobic interactions between the hydrophobic ligands immobilized on a stationary phase and the hydrophobic moieties on the macromolecule surface (Queiroz et al., 2001). There is a variety of stationary phases used in HIC, corresponding to organic polymers or silica. Their main characteristics are being chemically modifiable, highly porous, and of high moisturizing power. Among them, the most commonly used are polyacrylamide (BiogelP^{TM}), cellulose (Cellulafine^{TM}), dextran (Sephadex^{TM}), agarose (Sepharose^{TM}), and others. These supports are further modified by linking hydrophobic ligands that become a sort of "active group" that allows hydrophobic interaction with the macromolecule to be separated from a solution. The ligand is linked to the support through a spacer arm (usually glycidyl ether), so that there is no steric impediment for macromolecule-ligand interaction, and avoiding hydrophobic interaction between the ligands. Figure 1 depicts the retention of a protein to a HIC stationary phase.

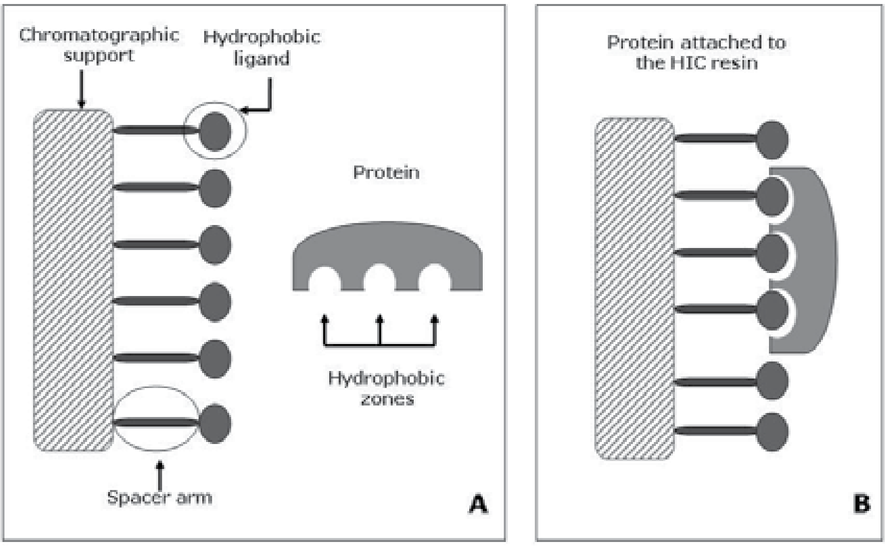


Fig. 1. Protein retention mechanism in HIC. (A) The basic structure of a HIC resin is depicted, and a protein is schematized highlighting the hydrophobic zones on the protein surface. (B) The protein gets in contact with the hydrophobic ligands of the resin, suffering a spatial reorientation. The hydrophobic ligands of the matrix interact with the exposed hydrophobic zones of the protein, and thus the protein is reversibly attached to the resin.

The most common hydrophobic ligands are alkyl or aryl groups of 4 to 10 carbons (Jennissen, 2000). The length of the carbon chain usually does not exceed 10 units in order to avoid self-folding. The nature of the hydrophobic ligand determines the performance of a HIC process. Figure 2 shows a scheme of stationary phases used in HIC and the chemical structure of the most commonly used alkyl and aryl groups, such as butyl (four carbons), octyl (eigth carbons) and phenyl (aromatic ring that promotes π-π interactions with the aromatic residues on a proteins surface). The hydrophobic interaction is directly proportional to the length of the alkyl chain. The most commonly used ligands in HIC resins are butyl, octyl and phenyl, in the following order in terms of relative interaction strength:

Phenyl > Octyl > Butyl

In the HIC process, retention is reinforced by the presence of a neutral salt. When a neutral salt is added to a solution consisting of a polar solvent, i.e. water, and a non-polar macromolecule, such as a protein, a competition for the water molecules that hydrate the macromolecule is observed, being more favorable to the salt. As a consequence, high salt concentration will reduce the number of solvent molecules that surround the macromolecules, thus favoring the hydrophobic interaction between them. Furthermore, if such solution comes in contact with a HIC resin, the interaction between the macromolecule and the hydrophobic ligand on the resin surface will be promoted, resulting in the adsorption of the macromolecule to the HIC stationary phase. From a process point of view, it is essential to choose the right type of salt and a concentration that minimizes macromolecule precipitation due to solubility decrease in the presence of high salt concentration (“salting-out”).

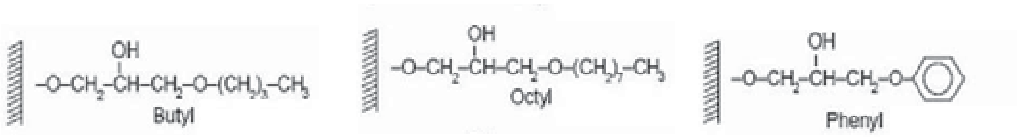
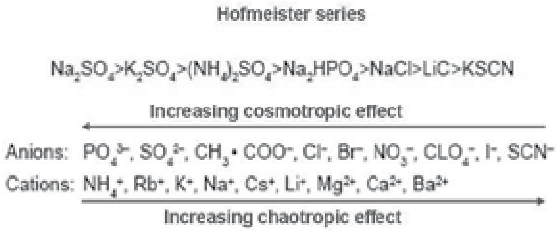

OH
$-O-CH_2-CH-CH_2-O-(CH_2)_3-CH_3$
Butyl
OH
$-O-CH_2-CH-CH_2-O-(CH_2)_7-CH_3$
Octyl
OH
$-O-CH_2-CH-CH_2-O-$
Phenyl

Fig. 2. Schematic representation of stationary phases used in HIC. Butyl is the shortest carbon chain used as HIC ligand and therefore the less hydrophobic one; octyl exhibits an intermediate hydrophobicity, and phenyl offers the strongest hydrophobic interaction.

The effect of different types of salt on macromolecule retention in HIC follows the Hofmeister (or lyotropic) series according to their positive influence in increasing the molal surface tension of water (Melander & Horvath, 1977). Besides, anions and cations exhibit cosmotropic or chaotropic properties. The salts at the beginning of the series are known as "cosmotropic" or "antichaotropic", since they promote hydrophobic interactions (as well as protein precipitation due to the "salting-out" effect) because of their water structuring ability. On the other hand, the salts at the end of the series, called "chaotropic", tend to randomize the structure of water and therefore they disfavor hydrophobic interactions. The salts ammonium sulfate and sodium chloride are most preferred in HIC.


Hofmeister series
$Na_2SO_4>K_2SO_4>(NH_4)_2SO_4>Na_2HPO_4>NaCl>LiC>KSCN$
Increasing cosmotropic effect
Anions: PO_4^{3-}, SO_4^{2-}, $CH_3 \cdot COO^-$, Cl^-, Br^-, NO_3^-, CLO_4^-, I^-, SCN^-
Cations: NH_4^+, Rb^+, K^+, Na^+, Cs^+, Li^+, Mg^{2+}, Ca^{2+}, Ba^{2+}
Increasing chaotropic effect

Once the macromolecule of interest is attached to the stationary phase, it is necessary to detach it in order to recover it as a bio-product. Desorption is most commonly accomplished by reducing the ionic strength in the mobile phase, by building a decreasing gradient of salt concentration (Fausnaugh et al., 1984). In this stage, the hydrophobic interaction between the macromolecule and the ligand is weakened as salt concentration diminishes in the mobile phase. As a consequence, the macromolecule is desorbed when a specific salt concentration is reached. This salt concentration, or ionic strength, depends on the physicochemical properties of the macromolecule. In this way, HIC can be used to selectively detach different macromolecules in a solution, thus becoming a powerful separation process.

Protein retention in HIC has been interpreted in the light of the underlying thermodynamic phenomena, by considering the effect of salt. Melander et al. (1989) proposed a thermodynamic model that describes protein retention in terms of electrostatic and hydrophobic interactions. This model describes protein retention due to only electrostatic interactions (case of ion Exchange Chromatography), only hydrophobic interactions (case of HIC), and both types of interactions (case of a weakly hydrophobic support or a chromatographic support bearing both hydrophobic and charged ligands). Simplifications of this model have been used to develop methodologies to predict protein retention in HIC. This model is described below.

2.3 Thermodynamic model for protein retention in HIC

The thermodynamic model proposed by Melander et al. (1989) to describe the effect of salt concentration on macromolecule retention in chromatography (IEC and HIC) can be applied to any stationary phase consisting of a highly hydrated surface modified with charged ligands (in the case of IEC), weakly hydrophobic moieties (in the case of HIC), or both. Electrostatic and hydrophobic interactions between the macromolecule and the stationary phase are treated separately. Electrostatic interaction is modeled based on the Manning's counter ion condensation theory (Manning, 1978), whereas hydrophobic interaction is treated by considering an adaptation of the Sinanoglu's solvophobic (Sinanoglu, 1982) theory that relates the salting out of proteins with their retention in HIC (Melander & Horvath, 1977). Figure 3 depicts protein retention due to hydrophobic interactions, electrostatic interactions, and both hydrophobic and electrostatic interactions.

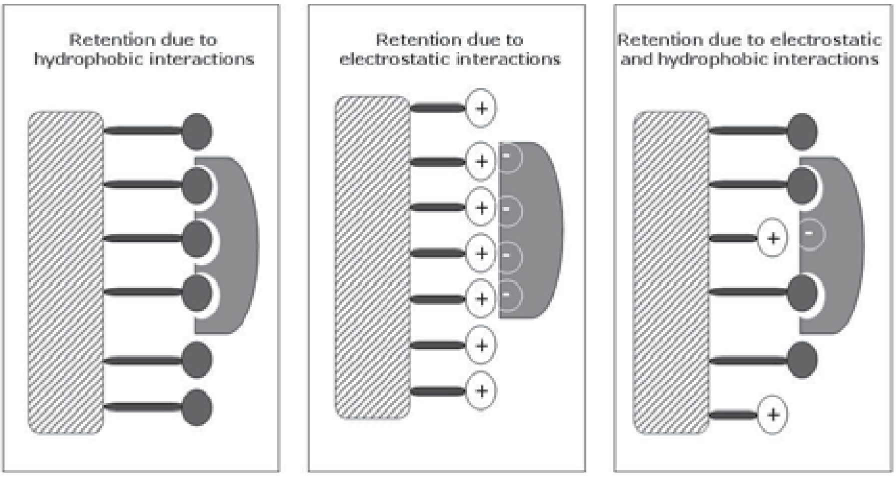

Retention due to hydrophobic interactions
Retention due to electrostatic interactions
Retention due to electrostatic and hydrophobic interactions

Fig. 3. Protein retention due to hydrophobic, electrostatic, and both interactions. Electrostatic interactions are long-range interactions, and then moieties with opposite charges do not need to be in physical contact. Hydrophobic interactions are short-range, and then interacting hydrophobic moieties must be in contact. As a consequence, when hydrophobic and charged moieties are present, both types of interactions may occur.

The main assumptions considered in the model are listed below:

i. The dimensions of the pores of the support are large with respect to the macromolecule size, and their shape is approximately a cylinder of infinite radius and the size-exclusion effects are negligible.
ii. The immobilized charges or hydrophobic moieties are uniformly spaced and equally accessible at the pore wall.
iii. The macromolecule is spherical and presents uniformly distributed and equally accessible fixed charges and hydrophobic patches on its surface.
iv. The macromolecule does not suffer conformational changes during the adsorption-desorption process.
v. Only a small fraction of the binding moieties on the stationary phase are occupied by the macromolecule.
vi. There are no specific interactions between the salt and the macromolecule.

2.3.1 Electrostatic interaction

The Gibbs energy of binding (ΔG^0_{es}) of the macromolecule to a stationary phase in presence of salt (that acts as a counter ion) is given by equation (2). Here m_s is the molal salt concentration, N_{Av} the Avogadro's number, "e" the base of the natural logarithm, "b" the average spacing of fixed charges on the surface, δ_p the thickness of the condensation layer over the surface of the stationary phase where each fixed charge occupies an area of b^2, δ_s the layer thickness of salt counter ion, Z_p the characteristic charge of the protein, Z_s the valence of the salt counter ion, and ξ a dimensionless structural parameter that characterizes the charged surface. R is the universal constant of gases and T the absolute temperature.

$$\frac{-\Delta G^0_{es}}{2.3\cdot R\cdot T}=\log\left(\frac{N_{Av}\cdot b^2\cdot\delta_p}{1000\cdot e}\right)+\frac{Z_p}{Z_s}\log\left(\frac{1000\cdot e}{\left(N_{Av}\cdot b^2\cdot\delta_s\cdot m_s\right)\cdot(1-Z_s\cdot\xi)}\right) \tag{2}$$

2.3.2 Hydrophobic interaction

The contact between the hydrophobic patches on the macromolecule surface that are exposed to the solvent and the hydrophobic ligands on the stationary phase, trigger the retention due to hydrophobic interaction. The Gibbs energy of hydrophobic interaction (ΔG_{hp}) is expressed in terms of the molal surface tension increment of the salt (σ_s), as shown in equation (3), which is valid only in the absence of specific salt effects. In Equation (3) m_s is the salt molality, ΔG^0_{aq} represents the reduction in Gibbs energy due to other effects different form hydrophobic interactions, $\Delta A'$ is the difference between the molecular surface area of the unbound macromolecule (A_M) and the molecular surface area of the macromolecule attached to the stationary phase (A_s). $\Delta A'$ corresponds to the surface contact area between the bound protein and the hydrophobic site of the matrix.

$$\Delta G^0_{hf}=\Delta G^0_{aq}-\Delta A'\cdot\sigma_s\cdot m_s \tag{3}$$

2.3.3 Combined electrostatic and hydrophobic interaction

The retention factor (k'), given in equations (4) and (5), is represented in terms of salt molality when both electrostatic and hydrophobic interactions are present. This is accomplished by combining equations (2) and (3) to give equation (6). In equation (5), K is the equilibrium constant and ϕ is the phase ratio (stationary phase mass / mobile phase mass). In equation (6) α is the phase volume ratio (stationary phase/mobile phase). Equation (6) can be written in a simplified form, as given by equation (7), where A is a constant determined by all the system characteristics, B the electrostatic interaction parameter and C the hydrophobic interaction parameter. In equation (7), the term C accounts for the hydrophobic surface contact area between the macromolecule and the stationary phase, and is given by equation (8).

$$\log K=\left(\frac{-\Delta G^0_{es}}{2.3\cdot R\cdot T}\right)-\left(\frac{-\Delta G^0_{hf}}{2.3\cdot R\cdot T}\right) \tag{4}$$

$$k'=\phi\cdot K \tag{5}$$

$$\log k' = \log\left(\frac{N_{Av}\cdot b^2\cdot\delta_p}{1000\cdot e}\right) + \frac{Z_p}{Z_s}\log\left(\frac{1000\cdot e}{\left(N_{Av}\cdot b^2\cdot\delta_s\cdot m_s\right)\cdot\left(1-Z_s\cdot\xi\right)}\right) + \frac{\Delta G^0_{aq}}{2.3\cdot R\cdot T} + \frac{\Delta A'\cdot\sigma_s\cdot m_s}{2.3\cdot R\cdot T} + \log\alpha \tag{6}$$

$$\log k' = A - B\cdot\log m_s + C\cdot m_s \tag{7}$$

$$C = \frac{\Delta A'\cdot\sigma_s}{2.3\cdot R\cdot T} \tag{8}$$

Equation (7) corresponds to the Simplified Thermodynamic Model for Electrostatic and Hydrophobic Interactions. This model is of practical usefulness, since its parameters can be obtained from experimental runs in a relatively simple manner, depending on the salt concentration present in the macromolecule solution. At low salt concentration, up to 0.5 molal, hydrophobic interactions can be neglected, and therefore the parameters A and B in equation (7) can be estimated by means of a linear regression between isocratic retention factors obtained at different salt molalilies. At high salt concentration, electrostatic interactions are negligible, and hence the parameters A and C can be obtained in a similar way, considering the isocratic retention factors. The hydrophobic contact area (ΔA′ in equation (8)) can easily be obtained from the slope of the limiting plot of log k′ versus molal salt concentration.

The simplified thermodynamic model has been used to investigate the effect of surface hydrophobicity distribution of proteins on retention in HIC (Mahn et al., 2004). The applicability of the model to predict protein retention time in HIC was demonstrated, and for the first time it was experimentally proven that surface hydrophobicity distribution has an important effect on protein retention in HIC. Furthermore, it was shown that the parameter ΔA′ that comes from equations (7) and (8) was able to represent the protein retention in HIC with salt gradient elution. However, the methodology proposed by Mahn et al. (2004) requires the generation of a considerable amount of experimental data, thus limiting its application.

3. Hydrophobic Interaction Chromatography process

The HIC process consists of injecting a macromolecule solution in a column packed with a stationary phase specifically designed to promote hydrophobic interaction with macromolecules such as proteins (solute). Usually retention is accomplished under high salt concentration conditions. Elution is achieved by decreasing the ionic strength in the mobile phase, building a decreasing salt gradient. At a microscopic level, the macromolecule enters in contact with the hydrophobic ligands at the pores surface of the resin, suffering a spatial reorientation. The hydrophobic ligands of the stationary phase interact with the hydrophobic zones of the macromolecule exposed to the solvent (usually aqueous solution), and thus the protein is reversibly attached to the resin.

Figure 4 shows a schematic representation of a HIC process. Here, A and B represent the vessels that contain the buffers used to manage the chemical environment in order to promote adsorption and desorption of the macromolecules present in the sample. The solution in A corresponds to a buffer with a low concentration of a neutral salt (usually 0.1 M), aiming to stabilize the macromolecular three-dimensional structure. The solution in B corresponds to buffer "A" added with a high salt concentration (usually higher than 1 M). Adsorption is promoted by using buffer "B", while desorption is induced by mixing both A and B forming a decreasing gradient salt concentration.

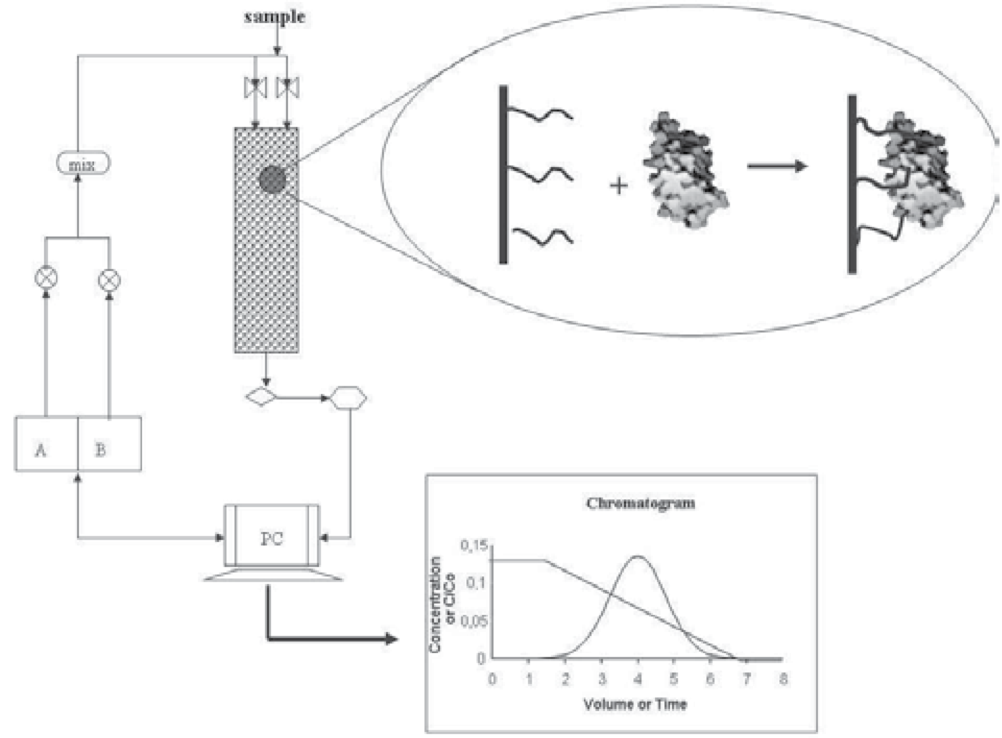


Fig. 4. Schematic representation of the HIC process. The HIC process consists of injecting a protein sample in a hydrophobic column under high salt concentration conditions such that hydrophobic interaction between the protein and the resin is promoted. Elution is achieved by decreasing the ionic strength in the mobile phase, building a decreasing salt gradient. In a microscopic level, the hydrophobic patches on the protein surface interact with the hydrophobic ligands of the resin, being reversibly attached to it. The protein concentration in the outlet is recorded as a function of time, and then a chromatogram is obtained.

The macromolecule concentration in the outlet solution is continuously determined through absorbance at 280 nm, and finally the elution curve or "chromatogram" is obtained. The chromatographic behavior in HIC can be characterized by several parameters, including the elution curve (most commonly by using the theoretical plate theory), the retention time or volume, or other parameters based on the preceding ones. To predict the behavior of proteins in HIC, the preferred parameter is the "Dimensionless Retention Time" (DRT), given by equation (9), where t_R is the time corresponding to the peak maximum, t_0 is the time at the beginning of the elution gradient, and t_f the time at the end of the gradient. In

HIC, the exploited property is hydrophobicity (Eriksson, 1998), and accordingly retention time is highly influenced by this property. Therefore, knowing macromolecule hydrophobicity allows predicting its behavior in HIC. Currently there is no universally agreed definition of protein hydrophobicity, but there is consensus in that it is determined by the hydrophobic contribution of the amino acids that compose the protein (Tanford, 1962).

$$DRT = \frac{t_R - t_0}{t_f - t_0} \tag{9}$$

On the other hand, protein retention in HIC is significantly affected by the operating conditions, which influence the resolution and selectivity of purification processes that include a HIC step (Ladiwala et al., 2006). From a process point of view, it is essential to count on methodologies and mathematical models to describe and to predict a protein behavior in HIC, ideally under varying operating conditions. Many efforts have been carried out to develop theories to explain this behavior based on protein properties, mainly protein hydrophobicity. At this point, controversial approaches have been proposed to theoretically estimate or experimentally determine protein hydrophobicity. These approaches include different amino acid hydrophobicity scales as well as diverse methodologies to perform calculations that use some scale to describe and predict protein retention time in HIC.

3.1 Protein hydrophobicity

3.1.1 Amino acid hydrophobicity scales

As stated above, protein hydrophobicity is determined by the hydrophobicity of the amino acids that compose it. Hence, it becomes necessary to quantify in any way the hydrophobic contribution of each amino acid. For this purpose, different approaches have been proposed to assign a hydrophobicity value to each one of the standard amino acids (Biswas et al., 2003; Kovacs et al., 2006). These methods are based on theoretical calculations and/or experimental determinations. Besides, the amino acid hydrophobicity scales differ in the hydrophobicity value assigned to each amino acid as well as in the relative position occupied by each one. These scales have been classified into several categories by different authors (Lienqueo et al., 2002; Mahn et al., 2009), based on their underlying principles.

Despite the differences between the hydrophobicity assigned to each residue by the different scales; it is clear a global tendency. Isoleucine shows the highest hydrophobicity in most scales, followed by Tryptophan. Glycine usually has an intermediate hydrophobicity level, i.e. neutral hydrophobicity, and the lowest level is mostly assigned to Aspartic acid (Lienqueo et al., 2007), i.e., this is the most hydrophilic amino acid. The suitability of the hydrophobicity scale depends on the use that will be given to the estimation of the protein or peptide hydrophobicity, as well as on the way to estimate this property. The scales proposed by Miyazawa & Jernigan (1996) and by Cowan & Whittaker (1990) are the most adequate to estimate protein hydrophobicity based on its three-dimensional structure (Lienqueo et al., 2007), regarding its behavior in HIC. Additionally, Salgado et al. (2005) proposed that the scale developed by Wertz & Scheraga (1978) is the most adequate to estimate protein hydrophobicity based on the amino acid composition of that protein.

	Cowan & Whittaker (1990)		**Miyazawa & Jernigan (1996)**		**Wertz & Scheraga (1978)**	
	Original	*Normalized*	*Original*	*Normalized*	*Original*	*Normalized*
ALA	0.420	0.660	5.330	0.391	0.520	0.375
ARG	-1.560	0.176	4.180	0.202	0.490	0.321
ASN	-1.030	0.306	3.710	0.125	0.420	0.196
ASP	-0.510	0.433	3.590	0.105	0.370	0.107
CYS	0.840	0.763	7.930	0.819	0.830	0.929
GLN	-0.960	0.323	3.870	0.151	0.350	0.071
GLU	-0.370	0.467	3.650	0.115	0.380	0.125
GLY	0.000	0.557	4.480	0.252	0.410	0.179
HIS	-2.280	0.000	5.100	0.354	0.700	0.696
ILE	1.810	1.000	8.830	0.967	0.790	0.857
LEU	1.800	0.998	8.470	0.908	0.770	0.821
LYS	-2.030	0.061	2.950	0.000	0.310	0.000
MET	1.180	0.846	8.950	0.987	0.760	0.804
PHE	1.740	0.983	9.030	1.000	0.870	1.000
PRO	0.860	0.768	3.870	0.151	0.350	0.071
SER	-0.640	0.401	4.090	0.188	0.490	0.321
THR	-0.260	0.494	4.490	0.253	0.380	0.125
TRP	1.460	0.914	7.660	0.775	0.860	0.982
TYR	0.510	0.682	5.890	0.484	0.640	0.589
VAL	1.340	0.885	7.630	0.770	0.720	0.732

Table 1. Amino acid hydrophobicity scales useful in HIC.

The Miyazawa & Jernigan (1996) scale is based on the three-dimensional structure of proteins, and it represents the contact energy between adjacent amino acids in folded protein. The Wertz & Scheraga (1978) scale is also based on knowledge of the folded protein structure, and it estimates the amino acid hydrophobicity as the ratio between the number of buried residues and the number of residues exposed to the solvent, for each type of standard amino acid. Both scales are based on clusters composed by a significant number of proteins whose three-dimensional structure had been elucidated through experimental methods. Both scales have been classified as indirect scales (Mahn et al., 2009). On the other hand, the Cowan & Whittaker (1990) scale, which has been considered a direct scale, assigned a hydrophobicity value to each standard amino acid based on the retention time of z-derivatives of each amino acid in HPLC. The scales mentioned above are presented in Table 1.

3.1.2 Estimation of protein hydrophobicity

There are different approaches to estimate protein hydrophobicity, which are based on different principles. The classical approach consists of estimating the "average surface hydrophobicity" ($\phi_{surface}$) based on the three-dimensional structure of the macromolecule in its native conformation (Lienqueo et al., 2002; Berggren et al., 2002). This approach considers only the amino acid residues that are accessible to the solvent at the protein surface, by using three-dimensional structural data. This method considers that each amino acid on the protein surface has a hydrophobic contribution proportional to its solvent accessible area, and the hydrophobicity of each residue is given by the amino acid hydrophobicity scale developed by Miyazawa & Jernigan (1996) or Cowan & Whittaker (1990), in their normalized form (see Table 1), as shown by equation (10).

$$\phi_{surface} = \frac{\sum \left(s_{aai} \cdot \phi_{aai} \right)}{s_p} \tag{10}$$

Here, $\phi_{surface}$ is the calculated value of the surface hydrophobicity for a given protein, i (i =1, . . ., 20; different i-values indicate different standard amino acids), s_{aai} is the solvent accessible area occupied by the amino acid i, ϕ_{aai} is the hydrophobicity value assigned to amino acid i by the hydrophobicity scale, and s_p is the total solvent accessible area of the entire protein. It has to be noted that for proteins with a prosthetic group s_p is bigger than the sum of the solvent accessible area occupied by the amino acids; and for proteins without prosthetic group, these values are equal. Table 2 shows the average surface hydrophobicity for a group of proteins using the amino acid hydrophobicity scales given in Table 1, and calculated by equation (10). This method for estimating protein hydrophobicity has proven to be valid in several cases (Lienqueo et al., 2002; Lienqueo et al., 2003; Lienqueo et al., 2007); however, this methodology is not valid for proteins that exhibit a highly heterogeneous distribution of the hydrophobic patches on their surfaces (Mahn et al., 2004).

Protein	Cowan & Whittaker	Miyazawa & Jernigan	Wetz & Scheraga
α-amylase	0.447	0.282	0.319
Citochrome c	0.362	0.185	0.171
Conalbumin	0.421	0.233	0.242
Concanavalin A	0.448	0.273	0.308
α-lactalbumin	0.491	0.318	0.304
β-lactoglobulin	0.468	0.279	0.284
Lysozyme	0.425	0.274	0.307
Myoglobin	0.392	0.214	0.220
Ovalbumin	0.457	0.257	0.270
Chymotrypsin	0.474	0.306	0.313
Chymotrypsinogen	0.468	0.298	0.305
Ribonuclease A	0.406	0.230	0.255
Thaumatin	0.464	0.269	0.279

Table 2. Surface hydrophobicity of proteins estimated by equation (9).

Genetic engineering is often used to improve the performance of separation and purification methods. Specifically in HIC, its performance has been improved by the fusion of short hydrophobic peptide tags such as T3, (TP)3, T3P2, T4, (TP)4, T6, T6P2, T8, (WP)2, (WP)4 to a protein of interest (Brandmann et al., 2000; Rodenbrock et al., 2000; Fexby & Bülow, 2004), thus increasing its original hydrophobicity. This genetic engineering strategy has the advantage that the structure/function changes are minimized in relation to the original properties of the native protein. Furthermore, the use of hydrophobic polypeptide tags allows investigating simple and less expensive stationary phases (in comparison with affinity chromatography supports), such as those used in HIC.

As a consequence, methods to calculate the surface hydrophobicity of tagged proteins have been proposed. One of those methods is the one proposed by Simeonidis et al. (2005) that allows computing the "tagged surface hydrophobicity" (ϕ_{tagged}), by equation (11). The surface hydrophobicity of the tagged protein is estimated as the average surface hydrophobicity of the original protein (without the tag) plus the hydrophobicity of the peptide tag. In this case, a fully exposed surface of the amino acids in the tag is assumed. In equation (11), n_k is the number of amino acids of "k" type (usually hydrophobic amino acids, such as tryptophan, leucine and isoleucine) in the tag, and s_{tag_aak} is the fully exposed surface of amino acid "k" in the tag.

$$\phi_{tagged} = \frac{\sum\left(s_{aai} \cdot \phi_{aai}\right)}{s_p} + \sum\left(\frac{\left(s_{tag\_aak} \cdot n_k\right)}{s_p + \sum\left(s_{tag\_aak} \cdot n_k\right)} \cdot \phi_{aak}\right) \tag{11}$$

Despite the remarkable results reached by the methods described above to estimate protein hydrophobicity, the need of knowing the three-dimensional structure appears as a serious disadvantage. This is especially clear from the ratio between the number of proteins of known three-dimensional structure available in the PDB database (Bermann et al., 2000) and the number of proteins sequenced in the UniProtKB/Swiss-Prot database (Bairoch et al., 2005). Currently (January 2011) this number is closer to 0.13 (70695/534420). This situation points out the need of a procedure based on low level information, such as the amino acidic composition. Salgado et al. (2005) developed a mathematical model to predict the average surface hydrophobicity of a protein based only on its amino acidic composition and, therefore, avoiding the use of its three-dimensional structure.

Equation (12) shows the basic structure of the model. In this equation, ASH represents the average surface hydrophobicity, n_i is the number of amino acids of class i in the protein, ^l is the normalized length of the protein sequence, and c_i correspond to adjustable parameters. The function f accounts for a correction of the amino acid composition of the protein according to different assumptions about the amino acids trend to be exposed to the solvent. The simplest form of f considers all the amino acids completely exposed. Parameters for building the function f were determined in a large set of non-redundant proteins by Salgado et al. (2005).

$$ASH = c_0 + \sum_{i=1}^{20} c_i \cdot f\left(n_i\right) + c_{21}\hat{l} \tag{12}$$

3.2 Methods for predicting retention time in HIC

The approaches discussed above to calculate protein hydrophobicity have been used to predict protein retention time by different methods. The simplest methodology uses straightforward quadratic models, whose parameters depend on the chromatographic conditions used in the HIC run (Lienqueo et al., 2007), and whose variables are DRT and the average surface hydrophobicity of the protein to be separated ($\phi_{surface}$). The most appropriate hydrophobicity scale was found to be that proposed by Miyazawa & Jernigan (1996), in its normalized form. The general model is given by equation (13), where A′, B′ and C′ are the model parameters that depend on the chromatographic conditions, such as type and concentration of salt and type of stationary phase. These parameters have been obtained from adjusting experimental data to the quadratic model. Table 3 shows the values of A′, B′ and C′ obtained for different operating conditions. The model given by equation (13) is useful for predicting retention times of structurally stable proteins that have a relatively homogeneous distribution of the surface hydrophobicity, such as ribonuclease A.

$$DRT = A' \cdot \phi_{surface}^{2} + B' \cdot \phi_{surface} + C' \tag{13}$$

Figure 5 shows a scheme of the methodology to predict DRT based on protein hydrophobicity. The procedure begins with the calculation of the protein surface accessible to the solvent, and the fraction of that surface occupied by each kind of amino acid. To calculate this, it is necessary to count on a PDB file, i.e. to know the spatial coordinates of each atom composing the macromolecule, preferably determined experimentally through X-ray crystallography or nuclear magnetic resonance (NMR). Experimentally determined structures can be obtained in The Protein Data Bank (PDB; www.rcsb.org/pdb) database (Bermann et al., 2000). Additionally, three-dimensional models can be found in other databases such as ModBase (http://modbase.compbio.ucsf.edu/modbase-cgi/search_form.cgi) (Pieper et al., 2009). Also it is required using a computational program or suit to perform the calculation, such as the software GRASP (Nicholls et al., 1991). With this information, the average surface hydrophobicity is calculated by means of equation (10) and using the Miyazawa & Jernigan hydrophobicity scale, in its normalized form. Finally, through a quadratic model like equation (13) the retention time of the protein can be estimated as DRT.

Resin	**Salt**	**Initial Salt molarity**	**A′**	**B′**	**C′**
Phenyl Sepharose	Ammonium sulfate	1	11.79	-0.29	0.35
Phenyl Sepharose	Ammonium sulfate	2	-12.14	12.7	-1.14
Phenyl Sepharose	Sodium chloride	2	-77.10	42.33	-5.13
Phenyl Sepharose	Sodium chloride	4	-65.01	37.55	-4.71
Butyl Sepharose	Ammonium sulfate	1	36.76	-16.07	1.73
Butyl Sepharose	Ammonium sulfate	2	10.02	0.45	-0.38
Butyl Sepharose	Sodium chloride	2	-12.05	6.51	-0.80
Butyl Sepharose	Sodium chloride	4	-1.74	5.55	-1.01

Table 3. Parameters of equation (12) for different operating conditions.

The surface hydrophobicity of tagged proteins (ϕ_{tagged}) has been used by Lienqueo et al. (2007) for predicting the DRT of cutinases tagged with hydrophobic peptides in different matrices for HIC, by means of equation (13) and the methodology represented in Figure 3. The coefficients of the linear model are constants for each set of operating conditions. This approach has proven to be effective in predicting the behavior of tagged proteins in HIC, since it showed a low deviation between predicted and experimental DRT (in the order of 2%), for the tagged cutinases that were studied. Finally, the ASH obtained from equation

(11) based on amino acidic composition was used to predict chromatographic behavior in HIC, resulting in a performance 5% better than that observed in the model based on the three-dimensional structure of proteins (equation (10)) (Salgado et al., 2008).

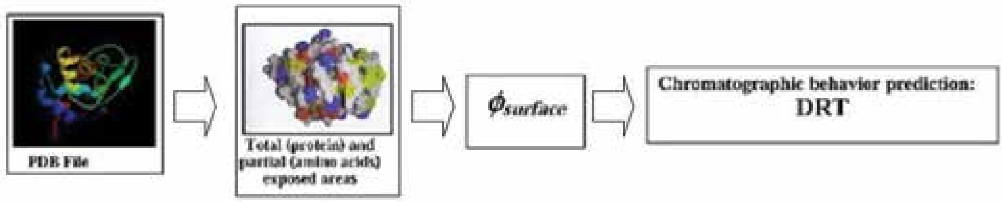

PDB File
Total (protein) and partial (amino acids) exposed areas
$\phi_{surface}$
Chromatographic behavior prediction:
DRT

Fig. 5. Methodology for predicting protein retention time in HIC based on surface hydrophobicity. Using a PDB file as input to the program GRASP, the total and partial accessible areas of the exposed amino acids is determined. Using an amino acid hydrophobicity scale and equation (12), the average surface hydrophobicity can be obtained. Then, through simple mathematical correlations the DRTof the protein can be estimated.

4. Applications in biomedical engineering

4.1 General applications

Currently, many proteins of pharmacological and industrial interest are obtained through highly optimized purification processes, typically consisting of two or three chromatographic separation stages. Usually these processes involve one or two IEC steps followed by a HIC step (Asenjo & Andrews, 2004). In addition, most recombinant proteins can be obtained at therapeutic grade of purity, by processes of the same structure (Asenjo & Andrews, 2008). Then, HIC often forms part of processes to yield a purified macromolecule of biomedical interest, such as therapeutic proteins (Seely & Richey, 2001), DNA vaccines (Diogo et al., 2000), and enzymes (Teng et al., 2010), among others. Besides, the use of HIC to purify protein complexes (McCue et al., 2008), as well as to study protein folding from a thermodynamic point of view (Geng & Wang, 2007), have been reported. Some applications of HIC for purifying enzymes and protein complexes, and to studying protein folding are described below.

4.1.1 Purification of proteins and enzymes by HIC

Recently, many strategies that involve a HIC step to purify proteins and enzymes of industrial and/or biomedical interest have been reported. For instance, Liu et al. (2010) developed a purification process to isolate and characterize an antifungal protein from *Bacillus subtilis,* which can be used as a bio-control agent. The process consisted of a preliminary precipitation with ammonium sulfate at 30-70% saturation, followed by HIC (using Phenyl Sepharose as stationary phase) and finally an IEC step. The process gave an overall recovery of 1.2% of total protein in the cell extract. The antifungal protein showed ribonuclease, protease and hemagglutinating activities.

On the other hand, Teng et al. (2010) purified and characterized an endo-β-1,4-glucanase from the giant snail (*Achatina fulica frussac*) by means of a process consisting of three chromatographic steps: size exclusion chromatography (SEC), anion exchange chromatography (AEC), and finally hydrophobic interaction chromatography. A 29-fold purity increase was achieved, and an overall recovery of 14.7% was reached. In addition, this novel enzyme has a particularly high stability at a broad pH range, acidic pH optimum, and a very high thermostability, and therefore it would have a great potential use in industry.

Lavery et al. (2010) reported the purification of a peroxidase from horseradish roots (*Armoracia rusticana*) by means of a three-step strategy, consisting of ultrasonication, ammonium sulfate precipitation, and HIC (using Phenyl Sepharose). In this strategy, the only high-resolution purification step corresponded to HIC. An overall yield of 71% and a 291-fold purification were achieved, thus demonstrating the high efficiency of this technique. The purified peroxidase was extremely stable in different media, and therefore its commercialization seems promising. Bhuvanesh et al. (2010) used a single-step method to purify a filarial protein (expressed heterologously in *E. coli*) with great potential as a vaccine for preventing human lymphatic filariasis. The purification method consisted of a HIC step. An overall recovery of 60% and 100% purity were achieved.

4.1.2 Purification of protein aggregates by HIC

The use of HIC to separate product-related impurities in the biopharmaceutical industry is well documented (Queiroz et al., 2001). This method is also used to separate multimers from monomeric forms of proteins of biomedical interest, since these conformations often differ in surface hydrophobicity. This difference owes to the fact that the stabilization of quaternary structures occurs due to hydrophobic interaction between the monomers, resulting in protein aggregation. In this way, the hydrophobic patches of a multimer are somewhat hidden, and therefore less accessible to the hydrophobic ligands of a HIC support, unlike the monomer whose hydrophobic patches are exposed to the solvent and, accordingly, accessible to the HIC ligands. The adsorption mechanism of protein aggregates in HIC is complex and not fully understood so far.

Mc Cue et al. (2010) developed a chromatography model to predict the separation of monomer and aggregate species. Equation (14) shows the Langmuir isotherm that describes equilibrium between the protein adsorbed to the resin and the protein that remains in solution. Here, C is the protein concentration in the mobile phase, q is the protein concentration in the stationary phase, q_m is the resin maximum capacity and k is the equilibrium constant. Equation (15) shows the mass balance used to describe the protein concentration profiles. Mass conservation was assumed and the intra-particle mass transfer was considered to be driven by homogeneous diffusion. In equation (15), D_{eff} is the effective diffusivity of the protein from the mobile phase bulk to the inner of the porous resin bead, t is time and r is the radial coordinate. The validity of the model was assessed by experimental determinations. A fraction of the aggregate proteins bound irreversibly to the HIC resin, becoming the major factor governing the process. This phenomenon was adequately described by the model.

$$C = \frac{q}{k \cdot (q_m - q)} \tag{14}$$

$$\frac{\partial q}{\partial t} = D_{eff} \cdot \left(\frac{\partial^2 q}{\partial r^2} + 2 \cdot \frac{\partial q}{r \cdot \partial r} \right) \tag{15}$$

4.1.3 Protein folding in HIC

Protein folding is relevant from a process point of view, since most recombinant proteins produced in bacteria such as *E. coli* are accumulated as inclusion bodies, and therefore protein refolding constitutes an additional stage in the production and purification process in order to yield a "functional" product (especially in the case of enzymes). Hydrophobic interaction chromatography has been used to study thermodynamics aspects of protein folding. For instance, Geng et al. (2005) performed calorimetric determinations on the enthalpy change ($\Delta H_{folding}$) of denatured lysozyme during its adsorption to a hydrophobic surface, with the simultaneous protein refolding. The surface consisted of PEG-600 made of a silica base HP-HIC (High Performance- Hydrophobic Interaction Chromatography) packing. At 25°C, $\Delta H_{folding}$ was found to be - 34 439 KJ/mol, involving adsorption, dehydration and molecular conformation enthalpies changes.

Later, Geng & Wang (2007) used the concept of "Protein Folding Liquid Chromatography" (PFLC), to describe a chromatographic process aiming to either raise the efficiency, or shortening the time of protein folding. Besides, an optimal PFLC should be able to simultaneously remove denaturant substances, separate contaminant proteins, promote refolding of the target protein, and ease denaturant recovery. Any type of chromatography can be used in PFLC, mainly Size Exclusion Chromatography, Ion Exchange Chromatography, Affinity Chromatography, and Hydrophobic Interaction Chromatography.

In HIC, the process is governed by thermodynamic equilibrium and so does the protein folding. PFLC provides the chemical equilibrium that favors the conversion from aggregate to desorbed protein, resulting in a higher refolding efficiency and shorter refolding time. The unfolded proteins, at a high ionic strength, are driven by hydrophobic interactions from the mobile phase to the HIC stationary phase, and the hydrophobic patches on the proteins surface get attached to the hydrophobic ligands, while the hydrophilic zones of the unfolded molecules remain in contact with the solvent. As a consequence, unfolded molecules are not able to aggregate. The unfolded molecules desorb from the HIC support as ionic strength in the mobile phase decreases. Protein molecules with incorrectly folded domains would be corrected by the spontaneous disappearance of the domains in the mobile phase due to their thermodynamic instability. After many HIC runs, the incorrectly folded domains will decrease, while the correctly folded molecules will predominate, resulting in protein refolding at high efficiency.

4.2 Applications in biomedical engineering

Biomedical applications of HIC are broad, since this technique offers some advantages over other chromatographic techniques, such as Affinity Chromatography (AC) and Reverse-Phase Chromatography (RPC). The use of AC depends on the availability of a specific ligand for the protein or group of proteins to be separated, thus limiting their applicability and raising its cost. The main disadvantage of RPC relies on the nature of the solvent in which the purified protein is recovered, usually an organic solvent not suitable for human or animal use. Then, HIC constitutes a purification tool suitable for biomedical applications, such as vaccines, therapeutic proteins, plasmids and mainly antibodies. In addition, the use of chromatography in high-throughput studies, such as proteomics and protein interactions, is increasing. Some of these Biomedical Engineering applications of HIC are discussed below.

4.2.1 Antibodies purification

At the beginning of the antibody industry, purification was performed through AC. For instance, protein A - AC was used for purifying monoclonal antibodies (MAbs), due to the extremely low MAb concentration in the initial solution (fermentation broth), and the high amount of contaminant proteins. Therefore, affinity chromatography was the most suitable technique, given its high selectivity and resolution. Unfortunately, this purification technique has a serious disadvantage given by the high affinity of the MAb for the ligand (such as protein A), making it difficult to release the MAb from the ligand, with the consequent economical detriment. Moreover, MAbs are highly hydrophobic macromolecules, and then the use of HIC has been suggested (Asenjo & Andrews, 2008). At the present time molecular biology advances have enabled reaching high concentrations of MAbs in the fermentation broth, making it possible to use less selective but cheaper purification techniques, such as HIC. Figure 6 depicts a monoclonal antibody (A) and the antibody attached to a HIC stationary phase (B).

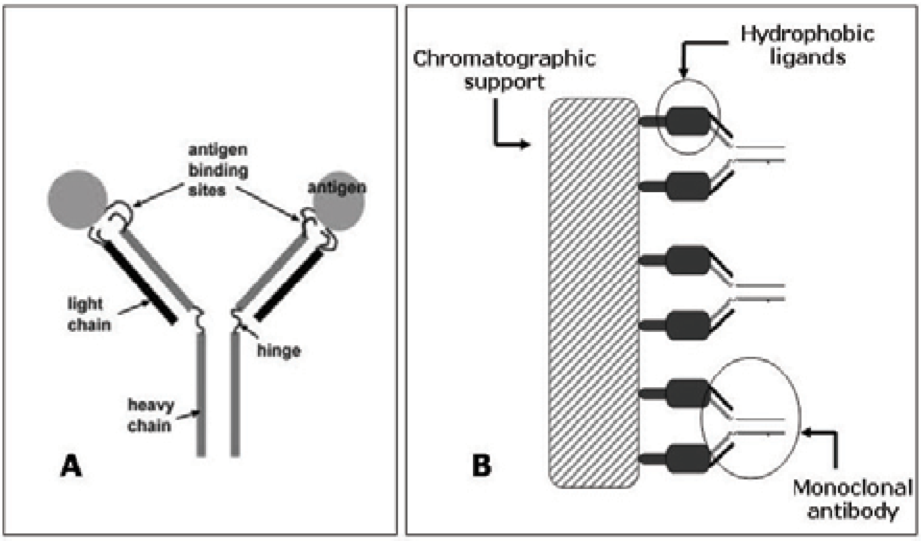


Fig. 6. (A) Schematic representation of a MAb. The antigen binding sites of the MAb are highlighted. Since this zone is characterized by an extremely high hydrophobicity, MAbs exhibit a high attraction for the hydrophobic ligands used in HIC resins. (B) Schematic representation of MAbs attached to a HIC resin. The antigen binding site interacts directly with the hydrophobic ligands of the HIC resin.

HIC is used as a polishing step in the purification processes of immunoglobulin-related products, since it has the ability to remove aggregated forms of the antibody (Rinderknecht & Zapata, 2006). Despite the high resolution offered by HIC, there are some drawbacks for its use in MAbs purification, given by the relatively low binding capacity of HIC supports and the consequent low yield in MAb recovery, compared to AC. Besides, MAb elution is usually achieved at a relatively high salt concentration, which implies that the solution containing the purified MAb also contains a high amount of salt that hinders sample manipulation and transitions during large-scale production.

This has encouraged research on HIC optimization, mainly regarding chromatographic supports. Recently, Chen et al., (2008) showed that the optimization of pore size of a HIC support significantly improved Immunoglobulin G binding capacity and also increased HIC efficiency, maintaining the MAb stability. Optimizing pore size facilitates mass transfer from mobile phase bulk towards the hydrophobic ligand. Kostareva et al. (2008) purified a heteropolymer (a kind of MAb consisting of a dual antibody conjugate) by HIC. They found that using a Propyl-HIC resin the heteropolymer was efficiently separated from free MAbs, thus confirming the ability of HIC for separating aggregates from monomers, and also its suitability for purifying MAbs.

4.2.2 Proteomics

Proteomics can be defined as the study of all the proteins codified by a genome, in a given tissue of a given organism at a given time. It involves studying how the concentration or "relative abundance" of the proteins change under a certain stimulus, protein conformational changes, protein – protein interactions (or "interactomics"), among others, as well as the use and development of experimental and bioinformatics technologies necessary to perform these studies. In this regard, protein separation techniques are essential. The fundamental separation methods used in proteomics are Sodium Dodecyl Sulfate- Polyacrylamide Gel Electrophoresis (SDS-PAGE) and/or Two-Dimensional Gel Electrophoresis (2DGE) and mass spectrometry (MS); the latter is used as separation but also as identification tool. Figure 7 depicts a classical proteomics experiment, starting from a biological sample, followed by preliminary fractionation by liquid chromatography and after that separation by 2DGE, and finally identification of protein spots by MS.

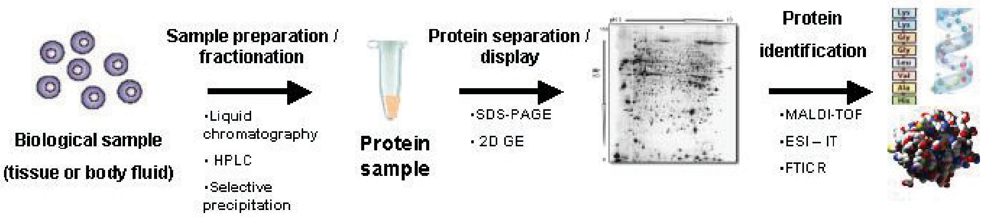


Fig. 7. Simplified representation of a gel-based proteomics experiment. Starting from a biological sample, a protein extract is obtained using different biochemical techniques to fractionate the sample. These fractionation steps allow the enrichment of protein fractions in low abundance proteins and to reduce the complexity of the sample. The protein fractions are then resolved by SDS-PAGE or 2D GE, and finally protein spots are excised form the gel and then analyzed by mass spectrometry in order to determine their identity and structural properties.

From a Biomedical point of view, proteomics is an important field in the task of discovering new biomarkers that reflect the health/disease status of living organisms. The use of proteomics with this purpose has been somewhat limited due to technical hurdles related to the high complexity of the biological samples to be analyzed, usually blood serum or plasma, but also cerebrospinal fluid, urine and tears. These samples show a wide dynamic range of protein concentration, exceeding 10^{10}. This means that the most abundant protein in plasma (albumin), for example, has a concentration 10^{10} times higher than that of the less abundant protein (such as transcription factors).

Two-dimensional gel electrophoresis can resolve a concentration range of up to 10^4, and therefore 2DGE images or "maps" of blood plasma are dominated by the highly abundant proteins, namely albumin, immunoglobulin, fibrinogen, among others, thus preventing the detection of low abundance proteins (Hoffmann et al., 2007). Mass spectrometry can resolve a range of 10^3 in a single spectrum, but combined with separation steps it can resolve a range of up to 10^6 (Jacobs et al., 2005). This range is still wide, and thus many proteins cannot be detected. Then, chromatographic separation steps should be used before 2DGE in order to reduce the dynamic range of proteins concentration, and consequently increase resolution.

The most abundant proteins in blood plasma are albumin, immunoglobulin, transferrin, haptoglobin, fibrinogen and α-1-antitrypsin, which amount to 90% of total protein mass. Then, total or partial depletion of these proteins allows detecting low abundance proteins. Different methods can be used to deplete these proteins, being liquid chromatography the most popular one (Nakamura et al., 2008). Different chromatographic strategies are available for this purpose, including affinity dye-based chromatography for albumin depletion, affinity to protein A and G for immunoglobulin depletion, specific antibody-affinity columns (Linke et al., 2007), and affinity columns containing lectins, peptides or inorganic ligands (Salih, 2005). Liquid chromatography has the advantage of being easy to use and to scale-up, but are relatively expensive, especially those involving affinity columns. Another drawback of affinity chromatography is the non-specific interactions that lead to the loss of some proteins, with the consequent loss of information (Altintas & Denizli, 2006).

In order to overcome the disadvantages of affinity chromatography for its use in blood plasma proteomics, several complementary strategies have been examined, such as sequential anion and cation exchange chromatography followed by 2DGE; and strong cation exchange chromatography followed by liquid-phase isoelectric focusing (Ottens et al., 2005; Barnea et al., 2005). Since these approaches considerably improve the capacity to detect low abundance proteins, it was suggested that the optimization of combinatorial processes by coupling immuno-affinity depletion with other conventional separation methods such as hydrophobic interaction chromatography will probably lead to significant advances in proteomics (Mahn et al., 2010). Despite the research conducted in this area, there is still a lack of optimized processes that ensure detection of the complete proteome of a tissue or cell.

4.2.2.1 Plasma fractionation by HIC

The applicability of HIC as a plasma fractionation method has been recently proposed. Geng et al. (2009) developed a two-dimensional liquid chromatography resin having two types of ligands, and hence that functions in two retention modes: cation exchange and hydrophobic interaction. This method could be applied to the fast fractionation of intact proteins before mass spectrometry analysis. The results obtained by HIC were similar to those obtained by ion exchange chromatography. On the other hand, a HIC matrix consisting of highly acetylated agarose has been used for the isolation of immunoglobulin from porcine serum, with a relative success (Ramos-Clamont et al., 2006).

Recently, Mahn et al. (2010) investigated if the performance of 2DGE could be improved by fractionating blood plasma through a HIC step, thus reducing the relative concentration of some highly abundant proteins in plasma. First, the hydrophobicity of the main 56 proteins present in blood plasma was determined. To do this, the amino acidic composition of the proteins was considered, and hydrophobicity was calculated by equation (16) based on the methodology proposed by Salgado et al. (2005). In equation (16), ϕ_{aai} is given by the Cowan–Whittaker hydrophobicity scale in its normalized form (see Table 1), n_i is the number of amino acids of type i in the protein, $s_{i,max}$ is the maximum solvent accessible area that an amino acid X can have when forming part of the G–X–G tripeptide in extended conformation (Miller et al., 1987).

$$ASH = \sum_{i=1}^{20} \left(\phi_{aai} \cdot \frac{n_i \cdot s_{i,\max}}{\sum_{j \in A} n_j \cdot s_{j,\max}} \right) \tag{16}$$

After that, a cluster analysis was performed in order to classify them as low, medium or high hydrophobicity proteins. This analysis showed that the highly abundant proteins, i.e. albumin, immunoglobulins, fibrinogen and haptoglobin, exhibited a medium hydrophobicity, and thus they fell in the same cluster. With this information, a HIC step was designed to deplete highly abundant proteins from rat plasma samples. The HIC step consisted of stepwise elution to separate the three groups of proteins (low, medium and high hydrophobicity) using a maximum concentration of 2 M ammonium sulfate, and concentration for elution of 0.6 M (to desorb low hydrophobicity proteins), 0.5 M (to desorb medium hydrophobicity proteins), and 0.0 M (to desorb the highly hydrophobic proteins).

Finally, the depleted samples were analyzed by 2DGE and the performance of the HIC pre-fractionation step was compared with that exhibited by a commercial immuno-affinity column. The reproducibility of 2DGE was similar to that obtained from immuno-affinity depleted plasma. However, HIC was more successful in depleting albumin and α-1-antitrypsin. Besides, HIC resulted in a much lower increment of immunoglobulin and haptoglobin abundances than the immuno-affinity column. Then, HIC depletion allowed detecting twice the number of protein spots than immuno-affinity depletion did. Therefore, HIC could be used as a depletion method complementary to affinity columns. The operating conditions in HIC could be optimized in order to maintain the high number of spots that are detected if HIC is used as the sole depletion method. Finally, given the relatively low cost of HIC supports and HIC operation, its use could be proposed as a convenient choice for depleting highly abundant proteins in plasma samples prior to 2DGE-based proteomics.

4.2.2.2 Analysis of protein interaction networks by HIC

Protein–protein interactions are essential in biological processes. All the interactions in a cellular system are known as protein interaction network or 'interactome'. In Biomedicine there is great interest in recognizing these interactions, aiming to establish the role they play in certain diseases. The traditional approaches to study protein-protein interactions are the antibody pull-down method (APD) and the yeast two-hybrid method (YTH). Despite their popularity, these methods have some disadvantages. It is very likely that a protein forms part of different complexes; then, in an APD experiment, antibodies targeting such a protein will pull down together all the complexes where the protein participates, making them appear to be part of a single large complex, confusing the biological interpretation of the results. The YTH is an "*in vivo*" method that allows detecting only binary interactions. It tends to give false positives and is limited to binary interactions. Therefore it is not useful in studying the dynamics of complex formation triggered by different stimuli (Corvey et al., 2005).

Liu et al. (2008) investigated the potential of chromatography to allow the simultaneous examination of multiple protein complexes along with comparing and validating results from the traditional methods. Since protein complexes remain intact during mild forms of elution in AC, a similar behavior should be expected in other chromatographic supports, such as IEC and HIC. They studied the extent to which protein interaction partners from yeast (*S. cerevisiae*) lysate remain associated during IEC, SEC and HIC. Most protein complexes remained intact, and all the proteins forming part of the complex migrated as a single unit. Protein complexes exhibited a chromatographic behavior different from that shown by the individual proteins that compose the complex. Accordingly, studying protein complexes could be easily performed by multidimensional chromatographic methods when at least one of the fractionation dimensions included SEC of native proteins. This method enables the study and recognition of several protein complexes simultaneously, avoiding the use of genetic engineering.

5. Conclusion

HIC is a powerful tool for purifying macromolecules of biomedical interest whose potential has been relatively under-exploited so far. Its applications are diverse, including industrial processes as well as analytical methods. The performance of HIC can be improved by optimizing the supports and the operation mode considering the hydrophobicity of the macromolecule to be separated. Research on optimization of HIC for biomedical applications should be encouraged, since this method allows reducing production cost of biopharmaceuticals such as antibodies and therapeutic proteins.

6. Acknowledgement

Fondecyt Programme.

7. References

Altintas, E.B., Denizli, A. (2006). Efficient removal of albumin from human serum by monosize dye-affinity beads. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 832, 2, 216-223.

Asenjo, J.A., Andrews, B.A. (2008). Challenges and trends in bioseparations. *Journal of Chemical Technology and Biotechnology*, 83, 117-120.

Asenjo, J.A., Andrews, B.A. (2004). Is there a rational method to purify proteins? From expert systems to proteomics. *Journal of Molecular Recognition*, 17, 236-247.

Bai, Q., Wei, Y.M., Geng, M.H., Geng X.D. (1997). High performance hydrophobic interaction chromatography - A new approach to separate intermediates of protein folding .1. Separation of intermediates of urea-unfolded alpha-amylase. *Chinese Chemical Letters*, 8, 67-70.

Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C., Redaschi, N., Yeh, L.S. (2005). The universal protein resource (UniProt). *Nucleic Acids Research*, 33, D154-D159.

Bandmann, N., Collet, E., Leijen, J., Uhlen, M., Veide, A., Nygren, P.A. (2000). Genetic engineering of the Fusarium solani pisi lipase cutinase for enhanced partitioning in PEG-phosphate aqueous two-phase systems. *Journal of Biotechnology*, 79, 161-172.

Barnea, E., Sorkin, R., Ziv, T., Beer, I., Admon, A. (2005). Evaluation of prefractionation methods as a preparatory step for multidimensional based chromatography of serum proteins. *Proteomics*, 5, 3367-3375.

Berggren, K., Wolf, A., Asenjo, J.A., Andrews, B.A., Tjerneld, F. (2002). The surface exposed amino acid residues of monomeric proteins determine the partitioning in aqueous two-phase systems. *Biochimica et Biophysica Acta-Protein Structure and Molecular Enzymology*, 1596, 253-268.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28, 235-242.

Bhuvanesh, S., Arunkumar, C., Kaliraj, P., Ramalingam, S. (2010). Production and single-step purification of Brugia malayi abundant larval transcript (ALT-2) using hydrophobic interaction chromatography. *Journal of Industrial Microbiology & Biotechnology*, 37, 1053-1059.

Biswas, K.M., DeVido, D.R., Dorsey, I.G. (2003). Evaluation of methods for measuring amino acid hydrophobicities and interactions. *Journal of Chromatography A*, 1000, 637-655.

Chaturvedi, R., Bhakuni, V., Tuli, R. (2000). The delta-endotoxin proteins accumulate in Escherichia coli as a protein-DNA complex that can be dissociated by hydrophobic interaction chromatography. *Protein Expression and Purification*, 20, 21-26.

Chen, J., Tetrault, J., Ley, A. (2008). Comparison of standard and new generation hydrophobic interaction chromatography resins in the monoclonal antibody purification process. *Journal of Chromatography A*, 1177, 272-281.

Corvey, C., Koetter, P., Beckhaus, T., Hack, J., Hofmann, S., Hampel, M., Stein, T., Karas, M., Entian, K.-D. (2005). Carbon source-dependent assembly of the Snf1p kinase complex in Candida albicans. *Journal of Biological Chemistry*, 280, 25323-25330.

Cowan, R., Whittaker, R.G. (1990). Hydrophobicity indices for amino acid residues as determined by high-performance liquid chromatography. *Peptide Research*, 3, 75-80.

Diogo, M.M., Queiroz, J.A., Prazeres, D.M. (2001). Studies on the retention of plasmid DNA and Escherichia coli nucleic acids by hydrophobic interaction chromatography. *Bioseparation*, 10, 211-220.

Diogo, M.M., Queiroz, J.A., Monteiro, G.A., Martins, S.A., Ferreira, G.N., Prazeres, D.M. (2000). Purification of a cystic fibrosis plasmid vector for gene therapy using hydrophobic interaction chromatography. *Biotechnology and Bioengineering*, 68, 576-583.

Eriksson, K., in: J.-C. Janson, L. Ryden (Eds.), Protein Purification: Principles, High-resolution Methods and Applications, 2nd ed., Wiley-Liss, New York, 1998, p. 151.

Fausnaugh, J.L., Regnier, F.E. (1986). Solute and mobile phase contributions to retention in hydrophobic interaction chromatography of proteins. *Journal of Chromatography*, 359, 131-146.

Fausnaugh, J.L., Kennedy, L.A., Regnier, F.E. (1984). Comparison of hydrophobic-interaction and reversed-phase chromatography of proteins. *Journal of Chromatography*, 31, 141-155.

Fexby, S., Bülow, L. (2004). Hydrophobic peptide tags as tools in bioseparation. *Trends in Biotechnology*, 22, 511-516.

Geng, X., Ke, C. , Chen, G., Liu, P., Wang, F., Zhang, H., Sun, X. (2009). On-line separation of native proteins by two-dimensional liquid chromatography using a single column. *Journal of Chromatography A*, 1216, 3553-3562.

Geng, X., Wang, C. (2007). Protein folding liquid chromatography and its recent developments. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 849, 69-80.

Haimer, E., Tscheliessnig, A., Hahn, R., Jungbauer, A. (2007). Hydrophobic interaction chromatography of proteins IV - Kinetics of protein spreading. *Journal of Chromatography A*, 1139, 84-94.

Hjerten, S. (1973). Some general aspects of hydrophobic interaction chromatography. *Journal of Chromatography*, 87, 325-331.

Hoffman, S.A., Joo, W.A., Echan, L.A., Speicher, D.W. (2007). Higher dimensional (Hi-D) separation strategies dramatically improve the potential for cancer biomarker detection in serum and plasma. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 849, 43-52.

Jacobs, J.M., Adkins, J.N., Qian, W.J., Liu, T., Shen, Y., Camp 2nd, D.G., Smith, R.D. (2005). Utilizing human blood plasma for proteomic biomarker discovery. *Journal of Proteome Research*, 4, 1073-1085.

Jennissen, H.P. (2000). Hydrophobic interaction chromatography. *International Journal of Bio-Chromatography*, 5, 131-138.

Kostareva, I., Hung, Campbell, F.C. (2008). Purification of antibody heteropolymers using hydrophobic interaction chromatography. *Journal of Chromatography A*, 1177, 254-264.

Kovacs, J.M., Mant, C.T., Hodges, R.S. (2006). Determination of intrinsic hydrophilicity/hydrophobicity of amino acid side chains in peptides in the absence of nearest-neighbor or conformational effects. *Biopolymers*, 84, 283-297.

Ladiwala, A., Xia, F., Luo, Q., Breneman, C.M., Cramer, S.M. (2006). Investigation of protein retention and selectivity in HIC systems using quantitative structure retention relationship models. *Biotechnology and Bioengineering*, 93, 836-859.

Lavery, C. B., MacInnis, M.C., MacDonald, M.J., Williams, J.B., Spencer, C.A., Burke, A.A., Irwin, D.J., D'Cunha, G.B. (2010). Purification of Peroxidase from Horseradish (Armoracia rusticana) Roots. *Journal of Agricultural and Food Chemistry*, 58, 8471-8476.

Lienqueo, M.E., Mahn, A. V., Asenjo, J.A. (2002). Mathematical correlations for predicting protein retention times in hydrophobic interaction chromatography. *Journal of Chromatography A*, 978, 71- 79.

Lienqueo, M.E., Mahn, A., Vasquez, L., Asenjo, J.A. (2003). Methodology for predicting the separation of proteins by hydrophobic interaction chromatography and its application to a cell extract. *Journal of Chromatography A*, 1009, 189-196.

Lienqueo, M.E., Salazar, O., Henriquez, K., Calado, C.R.C., Fonseca, L.P., Cabral, J.M. (2007). Prediction of retention time of cutinases tagged with hydrophobic peptides in hydrophobic interaction chromatography. *Journal of Chromatograohy A*, 1154, 460-463.

Linke, T, Doraiswamy, S., Harrison, E.H. (2007). Rat plasma proteomics: Effects of abundant protein depletion on proteomic analysis. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 849, 273-281.

Liu, B., Huang, L. Buchenauer, H., Kang, Z. (2010). Isolation and partial characterization of an antifungal protein from the endophytic Bacillus subtilis strain EDR4. *Pesticide Biochemistry and Physiology*, 98, 305-311.

Liu, X., Yang, W., Gao, Q., Regnier, F. (2008). Toward chromatographic analysis of interacting protein networks. *Journal of Chromatography A*, 1178, 24-32.

Mahn, A. V., Lienqueo, M.E., Asenjo, J.A. (2004). Effect of surface hydrophobicity distribution on retention of ribonucleases in hydrophobic interaction chromatography. *Journal of Chromatography A*, 1043, 47-55.

Mahn, A., Lienqueo, M.E., Salgado, J.C. (2009). Methods of calculating protein hydrophobicity and their application in developing correlations to predict hydrophobic interaction chromatography retention. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 1216, 1838-1844.

Mahn, A., Reyes, A., Zamorano, M., Cifuentes, W., Ismail. M. (2010). Depletion of highly abundant proteins in blood plasma by hydrophobic interaction chromatography for proteomic analysis. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 878, 1038-1044.

McCue, J.T., Engel, P., Ng, A., Macniven, R., Thömmes, J. (2008). Modeling of protein monomer/aggregate purification and separation using hydrophobic interaction chromatography. *Bioprocess and Biosystems Engineering*, 31, 261-275.

Melander, W.R., el Rassi, S., Horvath, Cs. (1989). Interplay of Hydrophobic and Electrostatic Interactions in Bio-polymer Chromatography - Effect of Salts on the Retention of Proteins. *Journal of Chromatography*, 469, 3-27.

Miller, S., Janin, J., Lesk, A.M., Chothia, C. (1987). Interior and surface of monomeric proteins. *Journal of Molecular Biology,* 196, 641-656.

Miyazawa, S., Jernigan, R. (1996). Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *Journal of Molecular Biology*, 256, 623-644.

Nakamura, T., Kuromitsu, J., Oda, Y. (2008). Evaluation of comprehensive multidimensional separations using reversed-phase, reversed-phase liquid chromatography/mass spectrometry for shotgun proteomics. *Journal of Proteome Research*, 7, 1007-1011.

Nicholls, A., Sharp, K.A., Honig, B. (1991). Protein Folding and Association - Insights from the Interfacial and Thermodynamic Properties of Hydrocarbons. *Proteins-Structure Function and Genetics*, 11, 281-291.

Ottens, A.K., Kobeissy, F.H., Wolper, R.A., Haskins, W.E., Hayes, R.L., Denslow, N.D., Wang, K.K. (2005). A multidimensional differential proteomic platform using dual-phase ion-exchange chromatography-polyacrylamide gel electrophoresis/reversed-phase liquid chromatography tandem mass spectrometry. *Analytical Chemistry*, 77, 4836-4845.

Pieper, U., Eswar, N., Webb, B.M., Eramian, D., Kelly, L., Barkan, D.T., Carter, H., Mankoo, P., Karchin, R., Marti-Renom, M.A., Davis, F.P., Sali, A. (2009). MODBASE, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Research*, 37, D347-D354.

Porath, J., Sundberg, L., Fornstedt, N., Olsson I. (1973). Salting-out in amphiphilic gels as a new approach to hydrophobic adsorption. *Nature*, 245, 465-466.

Queiroz J.A., Tomaz, C.T., Cabral, J.M.S. (2001). Hydrophobic interaction chromatography of proteins. *Journal of Biotechnology*, 87, 143-159.

Ramos-Clamont, G., Candia-Plata, M.C., Guzman, R., Vazquez-Moreno, L. (2006). Novel hydrophobic interaction chromatography matrix for specific isolation and simple elution of immunoglobulins (A, G, and M) from porcine serum. *Journal of Chromatography A,* 1122, 28-34.

Rinderknecht, E.H., Zapata, G.A. (2006). US Patent 7,038,017.

Rodenbrock, A., Selber, K., Egmond, M.R., Kula, M.R. (2000). Extraction of peptide tagged cutinase in detergent-based aqueous two-phase systems. *Bioseparation,* 9, 269-276.

Salgado, J.C., Andrews, B.A., Ortúzar, M.F., Asenjo, J.A. (2008). Prediction of the partitioning behaviour of proteins in aqueous two-phase systems using only their amino acid composition. *Journal of Chromatography A*, 1178, 134-144.

Salgado, J.C., Rapaport, I., Asenjo, J.A. (2005). Prediction of retention times of proteins in hydrophobic interaction chromatography using only their amino acid composition. *Journal of Chromatography A*, 1098, 44-54.

Salih, E. (2005). Phosphoproteomics by mass spectrometry and classical protein chemistry approaches. *Mass Spectrometry Reviews*, 24, 828-846.

Savard, J.M., Schneider, J.W. (2007). Sequence-specific purification of DNA oligomers in hydrophobic interaction chromatography using peptide nucleic acid amphiphiles: Extended dynamic range. *Biotechnology and Bioengineering*, 97, 367-376.

Seely, J.E., Richey, C.W. (2001). Use of ion-exchange chromatography and hydrophobic interaction chromatography in the preparation and recovery of polyethylene glycol-linked proteins. *Journal of Chromatography A*, 908, 235-241.

Shaltiel, S., Er-el, Z. (1973). Hydrophobic chromatography: use for purification of glycogen synthetase. *Proceedings of the National Academy of Sciences U.S.A.*, 70, 778-781.

Shepard, C.C., Tiselius, A. (1949). In "Chromatographic Analysis" p. 275. *Discussions of the Faraday Society*, 7. Hazell, Watson and Winey. London.

Simeonidis, E., Pinto, J.M., Lienqueo, M.E., Tsoka, S., Papageorgiou, L.G. (2005). MINLP models for the synthesis of optimal peptide tags and downstream protein processing. *Biotechnology Progress*, 21, 875-884.

Tanford, C. (1962). Contribution of hydrophobic interactions to the stability of globular conformation of proteins. *Journal of the American Chemical Society*, 84, 4240-4247.

Teng, Y., Yin, Q., Ding, M., Zhao, F. (2010). Purification and characterization of a novel endo-beta-1,4-glucanase, AfEG22, from the giant snail, Achatina fulica frussac. *Acta Biochimica et Biophysica Sinica*, 42, 729-734.

Wertz, D.H., Scheraga, H.A. (1978). Influence of water on protein structure. An analysis of the preferences of amino acid residues for the inside or outside and for specific conformations in a protein molecule. *Macromolecules*, 11, 9-15.

Zhang, Y., Martínez, T., Woodruff, B., Goetze, A., Bailey, R., Pettit, D., Balland, A. (2008). Hydrophobic interaction chromatography of soluble Interleukin 1 receptor type II to reveal chemical degradations resulting in loss of potency. *Analytical Chemistry*, 80, 7022-7028.

26

Development and Engineering of CSαβ Motif for Biomedical Application

Ying-Fang Yang
Biomedical Technology and Device Research Laboratories,
Industrial Technology Research Institute,
Republic of China

1. Introduction

Protein engineering is a process that modifies/creates functions or increases stabilities of proteins through artificial selection and evolution (Angeletti, 1998). For the fast development of molecular biology techniques, numbers of proteins have been successfully engineered to equip noval functions in the past several decades (Alahuhta et al., 2008;Bottcher & Bornscheuer, ;Ehren et al., 2008;Huston et al., 1988;Leta Aboye et al., 2008). These engineered proteins are developed either for academic research purposes or biomedical applications. The core of protein engineering is an appropriate protein scaffold (Hey et al., 2005;Pessi et al., 1993;Skerra, 2007). An excellent protein scaffold not only provides a platform for developing noval functions but also has many benefits, such as cost reduction during development/production or lasting efficacy of protein. An suitable protein scaffold should equip several characteristics (Hey et al., 2005;Pessi et al., 1993;Skerra, 2007). To introduce tailored functions into protein scaffolds is a major challenge and has a unique significance in protein design (Hey et al., 2005;Pessi et al., 1993).

A protein scaffold is a peptide framework that exhibits a high tolerance of its fold for modifications (Hey et al., 2005). Candidates for suitable protein scaffolds should exhibit a compact and structurally rigid core. The folding properties of the protein scaffolds should not be significantly changed when the side chains in a contiguous surface region are replaced or loops of varying sequence and length are presented (Skerra, 2007). Several additional priorities have to be considered if the scaffold is used in biomedical applications. The scaffold must display extra stabilities to environments, such as low pH, high concentration of chaotropic agents and high temperatures. Molecular weight of the scaffold should be low and small molecules have advantages in passing tissue barrier (Baines & Colas, 2006). The scaffold should highly resist protease degradation and this ensures the engineered proteins can be safe in the gastrointestinal tract and not degraded (Aharoni et al., 2005). Low immunogenecity is important to reduce unexpected side effects and damage of healthy tissue (Van Walle et al., 2007). Post-translational modification of protein is an important issue, too. Most of eukarytic proteins require post-translationally modified then gain their functions. These modifications includ glycosylation, cleavage of pre-peptide, formation of disulphide bridge and association of multiple peptides. Appropriate glycosylation also could reduce immunogenicity of proteins (Kosloski et al., 2009;Wang et al., 2010;Wu et al.). Currently, bacterial cells are the most conventional and convenious tools for mass production of recombinant proteins, but it is not easy for bacterial cells to undergo post-translational modication of eukaryotic proteins (Jacobs & Callewaert, 2009;Muir et al., 2009). Even expressed in eukaryotic cells, glycosylation of proteins is not exactly the same among difference specises (Perego et al., 2010). Unappropriatly modified recombinant proteins might lead to unexpected immune respones, if the proteins are used for medical purpose (Kosloski et al., 2009; Li & d'Anjou, 2009).

Protein designing challenges our understanding of the principles underlying protein structure and is also a good method to access our understanding of sequence-structure and structure-function relationship (Nikkhah et al., 2006). Rational design of proteins requires detailed knowledge of protein folding, structure, function, and dynamics (Chen et al., 2005). To build expression libraries, it is necessary to understand a protein scaffold in detail to amino acids usage on each residue position. This would reveal the key elements that affect functions and stabilities of a protein scaffold.

The appearance of new intellectual property, the breakthroughs in technology, or the increase in a market need are three major impacts to biopharma industry. Protein engineering is a branch of the biopharm industry, where intellectual property rights apparently play the most important role in the development and commercialization of final products. The intellectual property strategy to protect inventions of biopharma industry is to patent and to license them on an exclusive basis. The intellectual property also must be included in the linkage of research/development and business to ensure the commercial viability of biopharma products and to cope with the rapid changes of market. Currently, the intellectual property situation of biopharma industry is too complex and hampers the generation and production of recombinant protein/peptide drugs. Pantent analysis and patent map are necessary and helpful while planing the research and development and marketing. A well planed intellectual property strategy can not only protect the output of research and development but also defend market. A protein scaffold with simple legal situation will avoid knotty lawsuits. In recent years, shouts for alternative protein scaffolds is urgent and alternative protein scaffolds usually provide a favourable intellectual property situation. In this chapter, we will focus on a protein scaffold, cysteine-stabilized α/β (CSαβ) motif, and discuss protein engineering based on the scaffold.

2. Cysteine-stabilized α/β motif

2.1 Specific pattern and structural feature

CSαβ motif is a very unique protien scaffold. Proteins with a CSαβ motif express a high diversity in their protein primary structures (Figure 1) but share a common core sturcture that consists of a α-helix and an anti-parallel triple-stranded β-sheet (Figure 2 a to d). This is especially interesting not only for academic research but also very useful for applied utilities. From the protein sequence alignment analysis, six cysteines and one glycine (-C-X_l-CXXXC-X_m-GXC-X_n-CXC-) are exzactly conserved in all proteins containing a CSαβ motif (Carvalho & Gomes, 2009;Lay & Anderson, 2005). The cysteines form a framework and tightly connect the intramolecular structures. The most significant feature of the framework is two disulphide bonds which are formed with a patten of (*i, i+4*) and (*j, j+1*). The cysteines of (*i, i+4*) and (*j, j+1*) are located in the α-helix and the β3 strand of the β sheet; respectively (Figure 1). The helix and the sheet are connected by the disulphide bonds pairing with the pattern (*i*, *i*+*4*) and (*j*, *j*+*1*) (Assadi-Porter et al., 2000;Fant et al., 1999;Sun et al., 2002;Zasloff, 2002;Zhu et al., 2005). The third disulphide bond connects loop 1 and and β2 strand. In the β2 strand one glycin is conserved and locates in the central region of the motif. In plant defensins, there is usually a forth disulphide bond that seal the N- and C- terminee. The positively charged α-helix is another character. The positively charged residues are especially concentrated in the helix. β-sheet of the motif is formed by a hydrophobic amino acid clust and the sheet is composed by two or three anti-parallal strands. Loop regions of the mofit are highly diversed in lenght among proteins. There is a small turn that connects the helix and β2 strand.

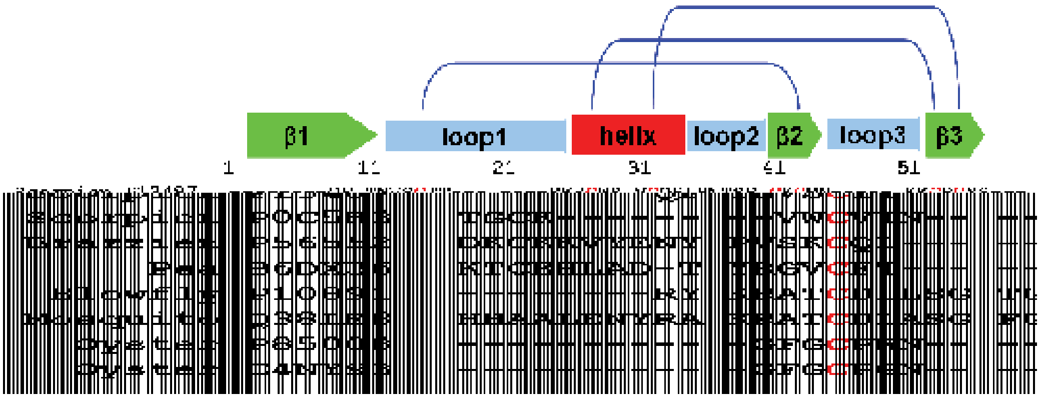
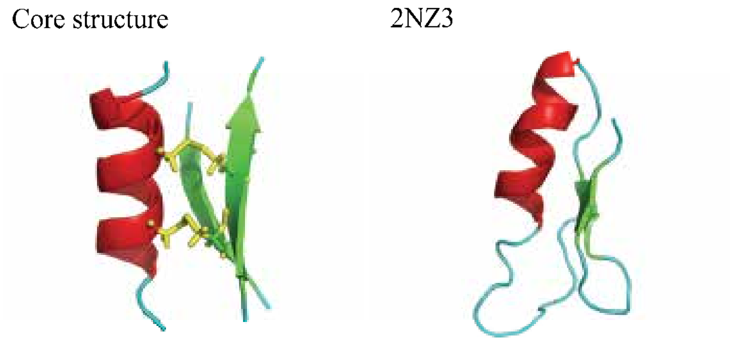


Fig. 1. The sequence alignment of proteins containing CSαβ motif. Sequences of two scorpion toxins (UniProt P13487, P0C5H3), two plant defensins (UniProt P56552, B6DX36), two insect defensins (UniProt P10891, Q38LE8) and two mollusk defensins (UniProt P85008, C4NY93) are aligned. The restrictly conservered residues, include six cysteins and one glycin, are high lighted in red. The green arrow box, red box and cyan boxs above the alignment represent β strands, α-helix and loops; respectively. The blue lines above on the top indcate the disulphide bond bwteen cysteines.

2.2 Biological distribution and functions

To date, there are at least four hundred proteins with a CSαβ motif have been discovered and deposited in databases. Proteins with the motif are widely distributed among plants, insects, arachnidia and mollusca (Sun et al., 2002;Zhu et al., 2005). They exhibit a wide spectrum of biological activities, including antimicrobial activity, enzyme inhibitory function, inhibition of protein translation and sweet taste (Assadi-Porter et al., 2000;Chen et al., 2005;Clauss & Mitchell-Olds, 2004;Spelbrink et al., 2004;Stec., 2006;Wong & Ng, 2005;Zhu et al., 2002). Proteins with the CSαβ motif usually serve a common function as defenders of their hosts (Lobo et al., 2007;Song et al., 2005;Zasloff, 2002).

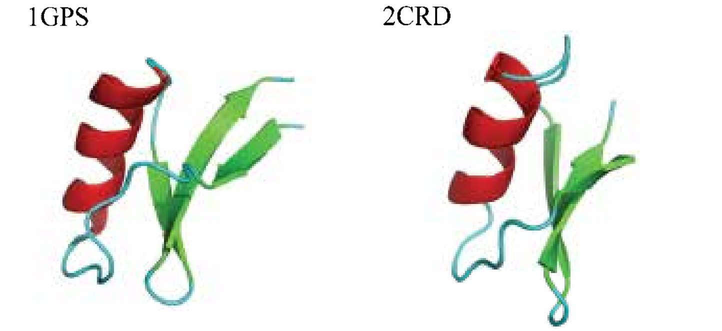
Before designing a unique function into a protein scaffold, it is important to understand the relationship between each part of the scaffold and its biological function. Based on the three-dimenional structure, protein scaffold containing a CSαβ motif can be devided into three parts: one α-helix, one β-sheet and three loops. It is well known that the three parts bearing different biological functions (Liu et al., 2006;Thevissen et al., 1996;Wong & Ng, 2005;Zhao et al., 2002).



Core structure

2NZ3

(a) The core structure of CSαβ motif. (b) Mosquito defensin (UniProt ID: Q17027)



1GPS

2CRD

(c) Wheat defensin (UniProt ID: P20158). (d) Scorpion toxin (UniProt ID: P13487)

Fig. 2. The core structure and three dimensional structrues of different specices. The structures are presented in color ribbon. Red: α-helix, green: β-sheet, and cyan: loop. Protein structures are retreved from Protein Data Bank and visualized with PyMol 0.99rc6.

The α-helix is related to antimicrobial ability. As described previously, the helix forms a positively charged cluster. When the positively charged residues were replaced with null or negatively charged amino acids, the anti-microbial ability of the proten is also changed. For its net positive charge, it is believed that proteins with the motif could interact with negatively charged cell membrane (Thevissen et al., 1996;Thomma et al., 2003). Several studies have demostrated that plant defensins are able to pass artificial membranes and lead to ions leaking from one side of the membrane. The mechanism about how plant defensin passing membrane is not revealed, yet.

Role of the β-sheet is less studied and disscussed. The direct mutagensis studies showed when the hydrophobic residues in the β-sheet are alanine substituted, biological function of mutated proteins are dromatically reduced (Walters et al., 2009;Yang et al., 2009). The maximal effects of mutated proteins are only 30-40% maximal effects of the wild type, even at a high protein concentration. Protein-protein docking model showed that the β-sheet could form interaction interface with their counterpart. These data imply that, the distribution of the hydrophobilc residues in the β-sheet plays a role in protein-protein interactions and β-sheet could relate to the protein-protein interaction specificity to their targets.

The length of loop regions are different from protein to protein and the loops connecting CSαβ motif can serve as functional epitopes (Figure 3) (Lay & Anderson, 2005;Wijaya et al., 2000;Zhao et al., 2002). For example, loop 1 of the *Arabidopsis thaliana* trypsin inhibitor (ATT_p) and loop 2 of cowpea thionine are the functional epitopes for trypsin inhibition (Wijaya et al., 2000;Zhao et al., 2002). The loop 3 of *Vigna radiate* defensin 1 (VrD1) is the functional loop for insect α-amylase inhibition (Lin et al., 2007;Liu et al., 2006).

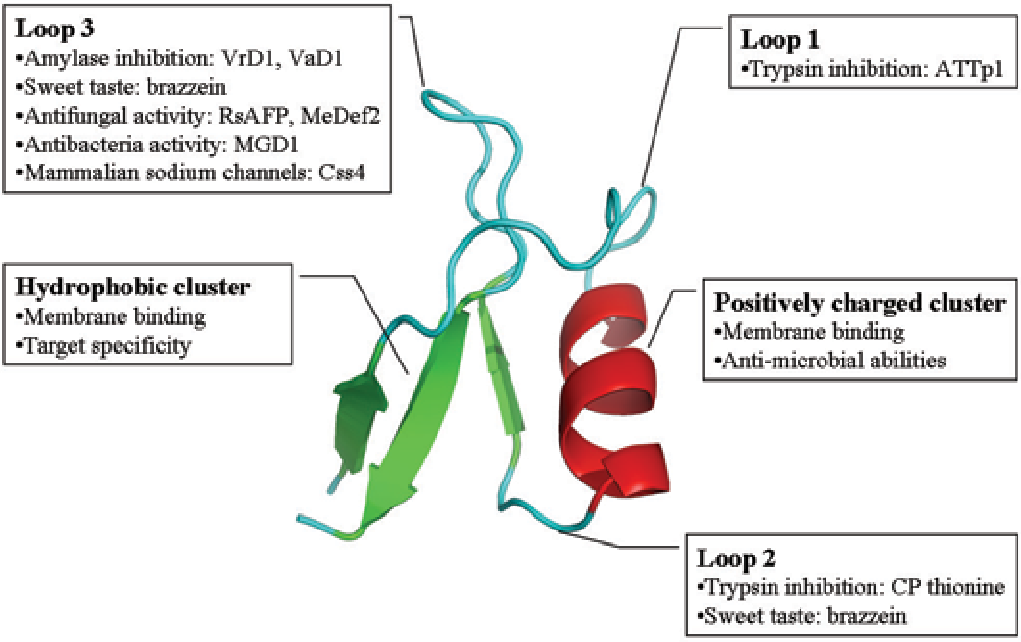

Loop 3
•Amylase inhibition: VrD1, VaD1
•Sweet taste: brazzein
•Antifungal activity: RsAFP, MeDef2
•Antibacteria activity: MGD1
•Mammalian sodium channels: Css4
Loop 1
•Trypsin inhibition: ATTp1
Hydrophobic cluster
•Membrane binding
•Target specificity
Positively charged cluster
•Membrane binding
•Anti-microbial abilities
Loop 2
•Trypsin inhibition: CP thionine
•Sweet taste: brazzein

Fig. 3. Structure-function relationship of peptide with a CSαβ motif. Red: α-helix, green: β strands, and cyan: loops.

2.3 Structural ultra-stability

A protein scaffold with ultra high stability can maintain its three dimensional structure in extreme environments and its functions can be preserved. Therefore, when a protein scaffold applied in biomedical applications, its structural stabilities must be considered. The scaffold should be able to pass a low pH environment in the stomach, resist protease digestion, endure chaotropic agents at high concentrations and so on. These criteria are to ensure the engineered proteins could reach their target sites and perform functions inside body, and are not destroyed. It has been reported that proteins with a CSαβ motif equip ultra stabilities to several extreme environments such as urea at a concentration of 6M, a temperature over than 95°C, and resist protease digestion without changing its structure (Malavasic et al., 1996;Yang et al., 2009). In Table 1, the properties of CSαβ motif and single-domain antibody are listed and compared (Holt et al., 2003;Skerra, 2007;Yang et al., 2009;Yang & Lyu, 2008). For their advantages, such as high specificity and affinity, antibodies still are the most popular protein scaffold for engineering. Antibodies are widely used both in routine laboratorial experiments and clinical diagnosis. In spite of their significant clinical success, several disadvantages, including high cost in manufacturing, large in size, undesired effector functions and complex intellectual property situations, obstruct their development and applications (Jones et al., 2008). Single-domain antibody is only constituted of antibody variable regions and does not have constant regions (Holt et al., 2003). Compared with single-domain antibodies, the proteins with a CSαβ motif have smaller molecular weights and higher structural stabilities (Holt et al., 2003;Skerra, 2007;Yang et al., 2009). As previously described, proteins with a CSαβ motif share low sequence identity but high structural similarity (Lin et al., 2007). Their functions are highly varied and sturctures are ultra-stable. There is a possibility to utilize CSαβ motif as an engineering scaffold for biomedical applications (Yang & Lyu, 2008).

Item	Molecule	
	Single-domain antibody	CSαβ motif protein
Molecular weight	11-15 kDa	5-7 kDa
Generation of expression library	B cell mRNA	Synthetic
Water solubility	Less	High
Ultra stability	Unfolded 6M Gdn-HCl T_m 60-78°C	Ultra stable 6M Gdn-HCl T_m > 96°C
Post-translational modification	Glycosylation	Disulfide bridge
High functional diversity	Yes	Yes
Tolerate to amino acid substitution	Loop regions	Structural and loop regions
Enzyme inhibition	Not certain	Direct inhibition
Membrane binding	Not certain	Direct binding
Legal problem	Very complex	Simple

Table 1. Comparison of properties bwtween CSαβ motif and single-domain antibody.

2.4 Amino acid usage of CSαβ motif

To understand the amino acid usage of a protein scaffold can reveal relationships among structures, functions and sequence residues (Kristensen et al., 1997;Yang et al., 2009). To completely understand the relationships could be an approach through extensive amino acid substitution and analysis of protein sequences (Corzo et al., 2007;Wang et al., 2006). Amino acid substitution have been performed in plant defensins, brazzein of *Pentadiplandra brazzeana* and VrD1 of *Vigna radiate*, and some key residue positions are discovered (Assadi-Porter et al., 2010;Yang et al., 2009). In both cases, amino acid substitution does not lead to the structure significantly being changed in all positions along the sequence but the replacement in some positions have effects on biochemical function (Assadi-Porter et al., 2010;Yang et al., 2009).

It has been noted that certain amino acids have preference to fold into a given secondary structure (Chan et al., 1995;Zhong & Johnson, 1992). Comprehending preference of amino acids usage will be really helpful to protein engineering and can be as a fundment for designing innovative peptides. The two major classes of CSαβ motif protein are plant defensins and scorpion toxins (Zhu et al., 2005). Currently, there are at least 140 sequences of scorpion toxin and 180 sequences of plant defensin deposited in the SwissProt database and the numbers are continuely increased. The peptide sequences of the scorpion toxin and plant defensin peptides deposited in the SwissProt database are retrieved and amino acid usage perferacnes are separately analyzed (Table 2 and Table 3).

In Table 2 and Table 3, the twenty amino acids are listed on the top row, the secondary structures and residue positions are listed on the left two columns. Use of amino acids in each position is counted and the usage frequency is calculated (form 0.00 to 1.00). Amino acids with a high usage frequency (> 0.20) are listed on the right column. The results are interesting and different between scorpion toxins and plant defensins. In scorpion toxins, peptides tend to employ similar amino acids in the same position. Therefore, sequences of scorpion toxins are less diverse and more uniformed. In plant defensins, sequences of peptides are more diverse. The major targets of scorpion toxins are ion channels on the neuron cell surface but the targets of plant defensins are various form peptide to peptide (Zhu et al., 2005). This might be the reason for the amino acid usage difference between the two classes of CSαβ motif peptides. In the helix region, scorpion toxins prefer using charged residues and plant defensins are more inclusive of different type of amino acids. The most interesting is the two residues flanking the exzactly conserved glycine on $\beta_2$2. In scorpion toxins, two tyrosins are prefered on $\beta_2$1 and $\beta_2$3, and the rates are 0.44 and 0.75; respectively. In plant defensins, a preceded glycine ($\beta_2$1) is prefered to the glycine on $\beta_2$2 and the frequency is 0.35. It is worth to clarify the role of the two tyrosins surrounding the glycine on $\beta_2$2 of scorpion toxins. Although the frequnectly used amino acids are listed, it just provides a reference for protein designing and there is no necessarity that combination of these amino acids forms a universal sequence.

	Position	Amino acids																				High frequency[a]
		A	G	V	L	I	F	Y	W	C	S	T	D	E	N	Q	R	H	K	M	P	
Strand 1	S1		0.81												0.12		0.03					G
	S2 (C)									1.00												C
	S3	0.05		0.03		0.01						0.03					0.02		0.85			K
	S4			0.34	0.13	0.10	0.03	0.22					0.02	0.01	0.02		0.10			0.03		V Y
	S5	0.05	0.07	0.02	0.00	0.06			0.04		0.51	0.06	0.03	0.03	0.02			0.02	0.05		0.01	S
Helix	H1	0.04	0.16	0.08							0.04	0.03	0.05	0.20	0.01	0.06	0.03	0.06	0.16		0.07	E
	H2		0.13		0.01	0.00	0.16	0.37	0.03		0.10		0.03	0.03	0.05	0.08		0.01				Y
	H3 (C)									1.00												C
	H4	0.09	0.05	0.03	0.05	0.10		0.01				0.01	0.25	0.06	0.22	0.03	0.03		0.03	0.01	0.01	D E
	H5	0.01	0.01	0.01	0.01	0.01					0.09	0.09	0.14	0.03	0.17	0.04	0.07	0.01	0.30		0.01	K
	H6	0.01		0.09	0.08	0.07	0.01			0.03	0.01	0.01	0.01	0.43	0.02	0.02		0.13	0.01	0.02	0.07	E
	H7 (C)									1.00												C
	H8	0.03	0.03	0.01	0.01						0.09	0.02		0.03	0.01	0.07	0.11	0.01	0.57			K
	H9	0.22	0.05	0.03	0.11	0.03	0.01	0.02			0.06	0.01	0.14	0.04	0.01		0.09		0.16	0.01		A
β_2 strand	$\beta_2$1	0.01	0.04	0.01		0.01	0.09	0.44			0.13	0.02	0.15	0.02	0.01			0.01	0.05			Y
	$\beta_2$2 (G)		1.00																			G
	$\beta_2$3					0.01	0.02	0.75			0.01	0.03	0.01	0.02		0.01	0.01	0.01	0.11			Y
	$\beta_2$4 (C)									1.00												C
	$\beta_2$5	0.01		0.01		0.01	0.02	0.34	0.26	0.00	0.03	0.01	0.03	0.01	0.00	0.01		0.10	0.01	0.07		Y W
β_3 strand	$\beta_3$1	0.49	0.05		0.03	0.00	0.01	0.01			0.05	0.01	0.01	0.02	0.00	0.04	0.02	0.01	0.25			A K
	$\beta_3$2 (C)									1.00												C
	$\beta_3$3	0.01			0.01	0.01	0.01	0.28	0.45				0.02		0.01	0.01		0.05	0.14	0.01		W Y
	$\beta_3$4 (C)									1.00												C
	$\beta_3$5	0.04	0.07		0.01	0.03	0.01	0.22	0.00		0.02	0.07	0.01	0.42	0.01	0.05			0.05	0.01		E Y

Table 2. The high frequent amino acids in the structural regions of scorpion β toxin.

	Position	Amino acids																				High frequency[a]
		A	G	V	L	I	F	Y	W	C	S	T	D	E	N	Q	R	H	K	M	P	
Strand 1	S1		0.02	0.08	0.20	0.07	0.01	0.03			0.01	0.27	0.02	0.08	0.01	0.02	0.07	0.01	0.04	0.04	0.02	T L
	S2 (C)									1.00												C
	S3	0.01		0.04	0.07	0.02	0.02	0.02	0.01		0.05	0.03	0.05	0.34	0.01	0.04	0.08	0.04	0.11	0.07	0.01	E
	S4	0.06	0.02	0.07	0.05	0.03					0.18	0.10	0.04	0.07	0.08	0.02	0.13	0.03	0.12		0.01	
	S5	0.05	0.04	0.05	0.12	0.05	0.01	0.04			0.06	0.05	0.03	0.07	0.00	0.12	0.08	0.01	0.08	0.01	0.13	
Helix	H1	0.07	0.10	0.02	0.01	0.02	0.01	0.01			0.15	0.15	0.05	0.03	0.14	0.05	0.06	0.04	0.08		0.04	
	H2	0.12	0.04	0.01	0.08	0.02	0.02	0.00	0.02		0.07	0.05	0.11	0.05	0.27	0.02	0.02	0.03	0.06	0.01	0.03	N
	H3 (C)									1.00												C
	H4	0.24	0.03	0.04	0.02	0.02	0.00	0.03			0.06	0.02	0.17	0.03	0.08	0.02	0.09		0.19			A
	H5	0.06	0.01	0.02	0.04	0.02	0.03		0.01		0.11	0.08	0.09	0.06	0.23	0.09	0.03	0.04	0.10		0.01	N
	H6	0.05	0.01	0.25	0.06	0.01	0.01	0.04		0.02	0.08	0.05	0.01	0.05	0.05	0.16	0.01	0.11	0.02	0.03	0.02	V
	H7 (C)									1.00												C
	H8	0.03	0.01	0.04	0.11	0.18	0.02	0.03	0.01		0.01	0.04	0.01	0.01	0.02	0.09	0.18	0.05	0.16	0.03	0.02	
	H9	0.06	0.02		0.01		0.01	0.01	0.01		0.09	0.21	0.02	0.07	0.19	0.11	0.08	0.01	0.10	0.01	0.02	T
β_2 strand	$\beta_2$1	0.05	0.35	0.05	0.01	0.02	0.01	0.01		0.01	0.11	0.03	0.08	0.02	0.05	0.01	0.05	0.14	0.05			G
	$\beta_2$2 (G)		1.00																			G
	$\beta_2$3	0.04	0.02	0.05	0.01	0.01	0.05	0.06			0.18	0.07	0.05	0.06	0.07	0.03	0.11	0.06	0.17	0.01		
	$\beta_2$4 (C)									1.00												C
	$\beta_2$5	0.02	0.01	0.07	0.05	0.06	0.04	0.01	0.01		0.06	0.02	0.14	0.04	0.17	0.03	0.16	0.09	0.06	0.01	0.01	
β_3 strand	$\beta_3$1	0.05		0.02	0.02	0.07	0.02	0.02			0.04	0.05	0.01	0.02	0.04	0.06	0.30	0.04	0.23	0.05		K
	$\beta_3$2 (C)									1.00												C
	$\beta_3$3	0.01	0.02	0.07	0.12	0.15	0.19	0.13	0.07			0.02	0.02	0.03	0.01	0.01	0.03	0.05	0.05	0.05		
	$\beta_3$4 (C)									1.00												C
	$\beta_3$5	0.02	0.02	0.03	0.04	0.06	0.02	0.22		0.04	0.07	0.30	0.03	0.02	0.03	0.04	0.04	0.01	0.04	0.02		Y

Table 3. The high frequent amino acids in the structural regions of plant defensin.

3. Applications of CSαβ motif peptides

Human beings have been using natural products containing CSαβ motif peptides for several hundreds years. The most well known natural products are sweet peptides of plants and scorpion venom. For their anti-microbial and pesticide abilities, in many cases, peptides with CSαβ motif are transfected into industrial crops to protect these transgenic plant from pathogens. In recent years, some of plant defensins are also screened and selected as antibiotics for therapeutic utilities. In this section, some peptides with therapeutic potential are selected and discussed.

3.1 Sweet peptide: Brazzein

Brazzein is isolated from *Pentadiplandra brazzeana*, a climbing plant plant of the West Africa. Berry fruits of *Pentadiplandra brazzeana* have sweetness and are traditional foods of African natives (Assadi-Porter et al., 2008;Yang & Lyu, 2008). For a long time, the sweetness of the berry fruits is a secret. Ding and Hellekant at University Wisconsin Madison isolated a small peptide with sweetness from *Pentadiplandra brazzeana* and named it as brazzein 1994 (Ming & Hellekant, 1994). Peptide sequence analysis shows that brazzein contains a set of (*i, i+4*) and (*j, j+1*) cysteine pattern. Its NMR structure is solved and confirms it as a CSαβ motif peptide 1998 (Caldwell et al., 1998). Brazzein is the smallest, most heat-stable and pH-stable protein known to have an intrinsic sweetness (Jin et al., 2006). Brazzein is 200 times sweeter than sugar and it can be used both in cold and hot drinks/foods without change its sweet taste (Jin et al., 2006). Brazzein is widely recurited as sugar substitute in low caloric dietary formulas or for diabetes patients. Brazzein is a food additive for several years and no side effects have been reported (Yang & Lyu, 2008).

3.2 Scorpion venom: Traditional remedies

Scorpion is therapeutically used in traditional Chinese medicine for hundreds of years. Figure 4 shows the therapeutic instruction of using scorpion described in the *Compendium of Materia Medica*, an ancient Chinese pharmacopoeia compiled in the 1500s AD. In traditional Chinese medicine, scorpion is recruited as one of the major treatments for stroke, rheumatism, epilepsy, hemiplegia, convulsions, cramps, twitchs, headaches, tetanus and scrofula. *Buthus martensii Karsch* is the most common species of scorpion that are processed as traditional Chinese medicine. Captured scorpions have to be preserved in salt. The preserved scorpions are usually boiled with herbs, such as licorice, to reduce toxicity after washing off the salt. Patients take the boiled extracts twice daily and the amount of scorpions should be controlled to under 10 grams for an adult daily. Most of proteins will be denatured and loss functions after being exposed to salt at a high concentration and water boiling temperature. The major effective composition of scorpion is scorpion venom. It has been demonstrated the venom can block cation channel of insects (Li et al., 2000;Wang et al., 2001). Several peptides containing CSαβ motif have been isolated from scorpion venom. The peptides can keep their structures and functions after be exposed to water boiling temperature and solvents of high salt concentration. BmK AEP, a peptide containing a CSαβ motif isolated from *Buthus martensii Karsch* venom, is the first anticonvulsant purified in the venom (Wang et al., 2001). BmK AEP has inhibitory effect to coriaria lactone-induced epilepsy model in rodents. It has a low toxicity to mice in doses up to 20 mg/animal (Rajendra et al., 2004). There is nearly no effect on heart rate, electroencephalogram, breathing rate and blood pressure (Rajendra et al., 2004;Wang et al., 2001). CII9, another peptide with a CSαβ motif isolated from *Centruroides limpidus Karsch*, can partially inhibit

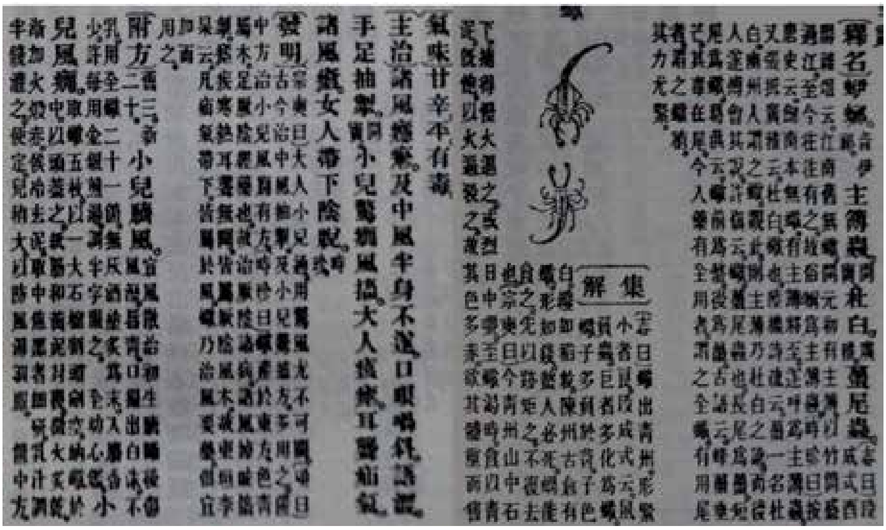


Fig. 4. The therapeutic instruction of scorpion described in the ancient Chinese pharmacopoeia: *Compendium of Materia Medica.*

sodium ion currents from superior cervical ganglion neurons of rats (Dehesa-Davila et al., 1996;Possani et al., 1999). In the rat model, CII9 also exerts a depressant effect on the behavior and electroencephalographic activity and antagonizes the epileptiform activity induced by penicillin (Dehesa-Davila et al., 1996;Gazarian et al., 2005;Possani et al., 1999).

3.3 Peptide antibiotics: Plectasin

Peptides with a CSαβ motif act as primary defenders for their host. Microbe-killing abilities of plant defensins have been demonstrated and transgenic industrial crops have been bred to carry specific defensins to increase their tolerance to plant pathogens. For directly binding to membranes or cell walls, target pathogens do not easily resist peptides with a CSαβ motif as they do to protein-targeting antibiotics. Pathogenic microbes would escape attacking from conventional antibiotics by mutating proteins but it is never easy to change the properties of cell membranes or cell walls. Scaffold of a CSαβ motif could be as a suitable platform for developing peptide antibiotics. Some plant defensins are tested and show activities againt human pathogenic microbes. For example, Rs-AFP2, a defensin isolated from seeds of *Raphanus sativus*, has an inhibitory activity against *Candida albicans* (Landon et al., 2004;Terras et al., 1992;Thomma et al., 2003). Scientists of Novozymes, a Dnaish biotech company, isolated a novel plant defensin, plectasin, with excellent antibiotics ability.

Plectasin is a 40-amino acid residue peptide isolated from saprophytic ascomycete, *Pseudoplectania nigrella*, and might be the first plant defensin with a CSαβ motif isolated from fungus (Mygind et al., 2005). Plectasin expresses anti-bacterial activities against a broad Gram-positive bacteria and has an inhibitory effect at a low concentration of 0.25 μg/ml to growth of *Streptococcus pneumoniae* (Mygind et al., 2005). Plectasin also has a comparable killing rate to *Streptococcus pneumoniae* as do conventional antibiotics (Mygind et al., 2005). *Streptococcus pneumoniae* is a major pathogenic bacteria and the most common cause of hospital/community-acquired pneumonia, bacterial meningitis, bacteremia, sinusitis, septic arthritis, osteomyelitis, peritonitis, and endocarditis (Whitney et al., 2000). Currently, antibiotics are major treatments to patients with infection of *Streptococcus pneumoniae*. In recent years, *Streptococcus pneumoniae* is more and more resistant to antibiotics and the demand for new drugs to cure *Streptococcus pneumoniae* infection is urgent (Baquero et al., 1991;Whitney et al., 2000). Another challenge of theraping *Streptococcus* infection is that *Streptococcus* can be an intracellular pathogen and avoids targeting by the immune system or drugs (Gordon et al., 2000;Talbot et al., 1996). Different from the conventional antibiotics, plectasin can directly act on the cell wall precursor lipid II of *Streptococcus* (Mygind et al., 2005). Studies also showed that plectasin has an intracellular activity against *Streptococcus aureus* both in human monocytes and in mouse peritonitis model without effecting the cells viability or inducing IL-8 production (Brinch et al., 2009;Hara et al., 2008).

4. Protein engineering based on CSαβ motif scaffold

Based on the scaffold, some CSαβ proteins have been engineered to exhibit new functions or changing of antimicrobial activities (Lin et al., 2007;Vita et al., 1999;Vita et al., 1995). Different approaches, including minimal residue substitution, functional epitope exchange, structural based modification and combinatorial chemistry have been employed to engineer the scaffold to exhibit new functions (Lin et al., 2007;Thevissen et al., 2007;Van Gaal et al., 2004;Vita et al., 1999;Vita et al., 1995;Zhao et al., 2004). Protein structure is important to the rational design of a protein scaffold. It does not only reveal the shape of a protein molecule but sidechain orientation of each residue. This information is crucial for investigating possible interactions between designed protein and its target.

4.1 Site-direct mutagenesis

In some studies, researchers focus on the relationships of structures, functions and each residue of peptides with a CSαβ motif. Extensive residue substitution is usually performed on the peptides and the changes of structures and functions are observed. The process is time-consuming and labor intensive but it is required for collecting basic information about the scaffold. Amino acid substitution has been extensively performed on two peptides with a CSαβ motif, VrD1 and brazzein (Assadi-Porter et al., 2010;Yang et al., 2009). In the two studies, molecular docking models are also established to investigate the interactions between receptors and ligands.

In both cases, similar positions along the structures are discovered that are crucial to functions of both peptides (Figure 5). These sites are widely distributed on β1 strand, loop 1, loop 2 and loop 3 (Assadi-Porter et al., 2000;Walters et al., 2009;Yang et al., 2009). From the molecular docking models, these sites either directly interact with their targets or play as functional epitopes and insert into the active site of their targets (Assadi-Porter et al., 2010;Liu et al., 2006;Yang et al., 2009). It is interesting that, when two negatively charged residues, D29 and E41, of brazzein are replaced by a positively charged amino acid, the sweetness of brazzein is greatly improved (Assadi-Porter et al., 2010). The mutants should be with high interests to industrial utilities. In the case of VrD1, there is no functional improvement observed on its mutants (Yang et al., 2009). Comparing conformations of the wild-type proteins and their mutations, there are only minimal shifts measured (Assadi-Porter et al., 2010). This implies that structure of peptide with a CSαβ motif is relatively stable and has high tolerance to amino acid substitution. Side chains of these residues on the interactive surface are crucial to biological functions of the peptides.

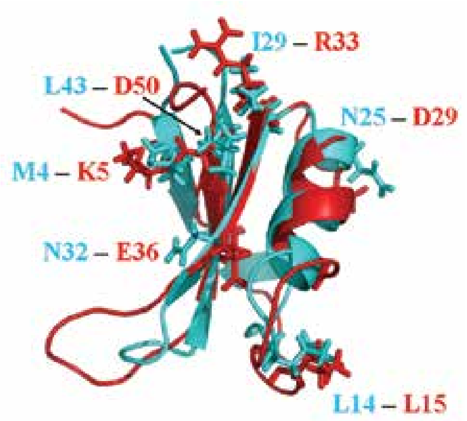


Fig. 5. Superposition of brazzein and VrD1. Some critical residues are overlapping between brazzein and VrD1. Red: brazzein, cyan: VrD1. Red and cyan labeled residues: critical residues of brazzein and VrD1; respectively. The two structures are aligned with SARST (http://sarst.life.nthu.edu.tw/iSARST/).

Multiple mutation also has been performed on the scaffold. Vital *et al* performed minimal residue substitutions on the charybodotxin (Chtx), an scorpion toxin containing CSαβ motif, to equip metal ion binding ability (Vita et al., 1995). Three residues, K27, M28 and R34, on the β sheet are substituted with histidines. The modified protein exhibts a chelatine property but has the same circular dichroism spectrum profile as the native Chtx does.

4.2 Functional epitope exchange

Grafting epitopes with known function is a straightforward stratage for a protein to gain new function. Functional epitopes could be exchanged between proteins have a high sequence identity or share a high structural similarity. Based on the structural homologous, the β2-β3 hairpin of scyllatoxin, a scorpion toxin containing the CSαβ motif, is replaced by the CDR2-like loop of human CD4. The chimeric protein can bind to the HIV-1 envelope glycoprotein and the affinity is increased 100 fold as compared with the native scyllatoxin (Vita et al., 1999). In another case, VrD1 and VrD2, two defensins of *Vigna radiate*, share 80% sequence identities. The major difference of sequences concentrats on the loop 3. When the loop 3 of VrD1 is grafted to VrD2, the chimeric peptide exhibits the enzyme inhibitory function as VrD1 does (Lin et al., 2007).

4.3 Combinatorial chemistry approach

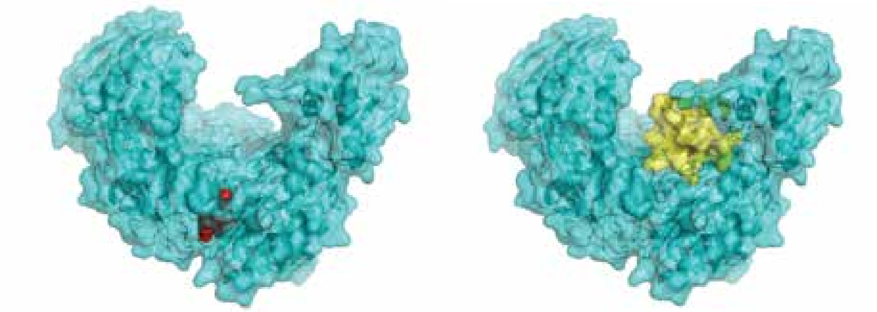
Combinatorial chemistry approach are powerful in screening and selecting binders with high affinity and high specificity (Hosse et al., 2006). It has been widely applied in antibody scaffold and thousands of antibodies are generated through the technology. Combinatorial chemistry approach also has been recruited to develop and isolate artificial proteins with new functions (Zhao et al., 2004). The approach could accelerate protein engineering based on a CSαβ motif to develop novel peptides with biomedical interesting (Thevissen et al., 2007;Van Gaal et al., 2004). Based on scaffold of insect defensin A, an expression library of peptides with 29 residues is constructed and used in screeening novel binders to targets. The expression library is artificially synthesized and amino acid of seven positions on the loops are randomized (Zhao et al., 2004). Tumor necrosis factor α (TNF-α), TNF receptor 1, TNF receptor 2 and antibody against BMP-2 are selected as targets and the screening results show significant enrichment in all cases.

5. The challenges of engineering of CSαβ motif

Although several successful engineered peptides based on the CSαβ motif scaffold have been reported, to design new functions into the scaffold does perplex us. In this section, we would like to discuss challenges in engineering a CSαβ motif to equip desired functions.

5.1 Lacking co-structural model

To introduce tailored functions into a protein scaffold is never an easy task. For its ultra stability and high diversity of function, there is no doubt that peptide with a CSαβ motif is a suitable scaffold to be engineered for biomedical applications. To date, there are only a few of successful cases are reported. Many structures of peptides with CSαβ motif have been solved and deposited in database, but knowledge about the motif is poor. In spite of hundreds of peptides with a CSαβ motif deposited in the database, it is hard to make conclusions relating to structures and biochemical functions of the motif. Key residues are revealed in some peptides with a CSαβ motif and molecular docking models provide reasonable explainations (Figure 6). The real interaction is not clarified, for lacking of a co-structural model. A co-structural model of the motif and its target will provide information about the dynamic interactions of the two protein molecules (Dumas et al., 2004;Thioulouse & Lobry, 1995). Structures of the complexes of peptides with a CSαβ motif and its counterparts will provide the situation of protein-protein interaction and have a great benefit to engineering of the scaffold.



(a) HIV reverse trascriptase with inhibitor (b) Docking model

Fig. 6. Docking model of plant defensin to HIV reverse transcriptase. Structures of HIV reverse transcriptase (3DRP, HIV RTase) and VrD2 (2GL1) are retrieved from the Protein data bank. Molecular docking is performed with PatchDock (http://bioinfo3d.cs.tau.ac.il/PatchDock/). (a) The crystal structure of HIV RTase with inhibitor. (b) Molecular docking model of VrD2 and HIV RTase. Cyan: HIV RTase, yellow: VrD2 and red: HIV RTase inhibitor.

5.2 Generation of gene library and construction of expression system

The advantages of combinatorial chemistry approach to protein engineering is its ability to associate every protein with its genetic material (Weng & DeLisi, 2002). There are two core parts about applying combinatorial chemistry approach in protein engineering: display techniques and gene expression library. These two core parts are equally important and affect each other. To date, several display techniques have been developed (Table 4) (Daugherty, 2007;Gronwall & Stahl, 2009). Natural properties of peptides with a CSαβ motif are as defenders for their native hosts and most of the peptides express anti-bacterial abilities, virus inhibitory abilities and inhibition of protein translation. How to design a expression library based on a CSαβ motif scaffold is a tough task.

Technique	Library limitation	Expression system	Disulphide bond
Bacterial display	10^{11}	Circular	No
Phage display	10^{11}	Circular	No
Ribosome display	10^{14}	Circular/Linear	Yes

Table 4. Comparison of three display techniques.

Peptides with a CSαβ motif are tightly held by disulphide bridges and may be not easily folded into appropriate structures inside bacterial cells (Villemagne et al., 2006). Ribosomal display technique can overcome the problems (Figure 7), but it requirs an ultra clean

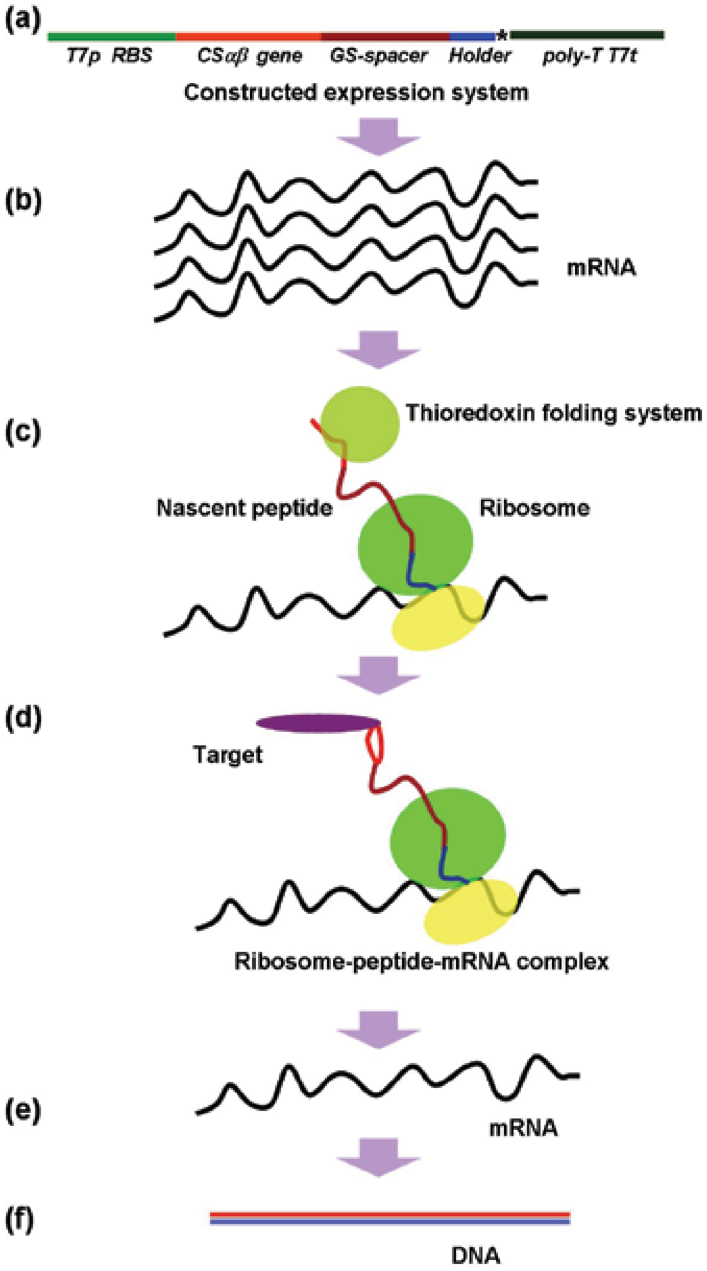

(a)
T7p RBS
CSαβ gene
GS-spacer
Holder
poly-T T7t
Constructed expression system
(b)
mRNA
(c)
Thioredoxin folding system
Nascent peptide
Ribosome
(d)
Target
Ribosome-peptide-mRNA complex
(e)
mRNA
(f)
DNA

Fig. 7. Concept of ribosome display techniques. (a) expression system, (b) & (c) transcription, translation and folding while translation, (d) binding and selection, (e) isolation of mRNA from the complex, (f) reverse transcription and construction of expression system for next cycle of selection. A specific holder gene usually is constructed to the expression system. The holder can arrest protein translation and the nascent peptide can form a stable complex with ribomsome.

environment to protect mRNA from degradation (Fennell et al., 2010;Villemagne et al., 2006). The gene library of a CSαβ motif has to be artificially synthesized then constructed to the expression system. There are two key issues of the synthesized gene library have to be considered, one is the size of library the other one is how many random residue positions it should have (Gronwall & Stahl, 2009;Michel-Reydellet et al., 2005). These two questions are hard to answer. A large library with more randomized residue positions could have more genotypes/phenotypes but it also has a high possibility to generate nonsense genes for the unvoided stop codon on unexpected positions. A library with a randomized loop 3 of a CSαβ motif scaffold has been constructed and successfully used in a noval binder to human TNF-α (Zhao et al., 2004). It also has been reported that all three loops of peptides with a CSαβ motif can equip unique functions and the structural regions also fine regulate the functions of the peptides (Assadi-Porter et al., 2010;Liu et al., 2006;Yang et al., 2009;Zhao et al., 2002). For a library of long length artificial genes, several cycles of PCR overlapping extensions are required and may reduce quality of the library. A circular or a linear expression systems can be recruited (Fennell et al., 2010;Shimizu et al., 2001;Villemagne et al., 2006). To construct the linear library, all required DNA elements, as promotor, ribosomal binding site and poly T tail, are synthesized or copied and constructed with the desired gene library through PCR (Katzen et al., 2005;Shimizu et al., 2001) . In the linear expression system, the library can have a maximal size up to 10^{13} (Table 4) (Gronwall & Stahl, 2009).When a linear expression system is recruited, the qualities of final PCR products should be well controlled. It is better for the library to be freshly constructed everytime the protein display is performed. The size of a library constructed with a circular expression system is much smaller for the genotype lost during construction, however, the library can be stably stored in a good condition after construction.

5.3 Mass production and folding of proteins

To screening possible funtions of a protein, it requires a large amount of the protein. The amount of proteins with a CSαβ motif in their native host is rare and to obtain eounght of the native proteins for assays requires a lot of raw materials. To produce recombinant proteins of a CSαβ motif in bacteria cells, the proteins have to be fused with a large tag to reduce their toxicity to bacteria. Protein folding is also another problem for apporpriately forming the disulphide bonds of a CSαβ motif. To overcome the problems, peptides with CSαβ motif can be fused with a thioredoxin tag and a cleavagable site is placed between the peptide and the tag. The recombinant proteins are expressed in *E. Coli* and purified. After the thioredoxin tag is cut off, the desired proteins are further purified. A simple method to cut off the thioredoxin tag is to place a acidohydrolysis site, -Asp-Pro-, between the peptide and the tag (Liu et al., 2006). A large amount of reconbinant proteins can be obtained by following the procedures and the native properties and functions of the proteins are preserved. The acidohydrolysis process is not suitable for proteins with a CSαβ motif containing proline for the proteins would be unstable in a heat and acidic environment (Cunningham & O'Connor, 1997;LeBlanc & London, 1997;Wilce et al., 1998). To resolve the problem, an alternative enzyme cleavage site can be introduced to replace the acidohydrolysis site.

5.4 Unpredictable biochemical function

Enzyme inhibitory function and microbial killing ability are the two greatest advantages of peptide with a CSαβ motif. The conventional bacterial/fungal killing assay is time consuming and it is possible for a performer to be exposed to pathogens during the operation. The peptides with a CSαβ motif have been shown to form pores on an artificial membrane, but the artificial membrane does not represent the membrane of different organisms. A new high throughput screening recruiting membrane of different pathogens will be a great help in resolving the problem (Yang & Lyu, 2008).

Not simple as binding assay, enzymes act on specific substrates and it is extremely difficult to unify experimental conditions for a broad spectrum of enzymes. *In vitro* enzyme inhibitory assays are costly and unable to represent the actually physiological conditions. Computer aided drug screening methods have been established to screen drugs with small molecule weights that can fit into active sites of target proteins (Barrons, 2004;Weideman et al., 1999). Drug dynamics also can be *in silico* simulated (Sinek et al., 2009;Zunino et al., 2009). An antibody can be as a simple binder and its binding targets usually are short fragments exposed to the surface of the proteins. The binding fragments are predicable for a huge amount of knowledge accumulated in the last several decades (Blythe & Flower, 2005;Kulkarni-Kale et al., ;Odorico & Pellequer, 2003). Currently, we do not well understand the interface between peptides with a CSαβ motif and their targets. How functional loops access active sites of the targets is needed to be decrypted. We know too little about the scaffold and it is to difficult to engineer it. It is hard to develop an algorithm to predict the biochemical functions of the peptides with a CSαβ motif. For the complexity of protein-protein interaction, it is a long way to screen peptide drugs *in silico*.

6. Business issue

To biopharma companies, protein engineering/evolution platforms must provide solutions to improve the throughput, safety, biodistribution, activity and frequency of dosing of designed protein, evading intellectual property issues and reducing manufacturing costs, especially in the late process of development and production stages. Intellectual property rights are definite issues, the intellectual property situations of antibodies are obscure as previously described. An alternative scaffold equips the same abilities as antibody does is a valuable prospect and desired to avoid the complex intellectual property landscape of antibody. A protein scaffold can access to a cardinal target and circumvent intellectual property issues that does not mean that it is necessarily a platform worth investing in. The successfully designed protein must equip an improved performance and provide solutions to the concerned points described above.

7. Conclusion

The knowledge about the scaffold is not enough to establish rules to predict its biochemical functions. There are still some technical limitations and the development of the scaffold is confined. Peptides with a CSαβ motif bear many excellent properties and can serve as an alternative protein scaffold for biomedical application. Natural products containing peptides with a CSαβ motif have been therapeutically used for hundreds of years. Some of the purified peptides are tested and have low toxicities to mammalian species. Prudential is needed when recruiting new proteins/peptides as treatments for diseases. The intellectual property situation of peptides with a CSαβ motif is not as complex as antibody, especially in biomedical application (Yang & Lyu, 2008). For its original expression in plant, the engineered proteins with medicinal effect can be expressed in transgenic crops and diseases can be controlled through foods. To predict the biochemical functions of peptides with a CSαβ motif is difficult, some prediction tools have been developed to predict functions of protein from the primary structure, but how proteins interact with their counterparts has to be considered. Folding and post-translational modification of proteins has to be thought over. We believe that peptides with a CSαβ motif will be an alternative scaffold and can be widely applied in biomedical utilities.

8. Acknowledgment

We sincerely thank Professor Ping-Chiang Lyu at the College of Life Science, National Tsing-Hua University for his kind and generous suggestions. We also thank Dr. Chung-Chang Liu, the senior expert of Industrial Technology Research Institute (ITRI), for his supporting part of the research of this work during his term as the director of Biomedical Engineering Research Laboratories, ITRI.

9. References

Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S. M., Roodveldt, C.&Tawfik, D. S. (2005) The 'evolvability' of promiscuous protein functions *Nature Genetics* Vol. 37 No. 1 pp. 73-76

Alahuhta, M., Salin, M., Casteleijn, M. G., Kemmer, C., El-Sayed, I., Augustyns, K., Neubauer, P.&Wierenga, R. K. (2008) Structure-based protein engineering efforts with a monomeric TIM variant: the importance of a single point mutation for generating an active site with suitable binding properties *Protein Engineering, Design and Selection* Vol. 21 No. 4 pp. 257-266

Angeletti, R. H. (1998) *Proteins: Analysis and Design*, Academic Press, 0120587858 New York

Assadi-Porter, F., Tonelli, M., Radek James, T., Cornilescu Claudia, C.&Markley John, L. (2008) How Sweet It Is: Detailed Molecular and Functional Studies of Brazzein, a Sweet Protein and Its Analogs. In: *Sweetness and Sweeteners*, pp. (560-572), American Chemical Society, 9780841274327, New York

Assadi-Porter, F. M., Aceti, D. J.&Markley, J. L. (2000) Sweetness Determinant Sites of Brazzein, a Small, Heat-Stable, Sweet-Tasting Protein *Archives of Biochemistry and Biophysics* Vol. 376 No. 2 pp. 259-265

Assadi-Porter, F. M., Maillet, E. L., Radek, J. T., Quijada, J., Markley, J. L.&Max, M. (2010) Key Amino Acid Residues Involved in Multi-Point Binding Interactions between Brazzein, a Sweet Protein, and the T1R2-T1R3 Human Sweet Receptor *Journal of Molecular Biology* Vol. 398 No. 4 pp. 584-599

Baines, I. C.&Colas, P. (2006) Peptide aptamers as guides for small-molecule drug discovery *Drug Discovery Today* Vol. 11 No. 7-8 pp. 334-341

Baquero, F., Beltran, J. M.&Loza, E. (1991) A review of antibiotic resistance patterns of Streptococcus pneumoniae in Europe *Journal of Antimicrobial Chemotherapy* Vol. 28 No. Suppl C pp. 31-38

Barrons, R. (2004) Evaluation of personal digital assistant software for drug interactions *American Journal of Health-System Pharmacy* Vol. 61 No. 4 pp. 380-385

Blythe, M. J.&Flower, D. R. (2005) Benchmarking B cell epitope prediction: Underperformance of existing methods *Protein Science* Vol. 14 No. 1 pp. 246-248

Bottcher, D.&Bornscheuer, U. T. (2010) Protein engineering of microbial enzymes *Current Opinion in Microbiology* Vol. 13 No. 3 pp. 274-282

Brinch, K. S., Sandberg, A., Baudoux, P., Van Bambeke, F., Tulkens, P. M., Frimodt-Moller, N., Hoiby, N.&Kristensen, H.-H. (2009) Plectasin Shows Intracellular Activity against Staphylococcus aureus in Human THP-1 Monocytes and in a Mouse Peritonitis Model *Antimicrob Agents Chemother.* Vol. 53 No. 11 pp. 4801-4808

Caldwell, J. E., Abildgaard, F., Dzakula, Z., Ming, D., Hellekant, G.&Markley, J. L. (1998) Solution structure of the thermostable sweet-tasting protein brazzein *Nature Structural & Molecular Biology* Vol. 5 No. 6 pp. 427-431

Carvalho, A. d. O.&Gomes, V. M. (2009) Plant defensins--Prospects for the biological functions and biotechnological properties *Peptides* Vol. 30 No. 5 pp. 1007-1020

Chan, H. S., Bromberg, S.&Dill, K. A. (1995) Models of Cooperativity in Protein Folding *Philosophical Transactions of the Royal Society B: Biological Sciences* Vol. 348 No. 1323 pp. 61-70

Chen, G.-H., Hsu, M.-P., Tan, C.-H., Sung, H.-Y., Kuo, C. G., Fan, M.-J., Chen, H.-M., Chen, S.&Chen, C.-S. (2005) Cloning and Characterization of a Plant Defensin VaD1 from Azuki Bean *Journal of Agricultural and Food Chemistry* Vol. 53 No. 4 pp. 982-988

Clauss, M. J.&Mitchell-Olds, T. (2004) Functional Divergence in Tandemly Duplicated Arabidopsis thaliana Trypsin Inhibitor Genes *Genetics* Vol. 166 No. 3 pp. 1419-1436

Corzo, G., Sabo, J. K., Bosmans, F., Billen, B., Villegas, E., Tytgat, J.&Norton, R. S. (2007) Solution Structure and Alanine Scan of a Spider Toxin That Affects the Activation of Mammalian Voltage-gated Sodium Channels *Journal of Biological Chemistry* Vol. 282 No. 7 pp. 4643-4652

Cunningham, D. F.&O'Connor, B. (1997) Proline specific peptidases *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* Vol. 1343 No. 2 pp. 160-186

Das, A., Wei, Y., Pelczer, I.&Hecht, M. H. (2011) Binding of small molecules to cavity forming mutants of a de novo designed protein *Protein Science* Vol. 20 No. 4 pp. 702-711

Daugherty, P. S. (2007) Protein engineering with bacterial display *Current Opinion in Structural Biology* Vol. 17 No. 4 pp. 474-480

Dehesa-Davila, M., Ramfrez, A. N., Zamudio, F. Z., Gurrola-Briones, G., Lievano, A., Darszon, A.&Possani, L. D. (1996) Structural and functional comparison of toxins from the venom of the scorpions Centruroides infamatus infamatus, centruroides limpidus limpidus and Centruroides noxius *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* Vol. 113 No. 2 pp. 331-339

Dumas, J. J., Kumar, R., McDonagh, T., Sullivan, F., Stahl, M. L., Somers, W. S.&Mosyak, L. (2004) Crystal Structure of the Wild-type von Willebrand Factor A1-Glycoprotein Ib Complex Reveals Conformation Differences with a Complex Bearing von Willebrand Disease Mutations *Journal of Biological Chemistry* Vol. 279 No. 22 pp. 23327-23334

Ehren, J., Govindarajan, S., Moron, B., Minshull, J.&Khosla, C. (2008) Protein engineering of improved prolyl endopeptidases for celiac sprue therapy *Protein Engineering, Design & Selection* Vol. 21 No. 12 pp. 699-707

Fant, F., Vranken, W. F.&Borremans, F. A. M. (1999) The three-dimensional solution structure of *Aesculus hippocastanum* antimicrobial protein 1 determined by ^{1}H nuclear magnetic resonance. *Proteins: Structure, Function, and Genetics* Vol. 37 No. 3 pp. 388-403

Fennell, B. J., Darmanin-Sheehan, A., Hufton, S. E., Calabro, V., Wu, L., Miller, M. R., Cao, W., Gill, D., Cunningham, O.&Finlay, W. J. J. (2010) Dissection of the IgNAR V Domain: Molecular Scanning and Orthologue Database Mining Define Novel IgNAR Hallmarks and Affinity Maturation Mechanisms *Journal of Molecular Biology* Vol. 400 No. 2 pp. 155-170

Fisher, M. A., McKinley, K. L., Bradley, L. H., Viola, S. R.&Hecht, M. H. (2011) *De Novo* Designed Proteins from a Library of Artificial Sequences Function in *Escherichia Coli* and Enable Cell Growth *PLoS ONE* Vol. 6 No. 1 pp. e15364

Gazarian, K. G., Gazarian, T., Hernodez, R.&Possani, L. D. (2005) Immunology of scorpion toxins and perspectives for generation of anti-venom vaccines *Vaccine* Vol. 23 No. 26 pp. 3357-3368

Gordon, S. B., Irving, G. R. B., Lawson, R. A., Lee, M. E.&Read, R. C. (2000) Intracellular Trafficking and Killing of Streptococcus pneumoniae by Human Alveolar Macrophages Are Influenced by Opsonins *Infection and Immunity* Vol. 68 No. 4 pp. 2286-2293

Gronwall, C.&Stahl, S. (2009) Engineered affinity proteins--Generation and applications *Journal of Biotechnology* Vol. 140 No. 3-4 pp. 254-269

Hara, S., Mukae, H., Sakamoto, N., Ishimoto, H., Amenomori, M., Fujita, H., Ishimatsu, Y., Yanagihara, K.&Kohno, S. (2008) Plectasin has antibacterial activity and no affect on cell viability or IL-8 production *Biochemical and Biophysical Research Communications* Vol. 374 No. 4 pp. 709-713

Hey, T., Fiedler, E., Rudolph, R.&Fiedler, M. (2005) Artificial, non-antibody binding proteins for pharmaceutical and industrial applications *Trends in Biotechnology* Vol. 23 No. 10 pp. 514-522

Holt, L. J., Herring, C., Jespers, L. S., Woolven, B. P.&Tomlinson, I. M. (2003) Domain antibodies: proteins for therapy *Trends in Biotechnology* Vol. 21 No. 11 pp. 484-490

Hosse, R. J., Rothe, A.&Power, B. E. (2006) A new generation of protein display scaffolds for molecular recognition *Protein Science* Vol. 15 No. 1 pp. 14-27

Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotna, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E.&Crea, R. (1988) Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli *Proceedings of the National Academy of Sciences of the United States of America* Vol. 85 No. 16 pp. 5879-5883

Jacobs, P. P.&Callewaert, N. (2009) N-glycosylation Engineering of Biopharmaceutical Expression Systems *Current Molecular Medicine* Vol. 9 No. 7 pp. 774-800

Jin, Z., Markley, J. L., Assadi-porter, F. M.&Hellekant, B. G. (2006) Protein sweetener. In., Wisconsin, Alumni Res Found (US), United States

Jones, D. S., Silverman, A. P.&Cochran, J. R. (2008) Developing therapeutic proteins by engineering ligand-receptor interactions *Trends in Biotechnology* Vol. 26 No. 9 pp. 498-505

Katzen, F., Chang, G.&Kudlicki, W. (2005) The past, present and future of cell-free protein synthesis *Trends in Biotechnology* Vol. 23 No. 3 pp. 150-156

Kosloski, M., Miclea, R.&Balu-Iyer, S. (2009) Role of Glycosylation in Conformational Stability, Activity, Macromolecular Interaction and Immunogenicity of Recombinant Human Factor VIII *The AAPS Journal* Vol. 11 No. 3 pp. 424-431

Kristensen, C., Kjeldsen, T., Wiberg, F. C., Schaffer, L., Hach, M., Havelund, S., Bass, J., Steiner, D. F.&Andersen, A. S. (1997) Alanine Scanning Mutagenesis of Insulin *Journal of Biological Chemistry* Vol. 272 No. 20 pp. 12978-12983

Kulkarni-Kale, U., Bhosle, S.&Kolaskar, A. S. (2005) CEP: a conformational epitope prediction server *Nucleic Acids Research* Vol. 33 No. suppl 2 pp. W168-W171

Landon, C., Barbault, F., Legrain, M., Menin, L., Guenneugues, M., Schott, V., Vovelle, F.&Dimarcq, J.-L. (2004) Lead optimization of antifungal peptides with 3D NMR structures analysis *Protein Science* Vol. 13 No. 3 pp. 703-713

Lay, F. T.&Anderson, M. A. (2005) Defensins - Components of the Innate Immune System in Plants *Current Protein & Peptide Science* Vol. 6 No. 1 pp. 85-101

LeBlanc, D. A.&London, R. E. (1997) Cleavage of the X-Pro Peptide Bond by Pepsin Is Specific for the trans Isomer? *Biochemistry* Vol. 36 No. 43 pp. 13232-13240

Leta Aboye, T., Clark, R. J., Craik, D. J.&Göransson, U. (2008) Ultra-Stable Peptide Scaffolds for Protein Engineering—Synthesis and Folding of the Circular Cystine Knotted Cyclotide Cycloviolacin O2 *ChemBioChem* Vol. 9 No. 1 pp. 103-113

Li, H.&d'Anjou, M. (2009) Pharmacological significance of glycosylation in therapeutic proteins *Current Opinion in Biotechnology* Vol. 20 No. 6 pp. 678-684

Li, Y.-J., Tan, Z.-Y.&Ji, Y.-H. (2000) The binding of BmK IT2, a depressant insect-selective scorpion toxin on mammal and insect sodium channels *Neuroscience Research* Vol. 38 No. 3 pp. 257-264

Lin, K. F., Lee, T. R., Tsai, P. H., Hsu, M. P., Chen, C. S.&Lyu, P. C. (2007) Structure-based protein engineering for α-amylase inhibitory activity of plant defensin *Proteins: Structure, Function, and Bioinformatics* Vol. 68 No. 2 pp. 530-540

Liu, Y. J., Cheng, C. S., Lai, S. M., Hsu, M. P., Chen, C. S.&Lyu, P. C. (2006) Solution structure of the plant defensin VrD1 from mung bean and its possible role in insecticidal activity against bruchids *Proteins: Structure, Function, and Bioinformatics* Vol. 63 No. 4 pp. 777-786 1097-0134

Lobo, D. S., Pereira, I. B., Fragel-Madeira, L., Medeiros, L. N., Cabral, L. M., Faria, J., Bellio, M., Campos, R. C., Linden, R.&Kurtenbach, E. (2007) Antifungal Pisum sativum Defensin 1 Interacts with Neurospora crassa Cyclin F Related to the Cell Cycle *Biochemistry* Vol. 46 No. 4 pp. 987-996

Malavasic, M., Poklar, N., Macek, P.&Vesnaver, G. (1996) Fluorescence studies of the effect of pH, guanidine hydrochloride and urea on equinatoxin II conformation *Biochimica et Biophysica Acta (BBA) - Biomembranes* Vol. 1280 No. 1 pp. 65-72

Michel-Reydellet, N., Woodrow, K.&Swartz, J. (2005) Increasing PCR Fragment Stability and Protein Yields in a Cell-Free System with Genetically Modified Escherichia coli Extracts *Journal of Molecular Microbiology and Biotechnology* Vol. 9 No. 1 pp. 26-34

Ming, D.&Hellekant, G. (1994) Brazzein, a new high-potency thermostable sweet protein from Pentadiplandra brazzeana B *FEBS Letters* Vol. 355 No. 1 pp. 106-108

Muir, E. M., Fyfe, I., Gardiner, S., Li, L., Warren, P., Fawcett, J. W., Keynes, R. J.&Rogers, J. H. (2009) Modification of N-glycosylation sites allows secretion of bacterial chondroitinase ABC from mammalian cells *Journal of Biotechnology* Vol. 145 No. 2 pp. 103-110

Mygind, P. H., Fischer, R. L., Schnorr, K. M., Hansen, M. T., Sonksen, C. P., Ludvigsen, S., Raventos, D., Buskov, S., Christensen, B., De Maria, L., Taboureau, O., Yaver, D., Elvig-Jorgensen, S. G., Sorensen, M. V., Christensen, B. E., Kjaerulff, S., Frimodt-Moller, N., Lehrer, R. I., Zasloff, M.&Kristensen, H.-H. (2005) Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus *Nature* Vol. 437 No. 7061 pp. 975-980

Nikkhah, M., Jawad-Alami, Z., Demydchuk, M., Ribbons, D.&Paoli, M. (2006) Engineering of β-propeller protein scaffolds by multiple gene duplication and fusion of an idealized WD repeat *Biomolecular Engineering* Vol. 23 No. 4 pp. 185-194

Odorico, M.&Pellequer, J.-L. (2003) BEPITOPE: predicting the location of continuous epitopes and patterns in proteins *Journal of Molecular Recognition* Vol. 16 No. 1 pp. 20-22

Perego, P., Gatti, L.&Beretta, G. L. (2010) The ABC of glycosylation *Nature Reviews Cancer* Vol. 10 No. 7 pp. 523-523

Pessi, A., Bianchi, E., Crameri, A., Venturini, S., Tramontano, A.&Sollazzo, M. (1993) A designed metal-binding protein with a novel fold *Nature* Vol. 362 No. 6418 pp. 367-369

Possani, L. D., Becerril, B., Delepierre, M.&Tytgat, J. (1999) Scorpion toxins specific for Na+-channels *European Journal of Biochemistry* Vol. 264 No. 2 pp. 287-300

Rajendra, W., Armugam, A.&Jeyaseelan, K. (2004) Neuroprotection and peptide toxins *Brain Research Reviews* Vol. 45 No. 2 pp. 125-141

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K.&Ueda, T. (2001) Cell-free translation reconstituted with purified components *Nature Biotechnology* Vol. 19 No. 8 pp. 751-755

Sinek, J., Sanga, S., Zheng, X., Frieboes, H., Ferrari, M.&Cristini, V. (2009) Predicting drug pharmacokinetics and effect in vascularized tumors using computer simulation *Journal of Mathematical Biology* Vol. 58 No. 4 pp. 485-510

Skerra, A. (2007) Alternative non-antibody scaffolds for molecular recognition *Current Opinion in Biotechnology* Vol. 18 No. 4 pp. 295-304

Song, X., Wang, J., Wu, F., Li, X., Teng, M.&Gong, W. (2005) cDNA cloning, functional expression and antifungal activities of a dimeric plant defensin SPE10 from Pachyrrhizus erosus seeds *Plant Molecular Biology* Vol. 57 No. 1 pp. 13-20

Spelbrink, R. G., Dilmac, N., Allen, A., Smith, T. J., Shah, D. M.&Hockerman, G. H. (2004) Differential Antifungal and Calcium Channel-Blocking Activity among Structurally Related Plant Defensins *Plant Physiol.* Vol. 135 No. 4 pp. 2055-2067

Stec., B. (2006) Plant thionins-the structural perspective *Cellular and Molecular Life Sciences (CMLS)* Vol. 63 No. 12 pp. 1370-1385

Sun, Y. M., Liu, W., Zhu, R. H., Wang, D. C., Goudet, C.&Tytgat, J. (2002) Roles of disulfide bridges in scorpion toxin BmK M1 analyzed by mutagenesis *Journal of Peptide Research* Vol. 60 No. 5 pp. 247-256

Talbot, U. M., Paton, A. W.&Paton, J. C. (1996) Uptake of Streptococcus pneumoniae by respiratory epithelial cells *Infection and Immunity* Vol. 64 No. 9 pp. 3772-3777

Terras, F. R., Schoofs, H. M., De Bolle, M. F., Van Leuven, F., Rees, S. B., Vanderleyden, J., Cammue, B. P.&Broekaert, W. F. (1992) Analysis of two novel classes of plant antifungal proteins from radish (*Raphanus sativus L.*) seeds *Journal of Biological Chemistry* Vol. 267 No. 22 pp. 15301-15309

Thevissen, K., Ghazi, A., De Samblanx, G. W., Brownlee, C., Osborn, R. W.&Broekaert, W. F. (1996) Fungal Membrane Responses Induced by Plant Defensins and Thionins *Journal of Biological Chemistry* Vol. 271 No. 25 pp. 15018-15025

Thevissen, K., Kristensen, H.-H., Thomma, B. P. H. J., Cammue, B. P. A.&Francois, I. E. J. A. (2007) Therapeutic potential of antifungal plant and insect defensins *Drug Discovery Today* Vol. 12 No. 21-22 pp. 966-971

Thioulouse, J.&Lobry, J. R. (1995) Co-inertia analysis of amino-acid physico-chemical properties and protein composition with the ADE package *Computer Applications In The Biosciences* Vol. 11 No. 3 pp. 321-329

Thomma, B. P. H. J., Cammue, B. P. A.&Thevissen, K. (2003) Mode of Action of Plant Defensins Suggests Therapeutic Potential *Current Drug Targets - Infectious Disorsers* Vol. 3 No. pp. 1-8

Van Gaal, L., Mertens, I., Ballaux, D.&Verkade, H. J. (2004) Modern, new pharmacotherapy for obesity. A gastrointestinal approach *Best Practice & Research Clinical Gastroenterology* Vol. 18 No. 6 pp. 1049-1072

Van Walle, I., Gansemans, Y., Parren, P. W. H. I., Stas, P.&Lasters, I. (2007) Immunogenicity screening in protein drug development *Expert Opinion on Biological Therapy* Vol. 7 No. 3 pp. 405-418

Villemagne, D., Jackson, R.&Douthwaite, J. A. (2006) Highly efficient ribosome display selection by use of purified components for in vitro translation *Journal of Immunological Methods* Vol. 313 No. 1-2 pp. 140-148

Vita, C., Drakopoulou, E., Vizzavona, J., Rochette, S., Martin, L., Menez, A., Roumestand, C., Yang, Y.-S., Ylisastigui, L., Benjouad, A.&Gluckman, J. C. (1999) Rational engineering of a miniprotein that reproduces the core of the CD4 site interacting with HIV-1 envelope glycoprotein *Proceedings of the National Academy of Sciences of the United States of America* Vol. 96 No. 23 pp. 13091-13096

Vita, C., Roumestand, C., Toma, F.&Menez, A. (1995) Scorpion toxins as natural scaffolds for protein engineering *Proceedings of the National Academy of Sciences of the United States of America* Vol. 92 No. 14 pp. 6404-6408

Walters, D. E., Cragin, T., Jin, Z., Rumbley, J. N.&Hellekant, G. (2009) Design and Evaluation of New Analogs of the Sweet Protein Brazzein *Chemical Senses* Vol. 34 No. 8 pp. 679-683

Wang, C.-G., He, X.-L., Shao, F., Liu, W., Ling, M.-H., Wang, D.-C.&Chi, C.-W. (2001) Molecular characterization of an anti-epilepsy peptide from the scorpion Buthus martensi Karsch *European Journal of Biochemistry* Vol. 268 No. 8 pp. 2480-2485

Wang, L. K., Schwer, B.&Shuman, S. (2006) Structure-guided mutational analysis of T4 RNA ligase 1 *RNA* Vol. 12 No. 12 pp. 2126-2134

Wang, W., Lu, B., Zhou, H., Suguitan, A. L., Jr., Cheng, X., Subbarao, K., Kemble, G.&Jin, H. (2010) The Hemagglutin protein 158N glycosylation and receptor binding specificity synergistically affect antigenicity and immunogenicity of a live attenuated H5N1 A/Vietnam/1203/2004 vaccine virus in ferrets *Journal of Virology* Vol. 84 No. 13 pp. 6570-6577

Weideman, R. A., Bernstein, I. H.&McKinney, W. P. (1999) Pharmacist recognition of potential drug interactions *American Journal of Health-System Pharmacy* Vol. 56 No. 15 pp. 1524-1529

Weng, Z.&DeLisi, C. (2002) Protein therapeutics: promises and challenges for the 21st century *Trends in biotechnology* Vol. 20 No. 1 pp. 29-35

Whitney, C. G., Farley, M. M., Hadler, J., Harrison, L. H., Lexau, C., Reingold, A., Lefkowitz, L., Cieslak, P. R., Cetron, M., Zell, E. R., Jorgensen, J. H.&Anne Schuchat. (2000) Increasing Prevalence of Multidrug-Resistant Streptococcus pneumoniae in the United States *New England Journal of Medicine* Vol. 343 No. 26 pp. 1917-1924

Wijaya, R., Neumann, G. M., Condron, R., Hughes, A. B.&Polya, G. M. (2000) Defense proteins from seed of Cassia fistula include a lipid transfer protein homologue and a protease inhibitory plant defensin *Plant Science* Vol. 159 No. 2 pp. 243-255

Wilce, M. C. J., Bond, C. S., Dixon, N. E., Freeman, H. C., Guss, J. M., Lilley, P. E.&Wilce, J. A. (1998) Structure and mechanism of a proline-specific aminopeptidase from Escherichia coli *Proceedings of the National Academy of Sciences of the United States of America* Vol. 95 No. 7 pp. 3472-3477

Wong, J. H.&Ng, T. B. (2005) Sesquin, a potent defensin-like antimicrobial peptide from ground beans with inhibitory activities toward tumor cells and HIV-1 reverse transcriptase *Peptides* Vol. 26 No. 7 pp. 1120-1126

Wu, C., Zhang, X., Tian, Y., Song, J., Yang, D., Roggendorf, M., Lu, M.&Chen, X. (2010) Biological significance of amino acid substitutions in hepatitis B surface antigen (HBsAg) for glycosylation, secretion, antigenicity and immunogenicity of HBsAg and hepatitis B virus replication *Journal General Virology* Vol. 91 No. 2 pp. 483-492

Yang, Y.-F., Cheng, K.-C., Tsai, P.-H., Liu, C.-C., Lee, T.-R.&Lyu, P.-C. (2009) Alanine substitutions of noncysteine residues in the cysteine-stabilized αβ motif *Protein Science* Vol. 18 No. 7 pp. 1498-1506

Yang, Y.-F.&Lyu, P.-C. (2008) The Proteins of Plant Defensin Family and their Application Beyond Plant Disease Control *Recent Patents on DNA & Gene Sequences* Vol. 2 No. 3 pp. 214-218

Zasloff, M. (2002) Antimicrobial peptides of multicellular organisms *Nature* Vol. 415 No. 6870 pp. 389-395

Zhao, A., Xue, Y., Zhang, J., Gao, B., Feng, J., Mao, C., Zheng, L., Liu, N., Wang, F.&Wang, H. (2004) A conformation-constrained peptide library based on insect defensin A *Peptides* Vol. 25 No. 4 pp. 629-635

Zhao, Q., Chae, Y. K.&Markley, J. L. (2002) NMR Solution Structure of ATT$_{\mathrm{p}}$, an *Arabidopsis thaliana* Trypsin Inhibitor *Biochemistry* Vol. 41 No. 41 pp. 12284-12296

Zhong, L.&Johnson, W. C. (1992) Environment affects amino acid preference for secondary structure *Proceedings of the National Academy of Sciences of the United States of America* Vol. 89 No. 10 pp. 4462-4465

Zhu, Q., Liang, S., Martin, L., Gasparini, S., Menez, A.&Vita, C. (2002) Role of Disulfide Bonds in Folding and Activity of Leiurotoxin I: Just Two Disulfides Suffice *Biochemistry* Vol. 41 No. 38 pp. 11488-11494

Zhu, S., Gao, B.&Tytgat, J. (2005) Phylogenetic distribution, functional epitopes and evolution of the CSαβ superfamily *Cellular and Molecular Life Sciences* Vol. 62 No. 19 - 20 pp. 2257-2269

Zunino, P., D'Angelo, C., Petrini, L., Vergara, C., Capelli, C.&Migliavacca, F. (2009) Numerical simulation of drug eluting coronary stents: Mechanics, fluid dynamics and drug release *Computer Methods in Applied Mechanics and Engineering* Vol. 198 No. 45-46 pp. 3633-3644

27

Application of Liposomes for Construction of Vaccines

Jaroslav Turánek[1], Josef Mašek[1], Milan Raška[2] and Miroslav Ledvina[3]
[1]*Veterinary Research Institute,*
[2]*Palacky University,*
[3]*Institute of Organic Chemistry and Biochemistry,*
Czech Republic

1. Introduction

Vaccinology as a scientific field is undergoing a dramatic development. Never before such sophisticated techniques and in-depth knowledge of immunological processes have been at hand to exploit fully the potential of protecting from as well as curing diseases through vaccination. In spite of great successes like eradication of smallpox in the 1970s and poliomyelitis elimination from all but six countries in the world (two important milestones in the medical history), new challenges have arisen to be faced. Rapidly changing ecosystems and human behaviour, an ever-increasing density of human and farmed animal populations, a high degree of mobility resulting in rapid spreading of pathogens in infected people and animals, new contacts between human and animals in endemic areas, poverty, and war conflicts in the third world, and many other factors contribute to the more frequent occurrence and rapid dissemination of new diseases. Three diseases that most heavily afflict global health are AIDS, tuberculosis, and malaria. As an example of new viral pathogens we can mention Ebola virus, SARS-coronavirus, or new strains of influenza virus (Wack & Rappuoli, 2005).

Among re-emerging diseases of the past few years, diphteria and cholera should be mentioned. Moreover, multi drug-resistant bacteria frequently occur as a result of overdosing on antibiotics. One of the most important future challenges will be to respond promptly to emerging diseases such as those mentioned above. Rapid sequencing of the genome of the pathogen implicated the speed of the development of diagnostic tools as well as the identification and expression of recombinant targets for vaccines and therapeutic agents development (Stadler et al., 2003). Immunotherapy of cancer represents a special field, where anticancer vaccines could be powerful weapons/tools for long-term effective treatment.

The progress in the vaccine development is closely related with the progress in immunology and molecular biology. A new term "Reverse vaccinology" was proposed by Rappuoli (Rappuoli, 2000) to specify a complex genome-based approach in the vaccine development. Unlike the conventional approach that requires a laborious process of a selection of individual components important for the induction of protective immune response, reverse vaccinology offers a possibility to use genomic information derived from *in silico* analysis of the sequenced organisms. This approach can significantly reduce the time necessary to identify the antigens for the development of a candidate vaccine and enables a systematic identification of all potential antigens of pathogens including those which are difficult or currently impossible to culture. Of course, this approach is limited to the identification of protein or glycoprotein antigens, omitting such important vaccine components as polysaccharides and glycolipids. Nevertheless, reverse vaccinology can enable scientists to systemically classify the potential protective antigens, thereby helping to improve the existing vaccines and to develop efficient preparations against virtually any pathogen that has had its genome sequence determined.

As regards the process of activation of the immune system to produce an adaptive immune response, it is generally observed that the antigen by itself may not be adequate as a stimulating agent. Many potential antigens have no apparent immunizing activity at all when tested alone. In general, seamy side of pure recombinant protein antigens and synthetic peptide antigens is their poor immunogenicity. Therefore, potent adjuvants are required for highly purified antigen-based vaccines to be effective.

2. Adjuvants

2.1 Toll-like receptors and pathogen-associated molecular patterns

The word *adjuvant* is derived from the Latin root of *adjuvare*, which means to help. Thus, an adjuvant can be defined as any product which increases or modulates the specific humoral or cellular immune response against an antigen. The interaction between the innate and the adaptive immune responses is paramount in generating an antigen-specific immune response. The initiation of innate immune responses begins with the interaction of pathogen-associated molecular patterns (PAMPs) on the pathogen side with pattern-recognizing receptors (PRR) such as Toll-like receptors (TLRs) on the host cells involved in the innate immunity (e.g., dendritic cells). A major functional criterion commonly used for the evaluation of various new adjuvants involves their ability to stimulate the innate immunity cells. This would include engaging and other PRRs and the co-receptors and intracellular adaptor signalling proteins with which they are associated. PAMPs and their derivatives are utilized by adjuvant developers to harness the power of innate immunity to channel the immune response in a desired direction.

Based on the identification of several TLRs and PAMPs recognized by them, various PAMP agonists were tested as adjuvans. Examples of TLR-PAMP specific interaction include bacterial or viral unmethylated immunostimulating CpG oligonucleotides interacting with TLR9, liposaccharide and its component monophosphoryl lipid A (MPLA) interacting with TRL4. These two types of adjuvants are in advanced stage of testing in clinical trials and some already licensed vaccines contain MPLA in liposomal form. Further liposomal or lipid-based particle formulations of both CpG and MPLA are under development and testing.

2.2 Muramyl dipeptide and other muropeptides

Very specific group of PAMPs is represented by peptidoglycans (PGN). Both Gram-positive and Gram-negative bacteria contain PGN which consists of numerous glycan chains that are cross-linked by oligopeptides. These glycan chains are composed of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) with the amino acids coupled to the muramic acid. Muropeptides are breakdown products of PGN that bear at least the MurNAc moity and one amino acid (Traub et al., 2006). One of the prominent muropeptides is muramyl dipeptide (MDP), which is known since the 1970s.

Recently, the molecular bases for MDP recognition and subsequent stimulation of the host immune system have been uncovered. Myeloid immune cells (monocytes, granulocytes, neutrophils, and also DCs) possess two types of intracellular receptor for MDP, namely NOD2 and Cryopyrin (inflammasome-NALP-3 complex) (Agostini et al., 2004; Girardin and Philpott, 2004; McDonald et al., 2005). These two receptors recognize MDP/MDP analogues minimal recognition motifs for bacterial cell wall peptidoglycans (Girardin & Philpott, 2004; McDonald et al., 2005). NOD2 is also expressed in specialised epithelial cells, Paneth cells, localised in crypts of Lieberkün, which are producers of antimicrobial peptides having direct antimicrobial activity together with signalling functions within the immune system. Induction of an innate immune response against *Cryptosporidium parvum* infection by the liposomal preparation of lipophilic norAbuGMDP was demonstrated by us in newborn goats (Turanek et al., 2005) and this data is supportive of the present view of the role of MDP recognition in inducing both specific and innate immune responses.

Here GMDP abbreviates N-acetylglucosaminylmuramyl dipeptide. Another recently reported sensor of MDP is Cryopyrin (also known as CIAS1 and NALP3), which is a member of the NOD-LRR family (Agostini et al., 2004). Cryopyrin is a part of the inflammasome complex that is responsible for processing caspase-1 to its active form. Caspase-1 cleaves the precursors of interleukin IL-1β and IL-18, thereby activating these proinflammatory cytokines and promoting their secretion. IL-1β is known to be a strong endogenous pyrogen induced by MDP. We showed that norAbu-MDP analogues were not pyrogenic even at a high concentration, much higher than the concentrations used for vaccination. We supposed that the modification introduced into the structure of MDP to get norAbuMDP analogues had not changed their affinity to NOD2 but had substantially decreased the affinity to cryopyrin. This hypothesis is in accordance with our data on pyrogenicity and is being currently tested in appropriate *in vitro* models. In addition, murabutide, another nonpyrogenic derivative of MDP was shown not to be able to induce detectable level of IL-1ß in sera of treated volunteers (Darcissac et al., 2001).

The expression of NOD2 in dendritic cells is of importance with respect to the application of MDP analogues as adjuvants. Nanoparticles like liposomes are able to provide a direct co-delivery of a danger signal (e.g., MDP) together with the recombinant antigen and therefore to induce an immune response instead of an immune tolerance. This is especially important for weak recombinant antigens or peptide antigens. Clearly, the recognition of MDP by DCs is crucial for the application of MDP analogues as adjuvants. Although the immuno-stimulatory effects of MDPs have been described for over three decades, the process of molecular recognition and binding of MDP/MDP analogues to NOD2 and cryopyrin receptors remains unclear. Within the cell, MDP/MDP analogues trigger intracellular signalling cascades that culminate in the transcriptional activation of inflammatory mediators such as the nuclear transcription factor NF-κB. The biological effects of muramyl peptides have been described for over three decades. The mechanism underlying their internalization of MDP to the cytosol, where it is sensed by NOD2 and cryopyrin, remains unclear. Liposomes probably play the role of efficient carriers for MDP and its analogues on the pathway from extracellular milieu into the cytosol, where they trigger intracellular signalling cascades that culminate in transcriptional activation of inflammatory mediators such as the nuclear transcription factor NF-κB pathway. In case of liposomal formulation of various MDP analogues, the relevant intracellular pharmacokinetics, molecular recognition and binding affinity towards NOD2 and Cryopyrin remain to be determined. Such differences found for various MDP analogues are responsible for their various biological activities (e.g., pyrogenicity, ability to induce the innate immune response etc.) and, therefore, could be utilised for a precise tuning of the intensity and type of immune response. Since the discovery and first synthesis of MDP, about one thousand various derivatives of MDP have been designed, synthesised, and tested to develop an appropriate drug for an immunotherapeutic application that would be free of the side effect exerted by MDP. The main side effects of MDP are pyrogenicity, rigor, headache, flue-like symptoms, hypertension etc. Only several preparations reached the stage of clinical testing and only Mifamurtide (Fig. 1) was approved for the treatment of osteosarcoma.

2.2.1 Mifamurtide (MTP-PE)

In pyrogenicity test in rabbits, pyrogenic activity of Mifamurtide i.v. was comparable to that of MDP. In several studies with cancer patients refractory to standard therapy, infused with liposomal Mifamurtide at a dose range of 0.01 – 1.8 mg/m^2/dose, dose-dependent fever (in common about 70% of patients) and rigor (about 50% of patients) were the most prominent from a number of acute systemic toxicities (Creaven et al., 1990).

Mifamurtide was also assayed as an additional immunomodulator in an MF59-adjuvanted influenza virus vaccine (Keitel et al., 1993) and HIV-1 vaccine (Keefer et al., 1996); systemic symptoms including fever, chill, and nausea made these vaccines unsuitable for clinical use. Today, the main interest lies in clinical trials for liposomal Mifamurtide as a component of three-drug chemotherapy of osteosarcoma (Anderson P.M, 2006; Anderson et al., 2010).

2.2.2 Romurtide (Muroctasine)

In healthy volunteers, s.c. administration of Romurtide (Fig. 1) at a dose of 200 µg induced - besides local pain and redness - an approximately 1° C increase in body temperature with great individual variability in the course of pyrogenicity curves; normalization occurred within 48 hours (Ichihara et al., 1988). Fever accompanied by chill and headache was also the most common adverse reaction in cancer patients treated with Romurtide at a dose range of 100 – 400 µg/dose s.c. for the restoraion of haemopoiesis after chemotherapy and/or radiotherapy (Tsubura et al., 1988) (Azuma & Seya, 2001; Tsubura et al., 1988).

2.3 nor-Muramyl glycopeptides

We found that a combination of structural modifications both in the saccharide and peptide moiety of MDP and GMDP molecules leads to significant suppression or elimination of pyrogenicity and potentiation of immune-stimulatory activity (Fig. 2). The substitution of muramic acid with normuramic acid and L-alanine with L-2-aminobutyric acid has lead, in the case of the norAbu-MDP molecule, to a decrease of pyrogenicity and, at the same time, to the potentiation of immunoadjuvant activity. If the same structural change is carried out in GMDP molecule, a non-pyrogenic and highly immunoadjuvant analog, norAbu-GMDP is obtained. Furthermore, it has been demonstrated that by the introduction of bulky lipophilic residues into the molecules of these analogs, immunomodulatory activity can be effectively profiled, while the favourable pharmacological parameters of the parent structures are retained. These facts motivated our aims to design and prepare the new groups of lipophilic analogs of norAbu-MDP and norAbu-GMDP, which differ in the character and topology of the lipophilic residue. We primarily aimed to modify their immunopharmacologic parameters. norAbu-MDP-Lys(L18), i.e. MT05, belongs to them. In accordance with our premise, all the new compounds were nonpyrogenic (rabbit test), and the character and topology of the lipophilic residue had a significant effect on their immunologic parameters. As an example, the structural differences between norAbu-MDP-Lys(18) (MT05) and Romurtide, which influenced the profile of the effects and lead to the elimination of pyrogenicity, are depicted in Fig. 3. (Ledvina M., Turánek J., Miller A.D., Hipler K.: Compound (Adjuvants): PCT appl., WO 2009/11582 A2, 2009.)

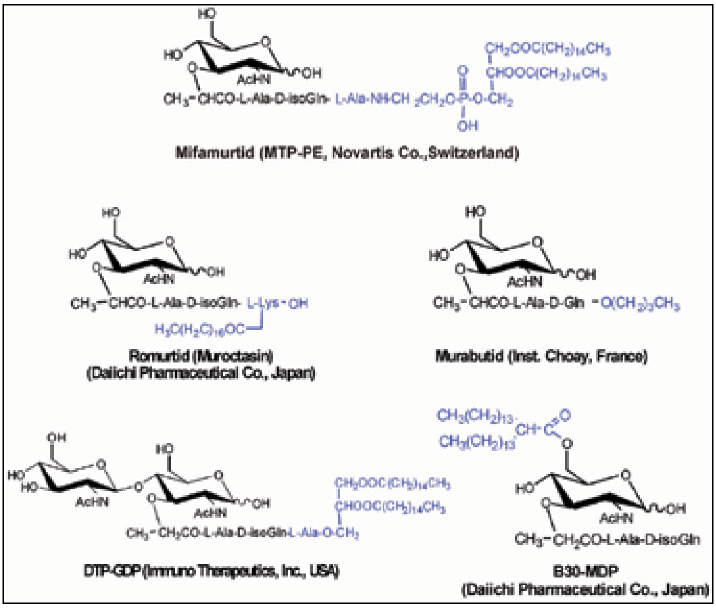

Mifamurtid (MTP-PE, Novartis Co.,Switzerland)
Romurtid (Muroctasin)
(Daiichi Pharmaceutical Co., Japan)
Murabutid (Inst. Choay, France)
DTP-GDP (Immuno Therapeutics, Inc., USA)
B30-MDP
(Daiichi Pharmaceutical Co., Japan)

Fig. 1. MDP derivatives developed by various pharmaceutical companies as adjuvants and immunotherapeutics.

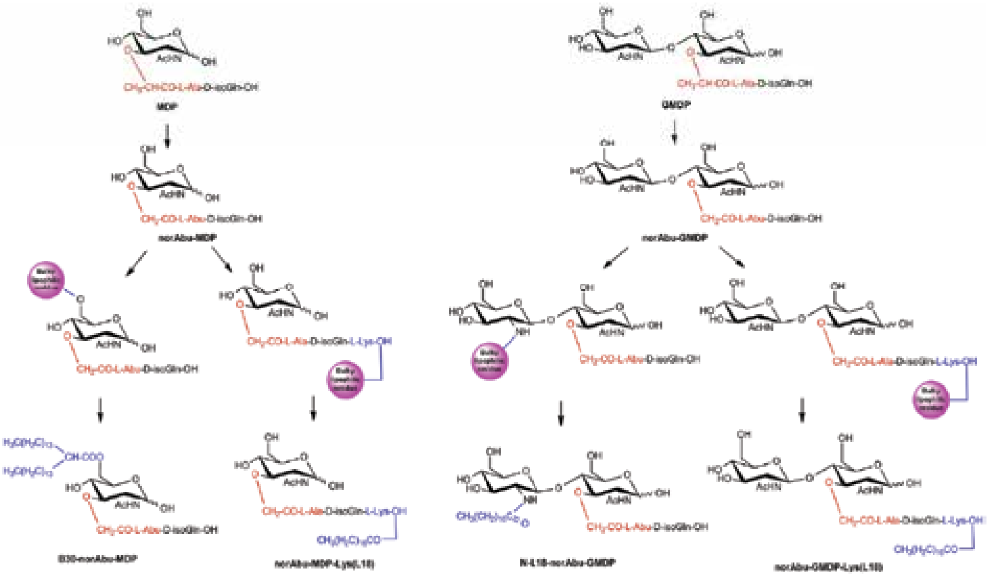
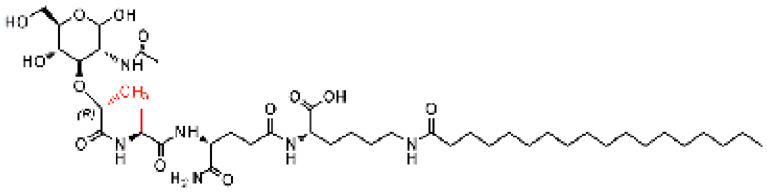
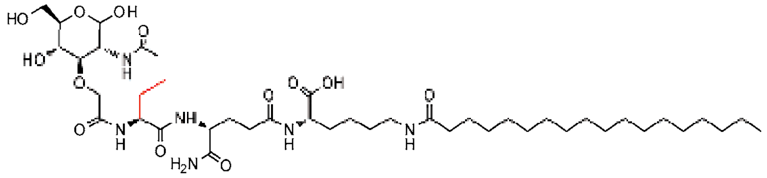

MDP
GMDP
norAbu-MDP
norAbu-GMDP
B30-norAbu-MDP
norAbu-MDP-Lys(L18)
N-L18-norAbu-GMDP
norAbu-GMDP-Lys(L18)

Fig. 2. Transformation of MDP into norAbu-MDP and GMDP into norAbuGMDP and formulae of their hydrophobised derivatives suitable for development of lipid-based adjuvants.



Romurtide (MDP-Lys(L18))



MT-05 (MDP-Lys(L18))

Fig. 3. Structural differences between Romurtide and MT05.

3. Liposomes

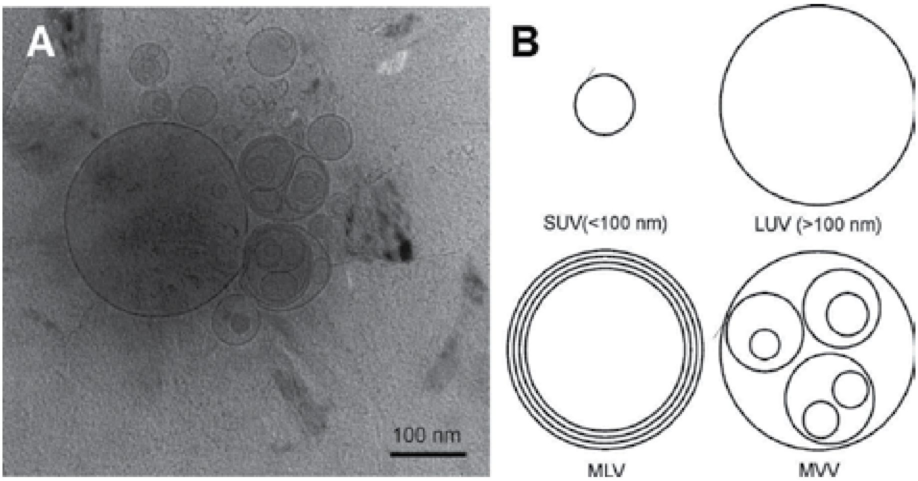
3.1 General characterisation of liposomes

Liposomes, membrane-like spherical structures consisting of one or more concentric lipid bilayers enclosing aqueous compartments were first formulated and described by Alex Bangham in 1965 (Bangham et al., 1965), and have become a useful tool and model in various areas of science. Liposomes represent the oldest and the most explored nano- and micro systems for biological studies on model membranes and for medical applications, especially for drug formulations, because they eliminate or suppress organ-specific toxic side-effects of various drugs (Allen, 1997). Through their 46-year history, liposomes have been approved as suitable delivery systems for applications ranging from cosmetics and dermatology, anti-infection and anticancer therapy and diagnostics up to human as well as veterinary vaccines (Gregoriadis, 1995).

Liposomes are classified in terms of number of bilayers enclosing the sequestered aqueous volume as follows: unilamellar, oligolamellar, and multilamellar. Unilamellar vesicles can be further divided into small unilamellar vesicles (SUVs) with a large curvature, and large unilamellar vesicles (LUVs) with a low curvature and hence, with properties similar to those of a flat surface. Multilamellar vesicles (MLVs) are liposomes that represent a heterogenous group in terms of size and morphology (Cullis et al., 1987). Lipid composition, size and morphology are variables determining the fate of liposomes in biological milieu; therefore, the selection of suitable method for the preparation of liposomal drugs and vaccines is of importance in respect to subsequent animal experiments and future successful marketing of the product. Schematic structures of various types of liposomes as well as realistic picture obtained by cryoelectron microscopy are presented in Fig. 4.

3.2 Liposome-based vaccines

The use of liposomes as vaccine adjuvants was first described by Allison and Gregoriadis in 1974 (Allison & Gregoriadis, 1974). Since that time, numerous studies were performed and proved that liposomes can be used to enhance the immune response towards a large variety of peptide and protein antigens derived from various microbial pathogens as well as tumours.



A) Photograph of various liposomal structures by cryoelectron microscopy. B) Schema of types of liposomes: SUV – small unilamellar vesicle, LUV - large unilamellar vesicle,, MLV – multilamellar vesicle, MVV – multivesicular vesicle

Fig. 4. Schematic representation of various morphological classes of liposomes and their real image obtained by cryoelectron microscopy.

The potential for the participation of liposome-based recombinant vaccines on the human and veterinary vaccine market is very promising (Adu-Bobie et al., 2003). Liposomal vaccines have been around for about 30 years and plenty of liposome variants have been developed; some of them with evident immune-stimulating properties and an attractive safety profile which resulted in registered products on the market or preparations in advanced stages of clinical testing. Liposomal hepatitis A vaccine is the first formulation of liposomes to become licensed for clinical use in humans (Gluck et al., 1992) (Hepatitis A - HepA, Epaxal http://www.crucell.com/Products-Epaxal). Epaxal liposomes contain influenza hemagglutinin protein which facilitates their binding and endocytosis by specific receptor on antigen presenting cells. Such forms of liposomes are called virosomes. Liposomes represent almost ideal carrier system for the preparation of synthetic vaccines due to their biodegradability and versatility as regards the incorporation of quite a number of various molecules having different physico-chemical properties (the size of the molecule, hydrophilicity or hydrophobicity, the electric charge).

The molecules and antigens can be either sterically entrapped into the liposomes (the internal aqueous space), or embedded into the lipid membrane (e.g., membrane-associated proteins/antigens) by hydrophobic interactions. Further, they can be attached to either the external or the internal membrane by electrostatic, covalent or metallo-chelating interactions. It is possible to encapsulate simultaneously various compounds into the liposomes: hydrophilised/lipophilised adjuvants (e.g., MPL A, CpG oligonucleotides, MDP and its analogues), soluble or membrane protein antigens, and ligands for the targeting to specific receptors on the antigen-presenting cells. Further, liposomes can be coated with mucoadhesive biopolymers, or undergo surface-charge modifications (e.g., by cationic lipids) (Altin & Parish, 2006).

As a great advantage, liposomes can be used for the preparation of self-assembling hybrid supramolecular nanosystems such as proteoliposomes, which can combine liposomal nanoparticles with suitable immunopotentiating/adjuvant molecules (e.g., MPL A, CpG oligonucleotides, MDP, and its analogues).

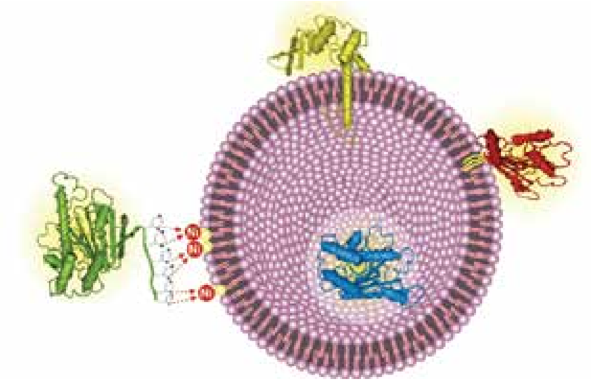
3.3 Liposomes as antigen carriers

Structural diversity of liposomes permits tailoring of liposome-based vaccines to obtain an optimal adjuvant effect for a particular antigen. Their safety profiles and ability to induce an immune response makes them likely to be included in vaccine formulations. Liposomal formulations offer several major advantages. They can (1) prevent degradation of the delivered antigens and adjuvants; (2) allow membrane proteins to reconstitute and preserve their antigen structure; (3) increase the antigenic effect of weak immunogens; (4) target antigen-presenting cells (APCs) and direct the antigen to MHCI or MHCII presentation; (5) contain the antigen and adjuvant molecules in one particle, thus functioning simultaneously as a delivery system and a vaccine adjuvant; and (6) reduce the antigen and adjuvant doses required for an immune response and controlled release. Therefore, liposomes provide a safe and effective platform for construction of subunit vaccines.

With respect to the physico-chemical nature of liposomes and phospholipid bilayers, the liposomes represent one of the most versatile structures for the preparation of drug delivery systems. Both hydrophobic and hydrophilic protein or peptide antigens can be associated with liposomes. Generally, antigens can be associated with liposomes in two ways and it is known that the encapsulated liposomal antigens induce a different immune response than the surface-linked antigens in both humoral (Shahum & Therien, 1988) and cell-mediated immunity (Fortin et al., 1996). If entrapped into the internal aqueous space of a liposome, the protein or peptide antigen is protected against proteolytic degradation and the antigen clearance is decreased. On the other hand, the liposomal membrane represents a barrier restricting the interaction of the antigen with and its recognition by B-cells. Especially, the stable multilamellar liposomes were found to be low immunogenic (Shek & Heath, 1983) and the antibody response reached is low or absent when the liposomes are made of lipids with a high transition temperature; in other words, when they are composed of saturated phospholipids (Gregoriadis et al., 1987). These liposomes are very stable in body fluids as well as in digestive tract and prevent a release of the entrapped antigen. Also the interaction of the encapsulated antigens with B-cells is limited. The fluidity of liposomes was found to be an important parameter also for the immune response towards a surface-linked antigen. Again, more fluidic liposomes composed of unsaturated phospholipids are more efficiently phagocytosed by APC and induce a one order of magnitude higher immune response than rigid liposomes composed of saturated phospholipids (Uchida & Taneichi, 2008).

Liposomes are potentially very useful for the construction of vaccination systems given their facile biodegradability and versatility as carriers for varieties of molecules having different physico-chemical properties (such as size, hydrophilicity, hydrophobicity, or net electrical charge). Liposomes also offer the possibility to associate or entrap simultaneously more than one type of molecules. Of particular interest to us has been the co-association of hydrophilic or lipophilic adjuvants (e.g., monophosphoryl lipid A [MPL A], CpG oligonucleotides, muramyl dipeptide (MDP), and/or MDP lipophilic analogues) with soluble or membrane protein antigens or ligands for the targeting of specific receptors on antigen-presenting cells. The molecules and antigens can be either sterically entrapped into the liposomes (the internal aqueous space), or embedded into the lipid membrane (e.g., membrane-associated proteins/antigens) by hydrophobic interactions (Fig. 5). The ligands for the targeting to specific receptors on the antigen-presenting cells can significantly enhance the intensity of the immune response (Altin & Parish, 2006). For the mucosal application, the liposomes can be coated with mucoadhesive biopolymers or modified with surface-charge modifiers (e.g., cationic lipids). In this way, liposomes become a versatile platform that represents a real multifunctional vaccination carrier.

The importance of liposomes for the effective co-administration of adjuvants could be demonstrated using MDP as an example. MDP has a weak immunoadjuvant activity in aqueous solution due to its rapid excretion into urine. Appropriate formulations of hydrophilic MDP in "water in oil" emulsions (Parant et al., 1979) or liposomes were used to harness its full adjuvant potenital (Tsujimoto et al., 1986). Some lipophilic derivatives of MDP like B30-MDP and MDP-Lys (L18) were synthesised and tested as adjuvants for recombinant hepatitis B surface antigen (Tamura et al., 1995) or influenza surface antigens hemagglutinin and neuraminidase (Nerome et al., 1990). We used new synthetic nonpyrogenic lipophilic analogues of norAbu-MDP modified at a peptide part by hydrophobic ligands (Fig. 2) and these well defined synthetic molecules were used for the first time in combination with metallochelating liposomes to construct an experimental recombinant vaccine. Surprisingly, we have found that at certain surface density of lipophilic analogues of norAbu-MDP (about 5 mol % of total liposomal lipid), the liposomes are promptly recognised and phagocytosed by human dendritic cells. The phagocytosis is about one order of magnitude higher than that of proteoliposomes or liposomes lacking norAbu-MDP adjuvant. This finding implicates an existence of receptors on dendritic cells, which can recognise some molecular pattern formed by the hydrophilic part of norAbu-MDP exposed on the liposomal surface (illustration of this phenomenon is in Fig. 12B).



a) physical entrapment inside the liposome (blue protein); b) reconstitution of membrane protein in lipid bilayer via hydrophobic transmembrane domain (yellow protein); c) anchoring of lipidised protein onto liposomal surface or attachment of recombinant protein onto the liposomal surface by covalent bond using activated lipids (red protein); d) attachment of recombinant protein onto the liposomal surface by non-covalent bond using metallochelating lipids

Fig. 5. Association of protein antigen with liposome.

3.4 Methods of liposome preparation

The laboratory and industrial procedures for the liposome preparation have been established and liposomes have been approved by FDA for biomedical applications. A variety of procedures for the preparation of various types of liposomes has been developed and reported in several reviews and monographs (Gregoriadis, 1992; Woodle &

Papahadjopoulos, 1988). To classify these methods, they were arranged in three categories: 1) mechanical dispersion methods such as hand shaking or vortexing, sonication, and high pressure homogenisation; 2) detergent-solubilizing dispersion methods including solubilized lecithin dispersion with sodium cholate or octylglucoside; and 3) solvent dispersion methods such as ethanol injection, ether infusion, and reverse-phase evaporation. These primary processes can be linked with secondary processes such as high-pressure homogenisation or extrusion through polycarbonate filters of various pore size, which are easy ways to prepare liposomes of a desired size and morphology (Barnadas-Rodriguez & Sabes, 2001; Berger et al., 2001; Cullis, 1987; Hope et al., 1985; Perrett et al., 1991; Turanek, 1994; Woodle & Papahadjopoulos, 1988; Schneider et al., 1995).

The stability of proteins is limited and not all the methods are useful for the preparation of proteoliposomes, especially if the protein is to be entrapped inside the liposome. The detergent dilution method is characterized by very mild conditions during the process and is suitable for the reconstitution of membrane proteins. Because many recombinant proteins tend to precipitate, this method is also useful to work with these protein antigens. Next paragraph describes in detail a modified detergent dilution method for the preparation of proteoliposomes.

3.5 Preparation and characterization of metallochelating liposomes

Because the preparation of metallochelating liposomes represents a post-forming modification of liposomes, it avoids a denaturation of proteins owing to the process used for the liposome production. Therefore, nearly all the methods mentioned above could be used to prepare more or less monodisperse liposomes. Here we describe in brief a modification of the detergent removal method, which is suitable for the preparation of very monodisperse unilamellar liposomes that are useful for structural studies by various techniques (e.g., TEM, dynamic light scattering, and gel permeation chromatography).

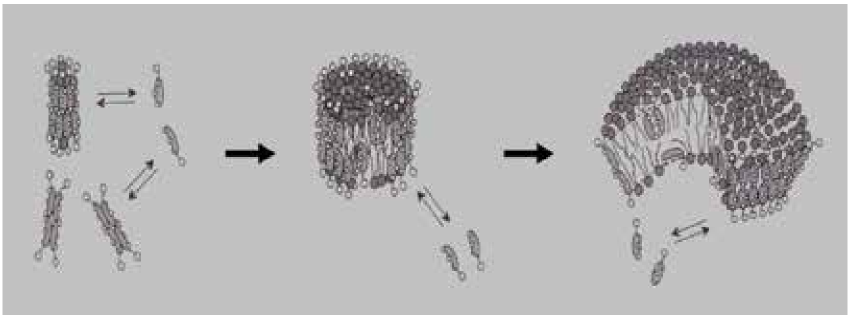
When essentially unilamellar monodisperse liposomes of spherical shape are needed (which is a prerequisite for a precise monitoring of the proteoliposome formation by dynamic light scattering), the detergent removal method is preferred to the other methods. The method is based on the transformation of phospholipid micelles stabilized by detergent with high critical micellar concentration (CMC) (e.g., cholate) to desk micelles and finally to vesicles during the process of detergent removal (Zumbuehl & Weder, 1981) (Fig. 6). The mild conditions provided by this method are advantageous for the preparation of proteoliposomes, especially for the reconstruction of membrane proteins (Rigaud & Levy, 2003) like viral or bacterial antigens or recombinant his-tagged proteins that are often prone to precipitation.

There are many variants of the detergent removal method, e.g., the dilution of the solution of mixed micelles, gel permeation chromatography, a simple dialysis or the controlled one in a special apparatus, cross-flow filtration, adsorption on beads, etc. (Schubert, 2003). The application of the flow-through ultrafiltration cell represents a new approach to the detergent removal method (Masek et al., 2011a). The linkage of the cell with systems like FPLC facilitates automation of the whole procedure and manipulation with the sample. The full control over the dialysis rate and the removal of the undesired residua (e.g., detergent, organic solvents, protein solubilizers) is ensured and various steps like an addition of required components through an injection valve during various stages of the liposome formation are easy to perform without breaching the sterile conditions. In this case, the sterile filter inserted in front of the cell inlet ensures that the sterility is kept during the whole process (Fig. 7). The low dead volume of the cell is of great importance for the preparation of liposomes and proteoliposomes in small laboratory scale. However, this arrangement enables also very easy up-scaling of the whole technology. A precise control over the rate of the detergent removal yields a final liposomal preparation of high monodispersity (PDI within the range of 0.05 - 0.06), which is shown to be reached routinely (Fig. 8). This monodispersity is better than those obtained by the dialysis method performed in the dialysis bags or slides (produced by Pierce) (PDI ≈ 0.08-0.12).

The size of the mixed micelles (≈5-6 nm; see Fig. 8) used by us for the preparation of liposomes is in good correlation with the Small's mixed micellar model proposing the structure of a small phospholipid bilayer disc stabilised at its hydrophobic edges by the molecules of cholate (Small, 1971; Schubert, 2003).

The process of the formation of the monodisperse liposomes is in good accordance with the proposed kinetic model of the micelle-vesicle transition based on a rapid formation of disk-like intermediate micelles followed by a growth of these micelles up to their critical size and their subsequent closure to form vesicles. The final size of the liposomal preparation could be controlled by ionic strength of the buffer used for the preparation of the micelles (Fig. 9). An increase of the NaCl concentration reduces CMC of cholate and shields the negative charge of the mixed micelles. These two factors are responsible for the formation and stabilisation of the large discoid bilayer micelles that are transformed into the larger liposomes (Schubert, 2003).

Various additives like bilayer stabilising sugars (e.g., sucrose) or recombinant protein solubilizers (e.g., urea, guanidine) are compatible with this method and can shift the size of the liposomes into the required range (Walter et al., 2000). Some recombinant proteins (e.g., circovirus envelope protein), which tend to precipitate in the absence of stabilizing buffers (imidazole and urea stabilizing buffer) were successfully linked onto metallochelating liposomes by one-step procedure based on the addition of the protein into the mixed micelle solution prepared in protein stabilizing buffer and transforming into proteoliposomes during the ultrafiltration procedure (Turánek, unpublished results).

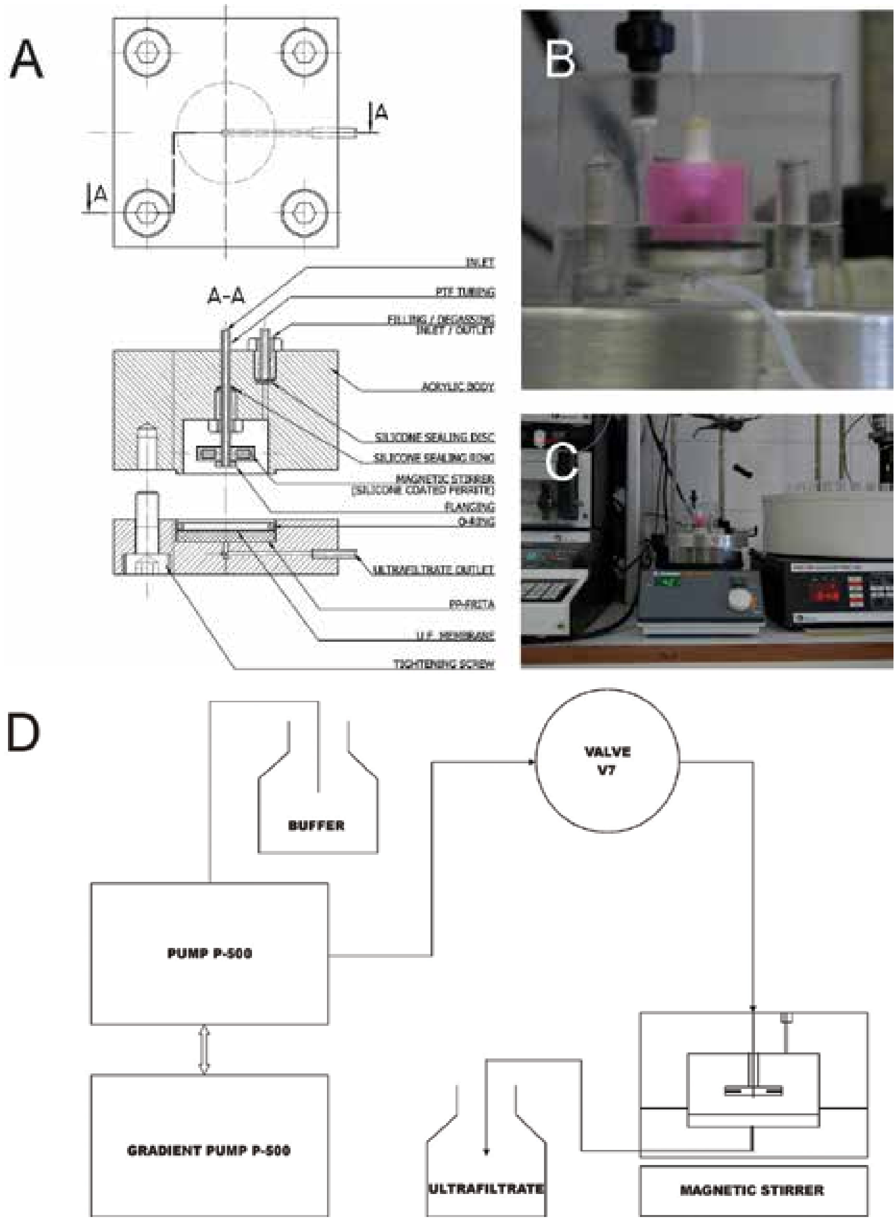


Adapted: R. Schubert, Methods Enzymol. 367 (2003) 46–70.

Fig. 6. Principle of detergent removal method and formation of liposomes from mixed bile salt-phospholipid micelles.

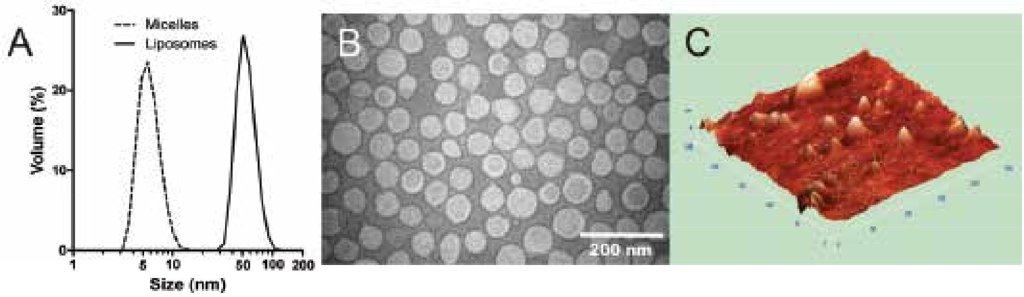
Small mixed micelles are fused in disc phospholipid micelles stabilized at edges by detergent. Further removal of detergent induces formation of large disc micelles which spontaneously vesiculate after reaching a critical size.

The transformation of micelles into liposomes during the ultrafiltration removal of cholate is a critical step affecting the final quality of liposomes. This process is easy to be monitored by DLS. The removal of cholate induced a formation of disc micelles, which was reflected by an increase of the micelle size and eventually by a formation of liposomes (Fig. 10). The process of liposome formation had been completed before the CMC of cholate was reached, as shown by the dashed vertical line. This line divides the flow-through volume axis into the left part, where micelles do predominantly exist and are transformed into liposomes, and the right part, where liposomes represent the main lipid form, while the residual detergent and other low molecular weight contaminants (e.g., traces of ethanol or tetrahydrofuran used to solubilize the lipids) are continuously removed by the process of ultrafiltration.


A
A
A
A-A
INLET
PTF TUBING
FILLING / DEGASSING INLET / OUTLET
ACRYLIC BODY
SILICONE SEALING DISC
SILICONE SEALING RING
MAGNETIC STIRRER (SILICONE COATED FERRITE)
FLANGING
O-RING
ULTRAFILTRATE OUTLET
PP-FRITA
U.F. MEMBRANE
TIGHTENING SCREW
B
C
D
VALVE V7
BUFFER
PUMP P-500
GRADIENT PUMP P-500
ULTRAFILTRATE
MAGNETIC STIRRER

(A) Schematic illustration of the ultrafiltration cell. (B) Photograph of the ultrafiltration cell in detail (pink: LR-PE-labelled liposomes inside the cell). (C) Schematic illustration of the linkage of the ultrafiltration cell with the FPLC system. (D) Photograph of the system

Fig. 7. System for preparation of liposomes by removal of detergent using ultrafiltration.



A) Size distribution of micelles and liposomes. The hydrodynamic diameters of the micelles and liposomes were determined by dynamic light scattering instrument NanoSizer NS (Malvern, UK) at 25 °C. Silica cuvette of 45-µl volume (Hellma, Germany) was used. (B) TEM micrograph of monodisperse liposomal preparation. (C) AFM micrograph of monodisperse liposomal preparation

Fig. 8. Size distribution of micelles and metallochelating liposomes analysed by dynamic light scattering and visualization of liposomes by TEM and AFM.

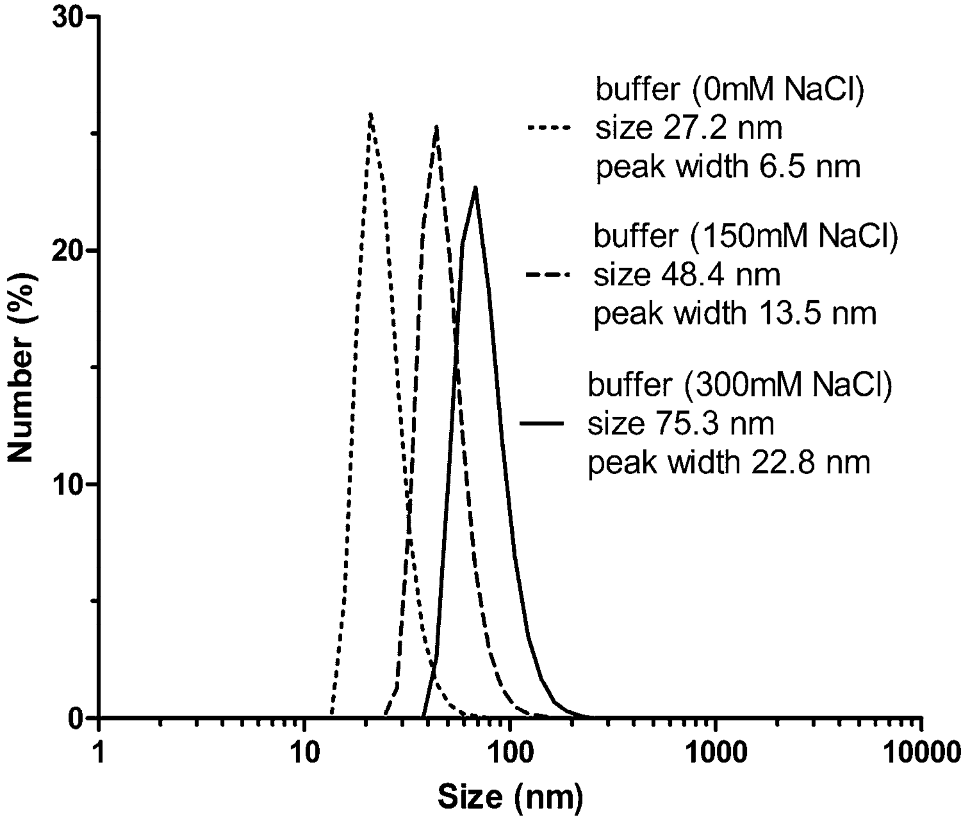


Fig. 9. Effect of ionic strength on the size distribution of liposomes prepared by detergent removal method.

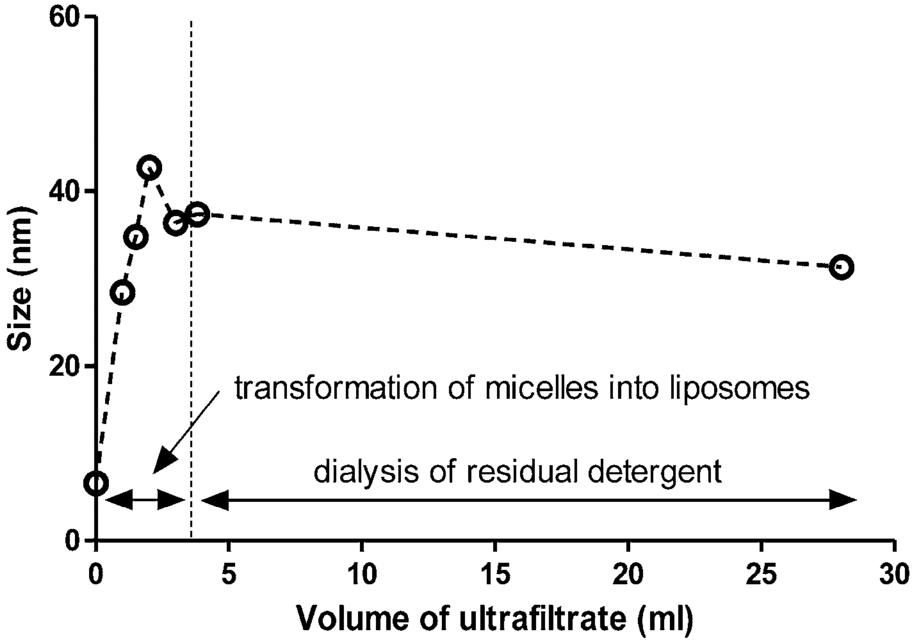


Fig. 10. Transformation of micelles into liposomes during ultrafiltration monitored by DLS. The dashed vertical line indicates the ultrafiltrate volume, when the CMC of sodium cholate was reached. This line divides the flow-through volume axis into the leftpart, where micelles do predominantly exist and are transformed into liposomes, and the right part, where liposomes represent the main lipid form, while residual detergent and ethanol/THF are continuously removed by the process of ultrafiltration.

4. Metallochelating bond and its application for construction of proteoliposomes

With respect to a potential application for the construction of vaccines, the question of *in vitro* and especially *in vivo* stability is of great importance. This problem could be divided into two fields. First, the stability of the liposomes themselves and second, the effect of the components presented in biological fluids (e.g., proteins and ions) on the stability of the metallochelating bond. It is beyond the scope of this chapter to address thoroughly this particular question. However, the GPC data indicate a good *in vitro* stability of the proteoliposomes containing recombinant His-tagged Outer surface protein C from *Borrelia burgdorferi* (rOspC) designated rOspC–HisTag during the chromatographic process, within which they experience a shear stress and dilution.

Also, the data on the incubation of rOspC-HisTag proteoliposomes in serum at 37 °C demonstrated the stability of the metallochelating bond linking the protein to the liposomal surface. In fact, *in vivo* fate of liposomes after the intradermal administration is different than that following an intravenous injection. First, dilution of proteoliposomes is not so rapid and second, the ratio of tissue fluid proteins to proteoliposomes is more favourable to proteoliposomes owing to their relatively high concentration at the site of application. Moreover, the flow rate of the tissue fluid within intradermal extracellular matrix is considerably lower than that of the muscle tissue or blood vessels. This fact is often overlooked. The stability of metallochelating bond probably depends also on the character of a particular protein. The study by Ruger shows that the single-chain Fv Ni-NTA-DOGS liposomes are unstable in human plasma and the majority of single-chain Fv fragments (anti CD 105) are released from the liposomal surface, which results in a lost of the specific targeting performance to the cells expressing a surface protein endoglin (CD 105) (Ruger et al., 2006). On the other hand, Ni-NTA3 –DTDA liposomes with single-chain Fv fragments (anti CD11c) bound onto the liposomal surface were able to target dendritic cells *in vitro* as well as *in vivo*. The application of the three-functional chelating lipid Ni-NTA3 –DTDA probably endows the metallochelating bond with a higher *in vivo* stability (van Broekhoven et al., 2004). Application of Ni-NTA-DOGS liposomes for the construction of experimental vaccine against sytemic *Candida* infection based on *Candida* Heat shock protein 90 kDa (rHSP90-HisTag) showed good stability in serum as well as strong immune response against recobinant rHSP90-HisTag antigen in mice (Masek et al., 2011b).

In vivo activity (immunogenicity) was also demonstrated for antigens associated with ISCOM particles via metallochelating lipid dipalmitoyliminodiacetic acid (Malliaros et al., 2004) and a peptide antigen associated with liposomes via Ni-NTA-DOGS (Chikh et al., 2002). Generally, metal ions, physico-chemical character of the metallochelating lipids and their surface density on the particles belong to the factors that could be optimized to get a required *in vivo* stability and, therefore, a strong immune response. The design and synthesis of new metallochelating lipids might accelerate a development and application of metallochelating liposomes for the construction of drug delivery systems and vaccines. Besides Ni^{2+}, other divalent ions such as Zn^{2+}, Co^{2+}, Fe^{2+}, and Cu^{2+}, have to be considered and experimentally tested as well.

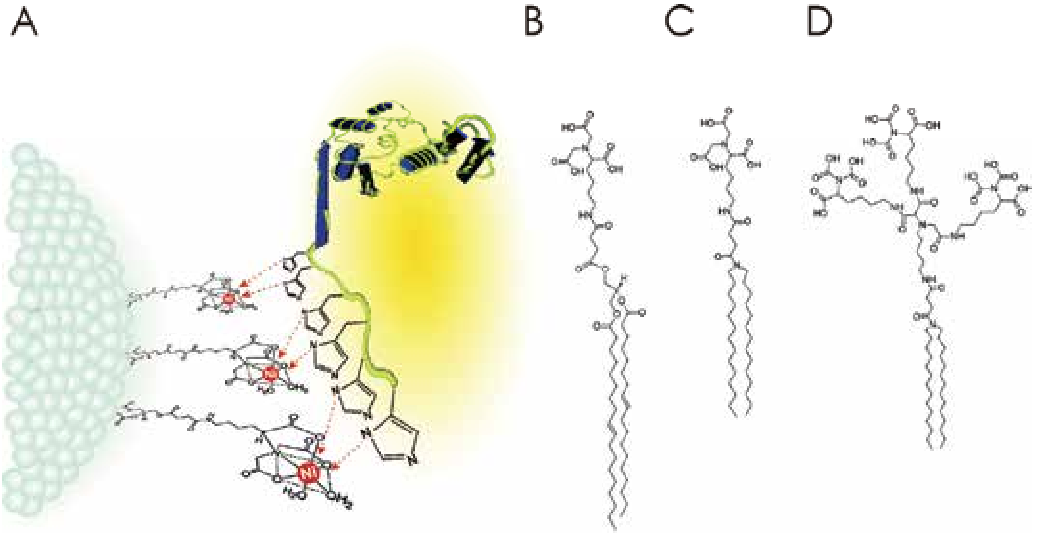


Fig. 11. Schematic illustration of recombinant His-Tagged protein bound onto the surface of metallochelating liposome (A) and formulae of metallochelating lipids NTA-DOGS (B), NTA–DTDA (C), trivalent NTA3 –DTDA (D).

5. Metallochelation liposomes for construction of experimental recombinant vaccines

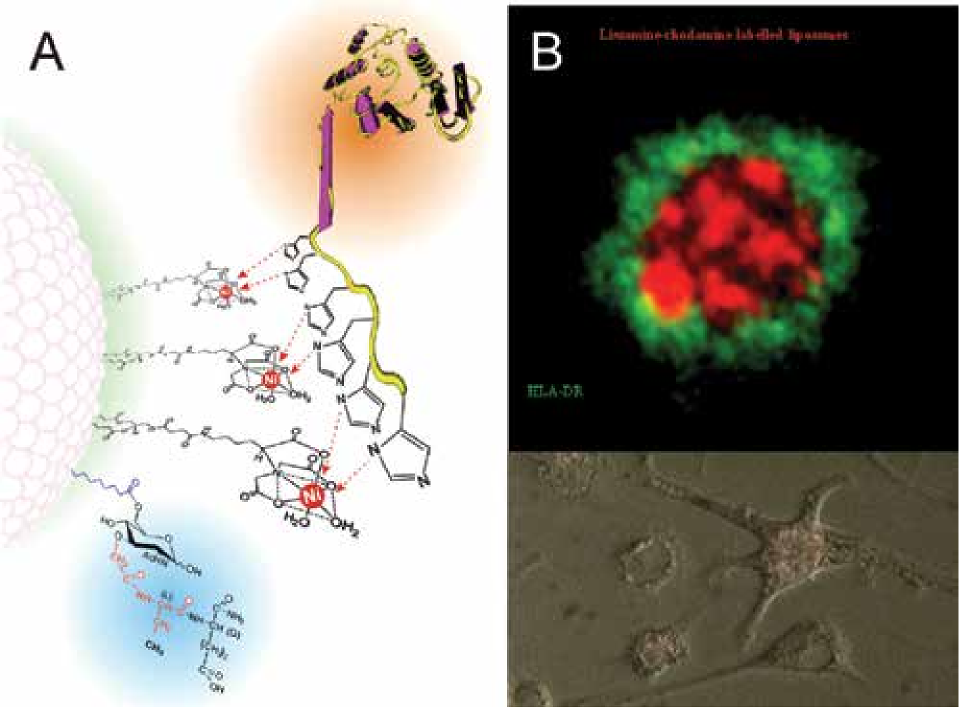
Only few papers report the implementation of the metallochelating lipids in the attachment of the recombinant proteins or synthetic peptides with His-Tag anchor (short peptide consisting of 4 to 6 molecules of histidine). Both reversible character and high affinity of the metallochelating bonds are very useful for the preparation of various self-assembling supra-molecular structures useful for the construction of experimental vaccines (Chikh et al., 2002; Malliaros et al., 2004; Masek et al., 2011a; Masek et al., 2011b). As an example of synthetic liposome-based recombinant vaccine we can use metallochelating liposomes and recombinant antigen rOspc-6HisTag derived from the pathogen Borrelia burgdorferi.

5.1 rOspC antigen *Borrelia* as an example for construction of metallochelation liposome-based vaccines

Lyme disease or Lyme borreliosis is an infectious disease caused by spirochetes of the *Borrelia burgdorferi* sensu lato complex vectored by ticks of the genus *Ixodes*. At least three species are pathogenic for humans, *B. burgdorferi* sensu stricto, *B. afzelii*, and *B. garinii*. The initial stage of Lyme disease is commonly associated with skin rash occurring within few weeks after the tick bites. Later the infection can spread to bloodstream and insult of joints, heart, and nervous system. Although not common, some patients experience late stage symptoms like arthritis, nervous system complications, or acrodermatitis chronica atrophicans. Here, slower response to antibiotics therapy, sometimes taking weeks or months to recover, or eventually, incomplete resolution is observed. In a few cases, antibiotic-refractory complications persist for months to years after antibiotic therapy, most likely due to infection-induced autoimmunity. Therefore, alternative approaches such as preventive immunisation are needed, mainly in the endemic areas (Krupka M, 2007; Tilly et al., 2008).

Protective immune response to *Borrelia* involves non-specific activity of complement, phagocytic cells and *Borrelia*-specific Th1-dependent response leading to production of complement-activating antibodies, in mouse presented mostly by IgG2a (IgG2b). During natural infection, nevertheless, *Borrelia* and tick saliva modulate the immune response toward non-protective Th2 type response, associated with production of neutralizing, poorly opsonizing *Borrelia*-specific antibodies (Vesely et al., 2009). *Borrelia* outer surface proteins OspA and OspC are among the most promising antigens for elicitation of opsonizing antibodies. The applicability of OspA antigen is limited because *Borrelia* expresses it mainly in the tick and the antibodies thus should act outside of the vaccinee's organism (Pal et al., 2000). Therefore continuously high level of OspA-specific antibodies is required to prevent *Borrelia* transfer. In contrast, OspC is expressed during the transfer and the initial stage of infection. In this case the vaccine-induced immune memory has enough time to initiate the production of opsonizing antibodies preventing *Borrelia* spreading (Tilly et al., 2006).

OspC antigen can be used here as an example of reverse vaccinology approach. Full length recombinant OspC is difficult to prepare in high yield and purity. Production of Osp-s for vaccination purposes is hindered by low yield of fully processed lipidized Osp antigens or low immunogenicity of their non-lipidized versions. In our experiments, removing of N′ terminal lipidation signal was associated with an increase of the recombinant protein yield and purity but, as demonstrated also for other *Borrelia* lipoproteins, a decrease in immunogenicity (Erdile & Guy, 1997; Gilmore et al., 2003; Lovrich et al., 2005; Weis et al., 1994). Induction of OspC-specific opsonizing antibodies to non – lipidised OspC could be enhanced by appropriate adjuvants and carriers like such as various modification of liposomes. It was reported that immunisation of mice with non-lipidated OspC in strong adjuvants (Complete Freund's Adjuvant, TiterMax, or Alum) could induce intense OspC-specific antibody responses (Earnhart et al., 2007; Earnhart and Marconi, 2007; Gilmore et al., 1996; Gilmore and Mbow, 1999; Ikushima et al., 2000). Here we demonstrated that similarly strong response could be elicited by immunisation of experimental mice with metallochelating nanoliposomes with entrapped lipophilic derivatives of norAbu-MDP as a potent adjuvant molecule.

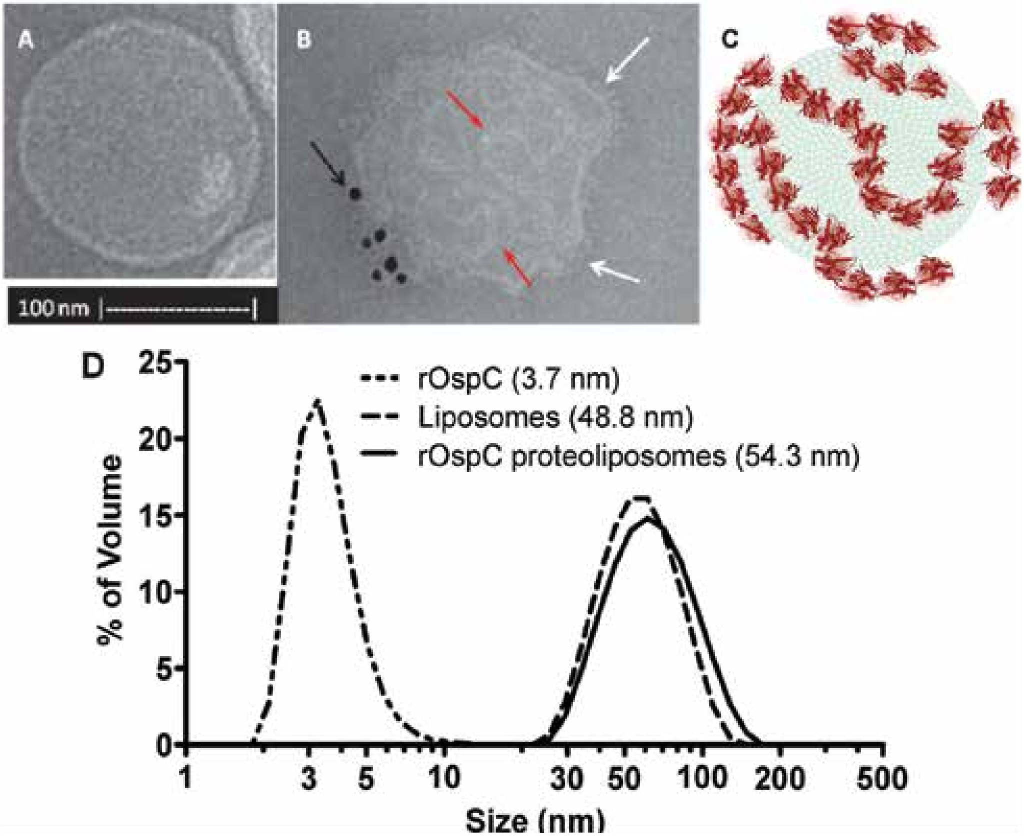


A) Scheme of metallochelating liposome with His-Tag recombinant protein antigen bound to the surface. Lipophilic nor-AbuMDP analogue as adjuvant is exposed on the surface of proteoliposome; B) Upper photograph (confocal microscop) shows human dendritic cell with phagocytosed prototypical vaccination nanoparticles cumulated in giant lysozomes (green – HLA-DR marked with antiHLA-DR antibodies; red – Lyssamine-rhodamine labelled vaccination particles). Lower photograph (Hofman modulation contrast and fluorescence microscopy) shows dendritic cells with accumulated vaccination particles inside (red fluorescence)

Fig. 12. Schematic drawing of prototypical recombinant protein vaccination nanoparticle and interaction with human dendritic cells.

5.2 Structure of OspC metallochelating proteoliposome

Protein-His-Tag/metallochelating lipid complex is anchored in phospholipid bilayer by its lipid moiety. Egg yolk or soya phosphatidyl choline contain large portion of unsaturated fatty acids, therefore DOGS-NTA lipid is freely miscible with these lipids and phase separation do not occur in liposomal bilayer. In other words, contribution of lipids to the formation of various protein structures on the liposomal surface is negligible.



A) Structure of plain metallochelating liposome (TEM) B) Structure of metallochelating liposome with OspC bound onto the surface (white arrows – OspC); black arrow – OspC marked by 10nm immunogold particles; small red arrow – bead chain of OspC molecules; C) Schematic presentation of bead chain model of proteoliposome D) Size distribution and hydrodynamic diameter of OspC, plain metallochelating liposomes and OspC proteoliposomes analysed by DLS. The size distribution of parent monodispersed metallochelating liposomes (dashed line) was compared with that of OspC proteoliposomes (full line). As a reference, size distribution of OspC (dotted line) is shown. Numbers in brackets represent mean hydrodynamic diameters of particles

Fig. 13. Characterisation of OspC preoteoliposomes.

This is advantageous for study of protein-protein interaction of proteins anchored on the surface of liposomes. Moreover, anchoring of proteins via metallochelating bond produces highly oriented binding of proteins because the proteins are attached to liposomes exclusively by His-tagged end of the polypeptide chain. If some interaction between liposomal surface bound proteins exists, we can observe formation of various structures which are conditioned by the character and number of protein-protein interactions. Another important feature of metallochelating proteoliposomes is relatively high surface concentration of proteins. This concentration could be set by changing the DOGS-NTA/phosphatidyl choline ratio in the lipid mixture used for preparation of liposomes.
The ultrastructure of the rOspC-His-tag proteoliposomes was revealed by TEM (Fig. 7). rOspC-His tag as an example of the preoteoliposomal structure "Bead chain" model. Binding of individual molecules of OspC protein onto the liposomal surface is clearly visible

on the rim of the liposome and immunogold staining confirmed the identity of OspC molecules as well as preservation of the epitopes recognised by polyclonal antibodies (Fig. 7b). In the case of recombinant OspC, TEM micrograph (Fig. 7b) showed that the individual molecules of OspC antigen were bound onto the surface of the liposome and some organisation in beads-like structures was revealed. This observation testifies against the simplification of the proteoliposomal structures and hence against accepting the simple schematic concept based on the random distribution on the liposomal surface.

Binding of OspC onto the surface of metallochelating liposomes was proved by an increase of hydrodynamic diameter as assayed by DLS. The increase of the size of proteoliposomes is well distinguished from plain liposomes, even if the increase of the size is only 5.5 nm (Fig.. 13D). This precise measurement was allowed by a preparation of parent monodispersed metallochelating liposomes and pointed to the importance of using monodispersive liposomal preparation for such a study (Fig. 8). In the case of homogenous coating of liposomes by OspC, the increase in the size should be of about 7.4 nm, theoretically. A lower increase of the size (5.5 nm) indicates only partial coating of the liposomal surface by the protein and this is in a good accordance with the structure revealed by TEM (Fig. 7b).

Binding of OspC onto metallochelating liposomes was confirmed also by GPC used as an independent method (Fig. 9) The liposomal fraction was separated from free protein and OspC was assayed by SDS PAGE followed by immunoblot (Fig. 9C). The vast majority of OspC was shown to be bound onto liposomes and was only slightly ripped from their surface by shearing forces taking place during penetrating through GPC column. The tailing character of the OspC elution profile is supportive to this explanation. Stability of the metallochelating bond in model biological fluid was studied by incubation in undiluted human serum. In spite of the presence of serum, it was estimated that more than 60% of OspC was still associated with liposomes. Based on this data, the half life of OspC proteoliposomes in serum was estimated to be at least 1 hour.

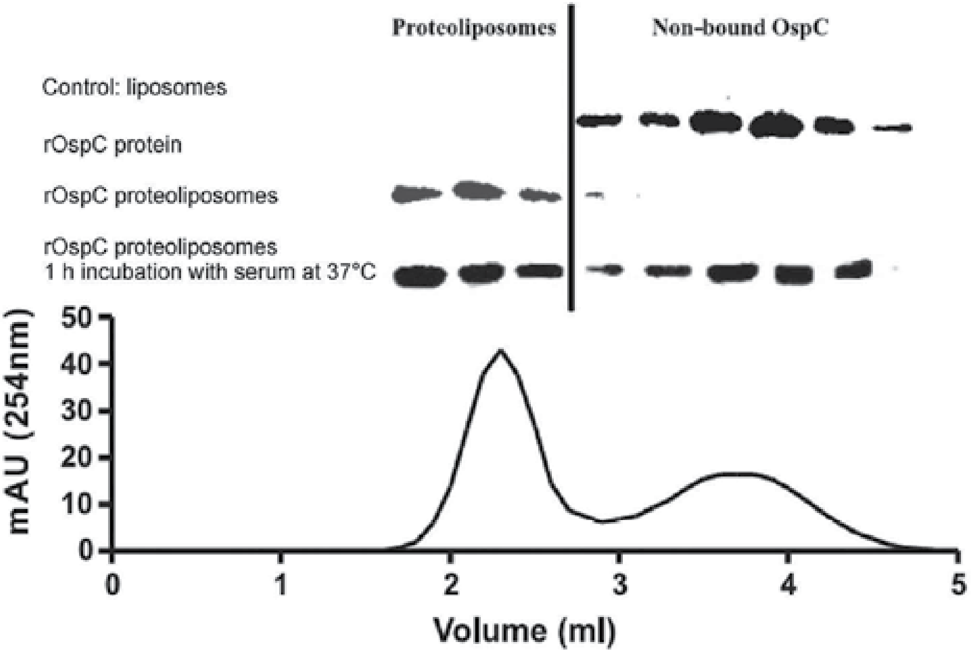


Fig. 14. Stability analysis of OspC proteoliposomes by GPC.

Non-bound OspC was separated from the proteoliposomes by gel permeation chromatography using Superose 6 column. OspC was detected in various fractions from GPC by immunoblot. OspC GPC elution profile is correlated with immunoblot assay.

5.3 Immunization experiments

New synthetic nonpyrogenic lipophilic analogues of NorAbu-MDP modified at a peptide part by two different hydrophobic ligands (Fig. 15) and for the first time these defined synthetic molecules were used in combination with metallochelating liposomes to construct an experimental recombinant vaccine. The important finding was that both MT06 and MT05 adjuvants exerted a high adjuvant effect comparable or better than MDP but proved itself as nonpyrogenic (rabbit pyrogenicity test) and safe. While alum induced stronger antibody response in IgG1 subtype, both MT06 derivatives and liposomal-MDP induced stronger immune response in both IgG2 subtypes. Interestingly, in comparison with MT06, the analogue MT05 induced stronger response in IgG1, IgG3 and IgM isotypes, respectively. This interesting finding pointed on the effect of lipophilic residues, which could not be supposed as the only accessory part of the molecule, but can significantly affect the quality of immune response. The position (peptide or sugar part) and the character (hydrophobicity and bulkiness) of the lipophilic function can affect the interaction with appropriate receptors as well as the metabolic degradation of the molecule. This aspect has not been described in the literature yet and is of interest for our understanding of the mechanism of action.

OspC itself did not elicit detectable OspC-specific antibodies of IgG, IgM, and IgA isotypes (Ig*). Similarly, when OspC bound onto the surface of liposomes was used for immunisation, only negligible increase of OspC-specific Ig* was detected after the second immunisation. In contrast, immunisation with OspC plus adjuvants (FCA, AlOH, MDP, MT05, and MT06) elicited strong OspC-specific antibody responses with ELISA titres of the same magnitude. The contributions of particular IgG isotypes for the immune response showed differences among various adjuvants (Fig. 15) (FCA was no tested for the IgG isotype response).

Mice (5 per group) were immunised by i.d. application of various liposome - adjuvant formulations of rOspC. Pooled sera from each group were used for ELISA analysis of specific antibodies titers. Naive sera were obtained before immunisation. ELISA plates were coated with 100 µl of rOspC (1 µg/ml), incubated with anti-mouse IgG1 (A), anti-mouse IgG2a (B), anti-mouse IgG2b (C), anti-mouse IgG3 (D), or anti-mouse IgM (E) and after addition of OPD plus H2O2, the absorbance was read at 490 nm on ELISA reader. The results are expressed as the end point titers +/- SD. Mean values are expressed in the table under each graph. Formulae of tested compounds in this experiment (F).

Although AlOH adjuvant induced strong OspC-specific antibody responses in total immunoglobulin level (Ig*) and IgG1 isotype (Fig. 15A), the response in complement-activating IgG isotypes (IgG2a and IgG2b; (Fig. 15B,C) was only modest. In comparison with AlOH, the synthetic adjuvant MT06 when combined with liposomes-bounded OspC induced strong OspC-specific response in isotypes IgG2a and at lesser extent in IgG2b (Fig. 15B). Application of another synthetic adjuvant MT05 was associated with dominancy of IgG3 and IgM (Fig. 15D,E). Furthermore, we compare responses to synthetic norAbu-MDP adjuvants with the response to liposomes-bounded OspC plus MDP, which elicits strongest OspC-specific antibodies in IgG2b and in lesser extent in IgG2a isotype (Fig. 15B,C).

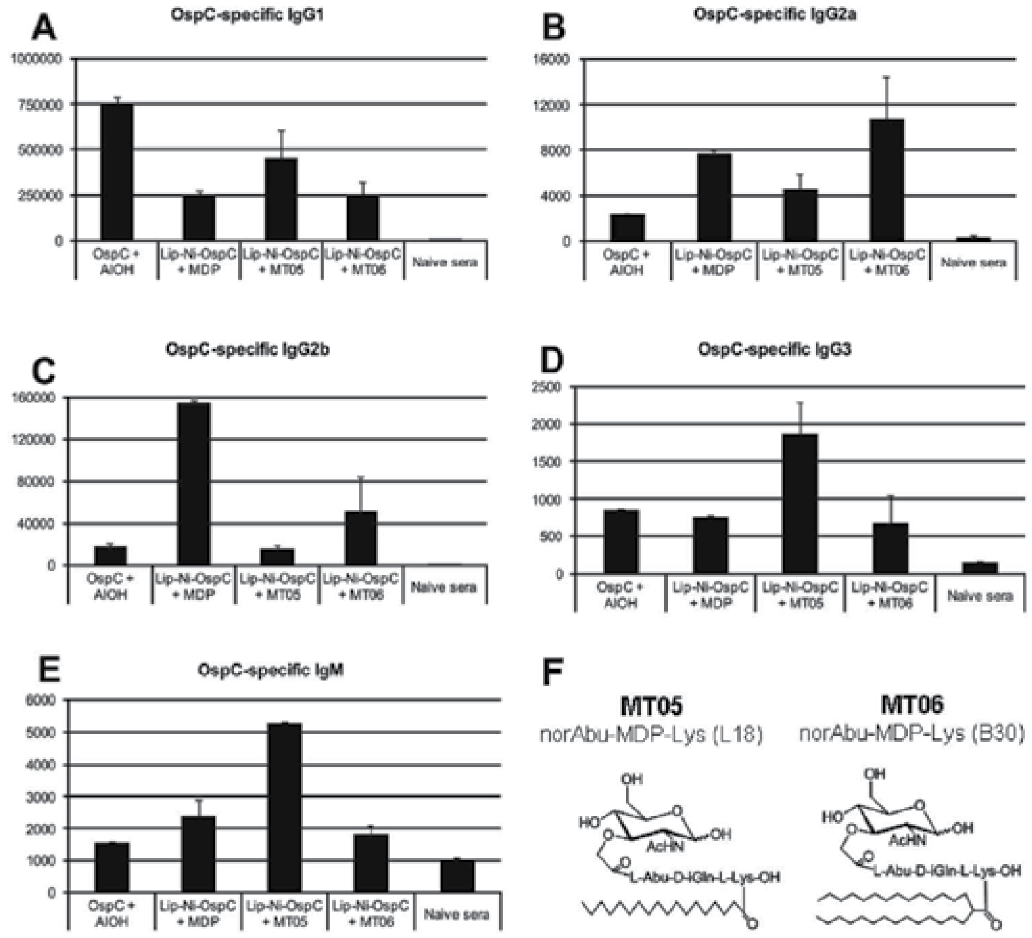


Fig. 15. ELISA analyses of specific antibody titers in sera of immunised mice.

6. Conclusion

Liposomes offer an interesting alternative to aluminium hydroxide and other adjuvant systems, especially in those cases, when the induction of cell immunity is of importance.

Various methods for preparation of proteoliposomal vaccination particles are developed and metallochelating liposomes represent the newest and versatile approach to this problem. New nonpyrogenic lipophilic derivatives of norAbu-MDP have been shown to be potent adjuvants for week antigens like recombinant nonlipidized OspC and are suitable adjuvant components, together with other synthetic adjuvants like MPLA and CpG oligos, for construction of vaccines based on various liposomal platforms. Adjuvant potency of these new MDP analogues is comparable to MDP but they lack side effects related to MDP, like strong pyrogenicity and flu like syndromes. Also other synthetic adjuvans like (MPL-A and CpG oligonucleotides) or could be entrapped in metallochelating liposomes alone or in their combinations to precisely tailor the immune response towards particular antigen, especially if the Th1 response is of interest.

Moreover, liposomes are also applicable for non-invasive routes of application as mucosal and transdermal ones, and if rationally designed and applied with new synthetic adjuvants derived from PAMP, they are able to steer the immune system towards desired effective response. The most important observation was that in all vaccinated animals liposomal based vaccines did not induce any side effects. Application of modern physico-chemical and microscopic methods for study of the structure and stability of proteoliposomal vaccination particles is indispensable part of successful development of modern safe and effective vaccines.

7. Acknowledgment

This work was supported by grants: GAČR P304/10/1951; MSM 6198959223; MZE 0002716202; KAN 200520703 AVČR and KAN 200100801.

8. References

Adu-Bobie, J., Capecchi, B., Serruto, D., Rappuoli, R. & Pizza, M. (2003). Two years into reverse vaccinology. Vaccine, Vol.21, No.7-8, pp. 605-610, ISSN 0264-410X

Agostini, L., Martinon, F., Burns, K., McDermott, M.F., Hawkins, P.N. & Tschopp, J. (2004). NALP3 forms an IL-l beta-Processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. Immunity, Vol.20, No.3, pp. 319-325, ISSN 1074-7613

Allen, T.M., (1997). Liposomes - Opportunities in drug delivery. Drugs, Vol.54, No.4, pp. 8-14. ISSN 0012-6667

Allison, A.C. & Gregoriadis, G. (1974). Liposomes As Immunological Adjuvants. Nature, Vol.252, No.5480, pp. 252, ISSN 0028-0836

Altin, J.G. & Parish, C.R. (2006). Liposomal vaccines - targeting the delivery of antigen. Methods, Vol.40, No.1, pp. 39-52, ISSN 1046-2023

Anderson, P.M., (2006). Liposomal muramyl tripeptide phosphatidyl ethanolamine: ifosfamide-containind chemotherapy in osteosarcoma. Future Oncol, Vol.2, No.3, pp. 333-343

Anderson, P.M., Tomaras, M. & McConnell, K. (2010). Mifamurtide in Osteosarcoma-A Practical Review. Drugs of Today, Vol.46, No.5, pp. 327-337, ISSN 1699-3993

Azuma, I. & Seya, T. (2001). Development of immunoadjuvants for immunotherapy of cancer. *International Immunopharmacology,* Vol.1, No.7, pp. 1249-1259, ISSN 1567-5769

Bangham, A.D., Standish, M.M. & Watkins, J.C. (1965). Diffusion of Univalent Ions Across Lamellae of Swollen Phospholipids. *Journal of Molecular Biology,* Vol.13, No.1, pp. 238-252, ISSN 0022-2836

Barnadas-Rodriguez, R. & Sabes, M. (2001). Factors involved in the production of liposomes with a high-pressure homogenizer. *International Journal of Pharmaceutics,* Vol.213, No.1-2, pp. 175-186, ISSN 0378-5173

Berger, N., Sachse, A., Bender, J., Schubert, R. & Brandl, M. (2001). Filter extrusion of liposomes using different devices: comparison of liposome size, encapsulation efficiency, and process characteristics. *International Journal of Pharmaceutics,* Vol.223, No.1-2, pp. 55-68, ISSN 0378-5173

Chikh, G.G., Li, W.M., Schutze-Redelmeier, M.P., Meunier, J.C. & Bally, M.B. (2002). Attaching histidine-tagged peptides and proteins to lipid-based carriers through use of metal-ion-chelating lipids. *Biochimica et Biophysica Acta-Biomembranes*, Vol.1567, No.1-2, pp. 204-212, ISSN 0005-2736

Creaven, P.J., Cowens, J.W., Brenner, D.E., Dadey, B.M., Han, T., Huben, R., Karakousis, C., Frost, H., Lesher, D., Hanagan, J., Andrejcio, K. & Cushman, M.K. (1990). Initial Clinical-Trial of the Macrophage Activator Muramyl Tripeptide Phosphatidylethanolamine Encapsulated in Liposomes in Patients with Advanced Cancer. *Journal of Biological Response Modifiers,* Vol.9, No.5, pp. 492-498, ISSN 0732-6580

Cullis, P.R., Hope, M.J., Bally, M.B., Madden, T.P. & Mayer, L.D. (1987). Liposomes as pharmaceuticals, In: *Liposomes: From Biophysics to Therapeutics,* Ostro, M.J., (Ed.), 39-69, Dekker, ISBN 0-8247-7762-X, New York/Basel, USA/Switzerland

Darcissac, E.C.A., Vidal, V., Guillaume, M., Thebault, J.J. & Bahr, G.M. (2001). Clinical tolerance and profile of cytokine induction in healthy volunteers following the simultaneous administration of IFN-alpha and the synthetic immunomodulator murabutide. *Journal of Interferon and Cytokine Research,* Vol.21, No.9, pp. 655-661, ISSN 1079-9907

Earnhart, C.G., Buckles, E.L. & Marconi, R.T. (2007). Development of an OspC-based tetravalent, recombinant, chimeric vaccinogen that elicits bactericidal antibody against diverse Lyme disease spirochete strains. *Vaccine*, Vol.25, No.3, pp. 466-480, ISSN 0264-410X

Earnhart, C.G. & Marconi, R.T. (2007). An octavalent Lyme disease vaccine induces antibodies that recognize all incorporated OspC type-specific sequences. *Human Vaccines,* Vol.3, No.6, pp. 281-289, ISSN 1554-8619

Erdile, L.F. & Guy, B. (1997). OspA lipoprotein of Borrelia burgdorferi is a mucosal immunogen and adjuvant. *Vaccine,* Vol.15, No.9, pp. 988-995, ISSN 0264-410X

Fortin, A., Shahum, E., Krzystyniak, K. & Therien, H.M. (1996). Differential activation of cell-mediated immune functions by encapsulated and surface-linked liposomal antigens. *Cellular Immunology,* Vol.169, No.2, pp. 208-217, ISSN 0008-8749

Gilmore, R.D., Bacon, R.M., Carpio, A.M., Piesman, J., Dolan, M.C. & Mbow, M.L. (2003). Inability of outer-surface protein C (OspC)-primed mice to elicit a protective anamnestic immune response to a tick-transmitted challenge of Borrelia burgdorferi. *Journal of Medical Microbiology,* Vol.52, No.7, pp. 551-556, ISSN 0022-2615

Gilmore, R.D., Kappel, K.J., Dolan, M.C., Burkot, T.R. & Johnson, B.J.B. (1996). Outer surface protein C (OspC), but not P39, is a protective immunogen against a tick-transmitted Borrelia burgdorferi challenge: Evidence for a conformational protective epitope in OspC. *Infection and Immunity,* Vol.64, No.6, pp. 2234-2239, ISSN 0019-9567

Gilmore, R.D. & Mbow, M.L. (1999). Conformational nature of the Borrelia burgdorferi B31 outer surface protein C protective epitope. *Infection and Immunity*, Vol.67, No.10, pp. 5463-5469, ISSN 0019-9567

Girardin, S.E. & Philpott, D.J. (2004). The role of peptidoglycan recognition in innate immunity. *European Journal of Immunology*, Vol.34, No.7, pp. 1777-1782, ISSN 0014-2980

Gluck, R., Althaus, B., Berger, R., Just, M. & Cryz, S.J. (1992). Development, safety and immunogeincity of new inactivated hepatitis A vaccines - Effect of adjuvants, In: *Travel Medicine 2,* Lobel, H.O., Steffen, R., Kozarsky, P.E., (Eds.), 135-136, International Society of Travel Medicine , ISBN 0-323-03453-5, Atlanta, USA

Gregoriadis, G. (Ed.). (1992). *Liposome Technology*, CRC Press, Inc., ISBN 9780849367090, Boca Raton, USA

Gregoriadis, G. (1995). Engineering liposomes for drug delivery: Progress and problems. *Trends in Biotechnology*, Vol.13, No.12, pp. 527-537, ISSN 0167-7799

Gregoriadis, G., Davis, D. & Davies, A. (1987). Liposomes As Immunological Adjuvants - Antigen Incorporation Studies. *Vaccine*, Vol.5, No.2, pp. 145-151, ISSN 0264-410X

Hope, M.J., Bally, M.B., Webb, G. & Cullis, P.R. (1985). Production of Large Unilamellar Vesicles by A Rapid Extrusion Procedure - Characterization of Size Distribution, Trapped Volume and Ability to Maintain A Membrane-Potential. *Biochimica et Biophysica Acta*, Vol.812, No.1, pp. 55-65, ISSN 0006-3002

Ichihara, N., Kanazawa, R., Sasaki, S., Ono, K., Otani, T., Yamaguchi, F. & Une, T. (1988). Phase I study and clinical pharmacological study of muroctasin. *Arzneimittel-forschung/Drug Research*, Vol.38-2, No.7A, pp. 1043-1069, ISSN 0004-4172

Ikushima, M., Matsui, K., Yamada, F., Kawahashi, S. & Nishikawa, A. (2000). Specific immune response to a synthetic peptide derived from outer surface protein C of Borrelia burgdorferi predicts protective borreliacidal antibodies. *Fems Immunology and Medical Microbiology*, Vol.29, No.1, pp. 15-21, ISSN 0928-8244

Keefer, M.C., Graham, B.S., McElrath, M.J., Matthews, T.J., Stablein, D.M., Corey, L., Wright, P.F., Lawrence, D., Fast, P.E., Weinhold, K., Hsieh, R.H., Chernoff, D., Dekker, C. & Dolin, R. (1996). Safety and immunogenicity of Env 2-3, a human immunodeficiency virus type 1 candidate vaccine, in combination with a novel adjuvant, MTP-PE/MF59. *Aids Research and Human Retroviruses*, Vol.12, No.8, pp. 683-693, ISSN 0889-2229

Keitel, W., Couch, R., Bond, N., Adair, S., Vannest, G. & Dekker, C. (1993). Pilot Evaluation of Influenza-Virus Vaccine (Ivv) Combined with Adjuvant. *Vaccine*, Vol.11, No.9, pp. 909-913, ISSN 0264-410X

Krupka, M., Raska, M., Belakova, J., Horynova, M., Novotny, R. & Weigl, E. (2007). Biological aspects of Lyme disease spirochetes: unique bacteria of the Borrelia burgdorferi species group. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*, Vol.151, No.2, pp. 175-186,

Lovrich, S.D., Jobe, D.A., Schell, R.F. & Callister, S.M. (2005). Borreliacidal OspC antibodies specific for a highly conserved epitope are immunodominant in human lyme disease and do not occur in mice or hamsters. *Clinical and Diagnostic Laboratory Immunology*, Vol.12, No.6, pp. 746-751, ISSN 1071-412X

Malliaros, J., Quinn, C., Arnold, F.H., Pearse, M.J., Drane, D.P., Stewart, T.J. & Macfarlan, R.I. (2004). Association of antigens to ISCOMATRIX (TM) adjuvant using metal chelation leads to improved CTL responses. *Vaccine*, Vol.22, No.29-30, pp. 3968-3975, ISSN 0264-410X

Masek, J., Bartheldyova, E., Korvasova, Z., Skrabalova, M., Koudelka, S., Kulich, P., Kratochvilova, I., Miller, A.D., Ledvina, M., Raska, M. & Turanek, J. (2011). Immobilization of histidine-tagged proteins on monodisperse metallochelation liposomes: Preparation and study of their structure. *Analytical Biochemistry*, Vol.408, No.1, pp. 95-104, ISSN 0003-2697

Masek, J., Bartheldyova, E., Turanek-Knotigova, P., Skrabalova, M., Korvasova, Z., Plockova, J., Koudelka, S., Skodova, P., Kulich, P., Krupka, M., Zachova, K., Czernekova, L., Horynova, M., Kratochvilova, I., Miller, A.D., Zyka, D., Michalek, J., Vrbkova, J., Sebela, M., Ledvina, M., Raska, M. & Turanek, J. (2011). Metallochelating liposomes with associated lipophilised norAbuMDP as biocompatible platform for construction of vaccines with recombinant His-tagged antigens: Preparation, structural study and immune response towards rHsp90. *J Control Release*, Epub Ahead of print

McDonald, C., Inohara, N. & Nunez, G. (2005). Peptidoglycan signaling in innate immunity and inflammatory disease. *Journal of Biological Chemistry*, Vol.280, No.21, pp. 20177-20180, ISSN 0021-9258

Nerome, K., Yoshioka, Y., Ishida, M., Okuma, K., Oka, T., Kataoka, T., Inoue, A. & Oya, A. (1990). Development of A New Type of Influenza Subunit Vaccine Made by Muramyldipeptide Liposome - Enhancement of Humoral and Cellular Immune-Responses. *Vaccine*, Vol.8, No.5, pp. 503-509, ISSN 0264-410X

Pal, U., de Silva, A.M., Montgomery, R.R., Fish, D., Anguita, J., Anderson, J.F., Lobet, Y. & Fikrig, E. (2000). Attachment of Borrelia burgdorferi within Ixodes scapularis mediated by outer surface protein A. *Journal of Clinical Investigation,* Vol.106, No.4, pp. 561-569, ISSN 0021-9738

Parant, M., Parant, F., Chedid, L., Yapo, A., Petit, J.F. & Lederer, E. (1979). Fate of the Synthetic Immunoadjuvant, Muramyl Dipeptide (C-14-Labeled) in the Mouse. *International Journal of Immunopharmacology*, Vol.1, No.1, pp. 35-41, ISSN 0192-0561

Perrett, S., Golding, M. & Williams, W.P. (1991). A Simple Method for the Preparation of Liposomes for Pharmaceutical Applications - Characterization of the Liposomes. *Journal of Pharmacy and Pharmacology*, Vol.43, No.3, pp. 154-161, ISSN 0022-3573

Rappuoli, R. (2000). Reverse vaccinology. *Current Opinion in Microbiology*, Vol.3, No.5, pp. 445-450, ISSN 1369-5274

Rigaud, J.L. & Levy, D. (2003). Reconstitution of membrane proteins into liposomes. *Liposomes, PT A,* Vol. 372, pp. 65-86, ISSN 0076-6879

Ruger, R., Muller, D., Fahr, A. & Kontermann, R.E. (2006). In vitro characterization of binding and stability of single-chain Fv Ni-NTA-liposomes. *Journal of Drug Targeting*, Vol.14, No.8, pp. 576-582, ISSN 1061-186X

Schneider, T., Sachse, A., Rossling, G. & Brandl, M. (1995). Generation of contrast-carrying liposomes of defined size with a new continous high-pressure extrusion method. International Journal of Pharmaceutics, Vol.117, No.1, pp. 1-12, ISSN 0378-5173

Schubert, R. (2003). Liposome preparation by detergent removal. Liposomes, PT A, Vol.367, pp. 46-70, ISSN 0076-6879

Shahum, E. & Therien, H.M. (1988). Immunopotentiation of the Humoral Response by Liposomes - Encapsulation Versus Covalent Linkage. *Immunology*, Vol.65, No.2, pp. 315-317, ISSN 0019-2805

Shek, P.N. & Heath, T.D. (1983). Immune-Response Mediated by Liposome-Associated Protein Antigens .3. Immunogenicity of Bovine Serum-Albumin Covalently Coupled to Vesicle Surface. *Immunology*, Vol.50, No.1, pp. 101-106, ISSN 0019-2805

Small, D.M. (1971). Chemistry, In: *The Bile Acids: Chemistry, Physiology, and Metabolism,* Nair, P.P., Kritchevsky, D., (Ed.), 249-356, Plenum Press, New York, USA

Stadler, K., Masignani, V., Eickmann, M., Becker, S., Abrignani, S., Klenk, H.D. & Rappuoli, R. (2003). SARS - Beginning to understand a new virus. *Nature Reviews Microbiology*, Vol.1, No.3, pp. 209-218, ISSN 1740-1526

Tamura, M., Yoo, Y.C., Yoshimatsu, K., Yoshida, R., Oka, T., Ohkuma, K., Arikawa, J. & Azuma, I. (1995). Effects of Muramyl Dipeptide Derivatives As Adjuvants on the Induction of Antibody-Response to Recombinant Hepatitis-B Surface-Antigen. *Vaccine*, Vol.13, No.1, pp. 77-82, ISSN 0264-410X

Tilly, K., Krum, J.G., Bestor, A., Jewett, M.W., Grimm, D., Bueschel, D., Byram, R., Dorward, D., VanRaden, M.J., Stewart, P. & Rosa, P. (2006). Borrelia burgdorferi OspC protein required exclusively in a crucial early stage of mammalian infection. *Infection and Immunity*, Vol.74, No.6, pp. 3554-3564, ISSN 0019-9567

Tilly, K., Rosa, P.A. & Stewart, P.E.(2008). Biology of infection with Borrelia burgdorferi. *Infectious Disease Clinics of North America*, Vol.22, No.2, 217-+, ISSN 0891-5520

Traub, S., von Aulock, S., Hartung, T. & Hermann, C. (2006). MDP and other muropeptides - direct and synergistic effects on the immune system. *Journal of Endotoxin Research*, Vol.12, No.2, pp. 69-85, ISSN 0968-0519

Tsubura, E., Nomura, T., Niitani, H., Osamura, S., Okawa, T., Tanaka, M., Ota, K., Nishikawa, H., Masaoka, T., Fukuoka, M., Horiuchi, A., Furuse, K., Ito, M., Nagai, K., Ogura, T., Kozuru, M., Hara, N., Hara, K., Ichimaru, M. & Takatsuki, K. (1988). Restorative Activity of Muroctasin on Leukopenia Associated with Anticancer Treatment. *Arzneimittel-Forschung/Drug Research*, Vol.38-2, No.7A, pp. 1070-1074, ISSN 0004-4172

Tsujimoto, M., Kotani, S., Kinoshita, F., Kanoh, S., Shiba, T. & Kusumoto, S. (1986). Adjuvant Activity of 6-O-Acyl-Muramyldipeptides to Enhance Primary Cellular and Humoral Immune-Responses in Guinea-Pigs - Adaptability to Various Vehicles and Pyrogenicity. *Infection and Immunity*, Vol.53, No.3, pp. 511-516, ISSN 0019-9567

Turanek, J. (1994). Fast-Protein Liquid-Chromatography System As A Tool for Liposome Preparation by the Extrusion Procedure. *Analytical Biochemistry*, Vol.218, No.2, pp. 352-357, ISSN 0003-2697

Turanek, J., Kasna, A., Koudela, B., Ledvina, M. & Miller, A.D. (2005). Stimulation of innate immunity in newborn kids against Cryptosporidium parvum infection-challenge by intranasal/per-oral administration of liposomal formulation of N-L 18-norAbu-GMDP adjuvant. *Parasitology*, Vol.131, No.5, pp. 601-608, ISSN 0031-1820

Uchida, T. & Taneichi, M. (2008). Clinical application of surface-linked liposomal antigens. *Mini-Reviews in Medicinal Chemistry*, Vol.8, No.2, pp. 184-192, ISSN 1389-5575

van Broekhoven, C.L., Parish, C.R., Demangel, C., Britton, W.J. & Altin, J.G. (2004). Targeting dendritic cells with antigen-containing liposomes: A highly effective procedure for induction of antitumor immunity and for tumor immunotherapy. *Cancer Research*, Vol.64, No.12, pp. 4357-4365, ISSN 0008-5472

Vesely, D.L., Fish, D., Shlomchik, M.J., Kaplan, D.H. & Bockenstedt, L.K. (2009). Langerhans Cell Deficiency Impairs Ixodes scapularis Suppression of Th1 Responses in Mice. *Infection and Immunity*, Vol.77, No.5, pp. 1881-1887, ISSN 0019-9567

Wack, A. & Rappuoli, R. (2005). Vaccinology at the beginning of the 21st century. *Current Opinion in Immunology*, Vol.17, No.4, pp. 411-418, ISSN 0952-791

Walter, A., Kuehl, G., Barnes, K. & VanderWaerdt, G. (2000). The vesicle-to-micelle transition of phosphatidylcholine vesicles induced by nonionic detergents: effects of sodium chloride, sucrose and urea. *Biochimica et Biophysica Acta-Biomembranes*, Vol.1508, No.1-2, pp. 20-33, ISSN 0005-2736

Weis, J.J., Ma, Y. & Erdile, L.F. (1994). Biological activities of native and recombinant Borrelia burgdorferi outer surface protein A: dependence on lipid modification. *Infection and Immunity,* Vol.62, No.10, pp. 4632-4636, ISSN 0019-9567

Woodle, M.C. & Papahadjopoulos, D. (Eds.). (1988). Methods in Enyzmology, Academic Press, San Diego, USA

Zumbuehl, O. & Weder, H.G. (1981). Liposomes of Controllable Size in the Range of 40 to 180 Nm by Defined Dialysis of Lipid-Detergent Mixed Micelles. Biochimica et Biophysica Acta, Vol.640, No.1, pp. 252-262, ISSN 0006-3002

28

iPS Cells: Born-Again Stem Cells for Biomedical Applications

Ambrose Jon Williams and Vimal Selvaraj
Cornell University,
USA

1. Introduction

1.1 Stemness

The fertilized egg, also known as the zygote, is a cell of total potential and plasticity and gives rise to the embryo and extra-embryonic tissues, and ultimately the whole adult organism. This property has since been termed *totipotency*, although the transition from fertilized egg to differentiated cells of the adult tissues (somatic cells) is not direct, progressing instead through lineages of successively more differentiated and committed intermediates towards the final cell type. Thus, the zygote gives rise to the trophoblast cells and inner cell mass (ICM) of the blastocyst stage embryo, the ICM gives rise to the primordial cells committed to the ectodermal, mesodermal and endodermal lineages. To give an example lineage, the ectoderm cells give rise to the neural crest stem cells, then neural stem cells, oligodendrocyte precursors, and finally oligodendrocytes which myelinate and form the white matter of the central nervous system. Each step is more specialized and less plastic than the base or 'stem' of the branch before it. Unlike the totipotent zygote, these 'stem cells' retain the ability to self-renew in addition to their plasticity. In addition to giving rise to somatic tissues during embryogenesis, the biological role of stem cells in an adult organism is to regenerate tissues lost to injury, disease or age.

1.2 History

The first stem cells discovered were the originator cells of teratomas, a rare tumor that comprised of multiple tissue types and was associated with embryonal carcinoma (EC). These EC stem cells have the then-unusual ability to self-renew indefinitely, as well as give rise to tissues from each of the three germinal layers (Kleinsmith and Pierce, 1964), a property termed *pluripotency.* Since then, many more stem cell types of varying potency have been discovered, including two more pluripotent cell types: the embryonic stem cell (ESC) is a non-cancerous analogue of the EC stem cell which is derived instead from the ICM of blastocyst stage embryos (Evans and Kaufman, 1981), and most recently induced pluripotent stem cells (iPSCs) which are produced from somatic cells that have been reprogrammed to a pluripotent state (Takahashi and Yamanaka, 2006). Pluripotent cells have since been demonstrated to have enormous potential for regenerative medicine, disease research and genetic engineering.

2. Applications of pluripotent stem cells

2.1 Animals from pluripotent cells

ESCs are the prototypical pluripotent stem cell and thus the most thoroughly characterized. They can self-renew indefinitely and are effectively immortal in cell culture. Although they lack the self-organizing capabilities of the fertilized egg, they can form any tissue in the adult organism as demonstrated by two key studies: injecting ESCs into blastocysts gives rise to chimaeric animals with tissues contributed by the injected ESCs as well as the original ICM (Moustafa and Brinster, 1972); injection of ESCs into blastocysts that have been rendered tetraploid (four genome copies, and therefore genomically incapable of forming a complete organism) produces animals wholly derived from the injected ESCs (Eggan et al., 2002). The latter technique is possible because tetraploid blastocysts retain the structural organization of a normal blastocyst, and although the tetraploid ICM will inevitably die out or senesce (and be replaced by the injected cells), the trophoblast component retains its function despite tetraploidy since trophoblasts eventually fuse and become polyploid anyway upon embryonic implantation into the maternal uterus. These properties are shared with all pluripotent cells, EC cells injected into blastocysts can also give rise to chimaeric animals (Mintz and Illmensee, 1975). Because of this potential, ESCs very quickly became a focus of applied research.

2.2 ESCs in genetic engineering and animal disease modeling

Modern reproduction techniques make it possible for a single ESC to give rise to a whole animal, greatly simplifying the process of genetically engineering animals. Previously, animals were bred extensively to isolate beneficial random mutations fertilized eggs were microinjected with DNA for random genomic integration (Gordon et al., 1980), or engineered animals were derived from nuclear-cloned somatic cells that had been engineered to the desired genotype; such a technique was used to generate cattle that lacked the prion protein and were thus made completely immune to bovine spongiform encephalopathy (BSE; mad cow disease, which transmits to humans as the variant Creutzfeldt-Jakob)(Richt et al., 2006). ESCs are easier to genetically engineer due to their infinite self-renewability, allowing a very small number of drug- or marker-selected cells to regenerate a whole culture or stable cell line. This technique has been used to generate a variety of mouse genetic models including sickle cell disease (Wu et al., 2006), thalassemia (Ciavatta et al., 1995), microcephaly (Pulvers et al., 2010), and T-cell lymphoma (Pechloff et al., 2010), as well as a p53-knockout rat for cancer research (Tong et al., 2010).

2.3 *In vitro* disease modeling using pluripotent cells

A major obstacle to disease research is the difficulty of acquiring diseased cells for study, usually because they are difficult to obtain from a living patient. For example, neurons are not easily obtained from a patient afflicted with Down syndrome, making detailed cell biology study of the neuronal basis for mental retardation impossible, and limiting our understanding of this disorder to more superficial behavioral neurological or postmortem pathological descriptions. However, a Down syndrome human ESC line as well as lines for other chromosomal trisomies have recently been derived (Biancotti et al., 2010), as has a human ESC line homozygous for Sickle Cell Disease (Pryzhkova et al., 2010). All were generated from embryos rejected by preimplantation genetic diagnosis (PGD) screens following *in vitro* fertilization (IVF). These lines allow cell culture study of diseased neurons, or any other cell type, by differentiating diseased ESCs into any cell type of interest; however researchers are still limited by the small number of diseased human ESC lines available.

Cloned embryos can be derived from adult cells using somatic cell nuclear transfer (SCNT), a technique made famous by the cloning of the sheep Megan, Morag and Dolly in the 1990s (Wilmut et al., 1997). It has been proposed that new diseased human ESC lines can be derived using this technique to make cloned embryos from diseased patients, and then harvesting them to create novel diseased ESC lines for disease study. At the time of this writing, SCNT for this application (Therapeutic Cloning) is currently legal in the United States and the European Union, but its legal status in these states as well as elsewhere across the world has been subject to numerous prior and continuing legal challenges. Although several large organizations continue to research this technology, it has been supplanted in recent years by alternate techniques for deriving patient-specific pluripotent stem cells.

2.4 Therapeutic potential of pluripotent stem cells

Pluripotent stem cells have been studied as, and shown great potential to be, a source of cell replacement therapies in a myriad of disease and injury models. Several human ESC lines have been differentiated into high-purity cardiomyocyte cultures that improve cardiac performance when transplanted into infracted rat hearts (Caspi et al., 2007). ESCs have also been differentiated into neural precursors and neurons including dopaminergic neurons which reverse the disease progression of Parkinsonian rats (Yang et al., 2008). In a model of spinal cord injury, ESC-derived oligodendrocytes transplanted into crushed rat spinal cords successfully restored locomotive function to the animals. Pancreatic beta cells, the insulin-secreting cells whose absence causes type I diabetes mellitus, have also been derived from ESCs and cure the diabetic phenotype of the mouse streptozotocin-induced model of diabetes upon transplantation (Kim et al., 2003). These are but a choice selection of the vast amount of scientific literature detailing the regenerative potential of ESCs.

At the time of this writing, two clinical trials are underway for ESC-based regenerative therapies in humans: an evaluation of human ESC-derived oligodendrocyte precursors to rescue neurologically complete spinal cord injury conducted by Geron Corporation, and ESC-derived retinal-pigmented epithelium for treatment of macular degeneration and Stargardt's macular dystrophy, which are major causes of blindness, conducted by Advanced Cell Technology Incorporated. A third proposed clinical trial is currently in the approval process between the Food and Drug Administration and applicant California Stem Cell Incorporated for ESC-derived motor neurons as a cure for type I spinal muscular atrophy, the leading genetic cause of infant mortality. These trials represent the first step in the direct evaluation of the therapeutic potential of pluripotent stem cells in human patients.

2.5 Pitfalls and obstacles to the use of ESCs

Transplants of ESC-derived tissues and biological devices are just as subject to immune rejection as conventional organ transplants, even more so due to the limited selection of human ESC lines. Although the engineering of non-immunogenic ESCs has been the subject of many academic initiatives and company startups, ongoing clinical and preclinical research for ESC-therapies is focused, for the mean time, on immune-privileged regions of the body: specifically the brain, eye and spinal cord. A second scientific concern is the purity of ESC-derived transplants because of the hazard posed by contaminating undifferentiated

ESCs that, if transplanted, can proliferate and form teratomas. The elimination of these leftover ESCs has been approached by several strategies: purification of differentiated cells by labeling and cell sorting (Pruszak et al., 2007), the engineering of special "suicide gene"-containing ESCs (Schuldiner et al., 2003), and the treatment of cells to be transplanted with chemotherapeutics (Bieberich et al., 2004). The concomitant destruction of stem cells by anti-cancer therapies reflects the generalized similarity between stem cells and cancer cells [reviewed in (Reya et al., 2001)].

As many as seven human embryos are sacrificed for each new human ESC line derived (Thomson et al., 1998); while the ethics of this are philosophically subjective they have nonetheless given rise to numerous high-profile legal challenges to continued ESC research and funding. In addition, the patent on derivation of human ESC lines is held by the Wisconsin Alumni Research Foundation. Until its expiration in 2016, commercial users wishing to use Wisconsin ESC ("WiCell") technologies might also be required to pay a royalty.

3. Induced pluripotency

3.1 Discovery

The laboratory of Shinya Yamanaka demonstrated in 2006 that somatic cells can be reprogrammed back to a primordial phenotype functionally identical to ESCs, and termed these reprogrammed cells induced pluripotent stem cells (iPSCs)(Takahashi and Yamanaka, 2006). These iPSCs have a morphology, growth and gene expression characteristics that are indistinguishable from ESCs. Like ESCs they also form teratomas consisting of tissues from all three germ layers when injected into immunodeficient animals (Takahashi and Yamanaka, 2006), and give rise to entire animals when injected into tetraploid blastocysts (Kang et al., 2009). iPSCs also have stable telomere lengths like ESCs (Marion et al., 2008) as well as an epigenetic state reflecting reversion back to pluripotency, although traces of the donor cell's epigenetic imprint are retained in early-passage iPSCs (Kim et al., 2010).

Pluripotency is typically induced by overexpressing in somatic cells the stem cell genes Oct3/4, Sox2, cMyc and Klf4 (Takahashi and Yamanaka, 2006) (collectively termed the Yamanaka factors) or by an alternate combination of Oct3/4, Sox2, Nanog and Lin28 (Yu et al., 2007) (the Thomson factors; this repertoire has not been extensively replicated in the literature). Retroviruses or lentiviruses are the standard vectors for inserting and over-expressing these transgenes for a period of 2-3 weeks. During which the formation of early ESC-like colonies are observed [Figure 1]. These colonies stain positively for alkaline phosphatase, a marker which distinguishes undifferentiated cells from fibroblasts, and when clonally selected and propagated they express the ESC markers SSEA-1 and Oct3/4 and assume a phenotype indistinguishable from ESCs (Takahashi and Yamanaka, 2006).

3.2 Molecular mechanisms of induced pluripotency

Induced pluripotency is a remarkably successful technique, although our understanding of the underlying mechanisms are limited. The Yamanaka combination of reprogramming factors wasn't arrived at by a serendipitous leap of understanding, but instead careful and methodical experimentation. When Yamanaka sought to reprogram skin cells into ESCs, he began with a list of 24 candidate genes identified by extensive review of ESC literature. Overexpression in fibroblasts for two weeks gave rise to ESC-like colonies expressing the pluripotency marker Fbxo15. After a yearlong process of elimination, his lab was able to replicate this result with just 4 genes: Oct3/4, Sox2, cMyc and Klf4 (Takahashi and Yamanaka, 2006).

It has been known for some time that Oct3/4, Sox2 and Nanog comprise the core of the pluripotency transcriptional network, as the deficiency in either one causes ES cells to lose pluripotency (Avilion et al., 2003; Mitsui et al., 2003; Nichols et al., 1998). It is interesting to note, however, that too much Oct3/4 or Sox2 can also disrupt pluripotency. As little as a two-fold excess in either causes ESCs to differentiate (Kopp et al., 2008; Niwa et al., 2000).

Oct3/4, Sox2 and Nanog all occupy each others' promoters, and more than 90% of promoters bound to by Oct3/4 and Sox2 are also occupied by Nanog (Boyer et al., 2005). Although all three are required for pluripotency, Nanog overexpression is not required for induced pluripotency. Adding Nanog to the mix, however, increases reprogramming, as does combining the Yamanaka and Thomson reprogramming repertoires (Liao et al., 2008).

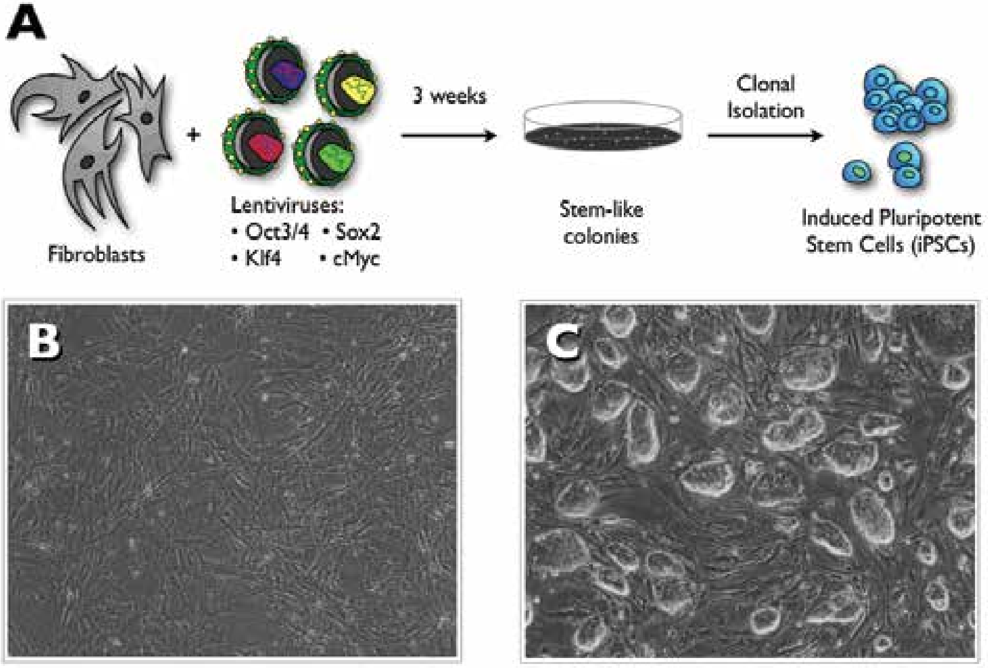


Fig. 1. Generation of iPSCs. A: Although there are a number of ways to generate iPS Cells, the model reprogramming experiment uses lentiviral vectors to integrate the Yamanaka factor transgenes into skin fibroblasts and over-express them for a period of three weeks. After this period of time, stem cell-like colonies become apparent in the reprogrammed culture which, following selection and characterization, will give rise to stable iPS Cell lines. B: Fibroblasts before reprogramming have typical morphology and grow in confluent cell monolayers. C: iPS Cells, however, grow in dense, elevated and round colonies with the characteristic "glass edge." They are microscopically indistinguishable from ESC cultures.

The use of the RNA-binding protein Lin28 as one of the Thomson factors suggested a role for microRNAs during the reprogramming process, and since then a number of pluripotency-regulating microRNAs have been identified (Zhong et al., 2010). Of particular interest is the miR-302 family of microRNAs, which induce stem cell-like plasticity when overexpressed in skin cancer cells (Lin et al., 2008).

Reprogramming to pluripotency is accompanied by the demethylation of promoter regions of known pluripotency genes, and this is observed in both iPSCs (Park et al., 2008b) and cloned embryos generated by SCNT (Lan et al., 2010). In partially or incompletely reprogrammed cells generated by either method, this demethylation is incomplete (Bourchis et al., 2001; Takahashi and Yamanaka, 2006). Likewise, chromatin alterations are also observed in reprogrammed cells as well as cloned embryos, and both of these processes are enhanced by histone deacetylase inhibitors (Han et al., 2010). Histone demethylation, particularly at the promoters of pluripotency genes, is also observed but it is not understood how this occurs during induced pluripotency.

3.3 The cancer generalization

Despite the success and reproducibility of this reprogramming technique, the permanent integration of additional copies of stem cell genes with high expression promoters poses a significant oncogenic hazard; in the earlier studies one in five chimaeric mice derived from iPSCs died from tumors resulting from spontaneous reactivation of reprogramming genes (Okita et al., 2007). Although the Yamanaka and Thomson factors are highly expressed in ESCs, they are either oncogenes themselves or associated with a poor clinical outcome when detected in cancers. cMyc specifically is one of the most well-characterized oncogenes, but Oct3/4 expression in animals also results in death due to extreme proliferation of undifferentiated progenitors (Hochedlinger et al., 2005). [Oct3/4 actually has no known role outside of pluripotent biology and when conditionally deleted in adult animals results in no detectable phenotype or defect in healing (Lengner et al., 2007)]. Oct3/4 (Gidekel et al., 2003) and Sox2 (Gangemi et al., 2009) are associated with cancer cell proliferation and tumor progression, while Klf4 has been linked to an invasive progression and metastasis in epithelial cancers (Pandya et al., 2004).

While the carcinogenic hazard introduced by genetic insertion of the Yamanaka factors led to a search for alternative reprogramming techniques, the generalization that ESCs and iPSCs biologically resemble cancer cells gave rise to a new line of thought: emulating oncogenesis to enhance reprogramming. Both SV40 Large T Antigen and TERT have been shown to enhance reprogramming when included in the Yamanaka factor repertoire (Park et al., 2008b), as has the knockout or knockdown of the tumor suppressor p53 [simultaneously discovered by 5 separate groups and reviewed in (Krizhanovsky and Lowe, 2009)]. The finding that adult stem cell populations, including hematopoietic (Eminli et al., 2009), keratinocyte (Aasen et al., 2008) and neural (Kim et al., 2008) stem cells, reprogram more easily than the more differentiated cells further down their lineages is also consistent with the understanding that adult stem cells are most prone to becoming cancerous. Although these studies contribute greatly to our understanding of induced pluripotency and stem cell biology, incorporating them into current techniques to enhance reprogramming has, until quite recently, been impossible, as doing so would greatly enhance the oncogenic hazards.

In recent years, a number of alternative techniques have emerged that induce pluripotency without genomic integration of the Yamanaka factors. Plasmid vectors have been used to induce pluripotency (Okita et al., 2008), however this technique has low reprogramming efficiency and half of the putative iPSCs generated contained some form of genomic integration. Adenoviruses also achieve reprogramming at a lowered efficiency; however a fraction of reprogrammed cells displayed karyotypic abnormalities (Stadtfeld et al., 2008).

Two other approaches favor transgene insertion followed by excision upon completion of reprogramming but also have their shortfalls: a retrotransposon vector which very rarely completely excises from the genome (Woltjen et al., 2009), and Cre-Lox recombination leaving behind residual sequences (Soldner et al., 2009). Most promising, however, are two DNA-free methods of reprogramming: direct delivery of recombinant transcription factors to the donor cells (Kim et al., 2009) and transfection with modified mRNAs encoding the Yamanaka factors (Warren et al., 2010). Although these last two methods achieve reprogramming with reduced efficiency, they circumvent the permanent oncogenic hazard presented by genomic integration of the Yamanaka factors (and any supplemental genes as well) and are most likely to be used for translational iPSC applications.

3.4 Advantages of induced pluripotency

Because they are derived from somatic cells and thus genetically autologous to the donor, iPSCs circumvent most of the obstacles, which have prevented clinical implementation of ESC technology. Moreover, being patient-specific they are not subject to immune rejection, and because induced pluripotency is an embryo-free method of deriving new pluripotent cell lines they are not subject to the funding restrictions or ethical controversies on ESC-derivation. The field of iPSC research is unlikely to see the sort of legal challenges that ESC research has, having drawn endorsements from social conservatives including Republicans in America and the Catholic Church.

iPSCs have several advantages in addition to overcoming ESC-specific obstacles. On a technique level, iPSCs are easier to derive than ESCs, and iPSC lines have already been derived from several species for which no ESC lines exist (Esteban et al., 2009; Li et al., 2009; Tomioka et al., 2010; Wu et al., 2009). This is because the optimal conditions for deriving ESCs vary across species [for example, human ESCs are maintained with basic fibroblast growth factor, while mouse ESCs are maintained with leukemia inhibitory factor], while induced pluripotency is conserved across mammals. Practically, this means induced pluripotency can facilitate easier genetic engineering of animals, as the most consistent and controlled techniques involve engineering of pluripotent stem cells. iPSC-based engineering of cattle and pigs is therefore becoming a new focus of the field [reviewed in (Telugu et al., 2010)].

Induced pluripotency has also become the key technique for deriving diseased pluripotent cells. Whereas several ESC lines modeling karyotypic abnormalities (Biancotti et al., 2010) and sickle cell disease (Pryzhkova et al., 2010) exist, derivation of new diseased ESC lines is limited to PGD-rejected embryos from *in vitro* fertilization. However, in the three years since induced pluripotency was first described in humans, iPSC lines representative of a large number of genetic diseases have been derived including amyotrophic lateral sclerosis, adenosine deaminase severe combined immunodeficiency, Shwachman-Bodian-Diamond syndrome, Gaucher disease type III, muscular dystrophies, Parkinson's disease, Huntington's disease, juvenile-onset type-1 diabetes, Down's syndrome, Lesch-Nyhan syndrome carrier, Fanconi anemia, spinal muscular atrophy, long-QT syndrome and familial dysautonomia and LEOPARD syndrome (Dimos et al., 2008; Ebert et al., 2009; Park et al., 2008a; Raya et al., 2009; Lee et al., 2009; Carvajal-Vergara et al., 2010; Moretti et al., 2010). The utility of these lines is their ability to give rise to diseased tissue *in vitro* for study as well as drug testing [Figure 2], whereas previously, research on these diseases and many others has been impeded.

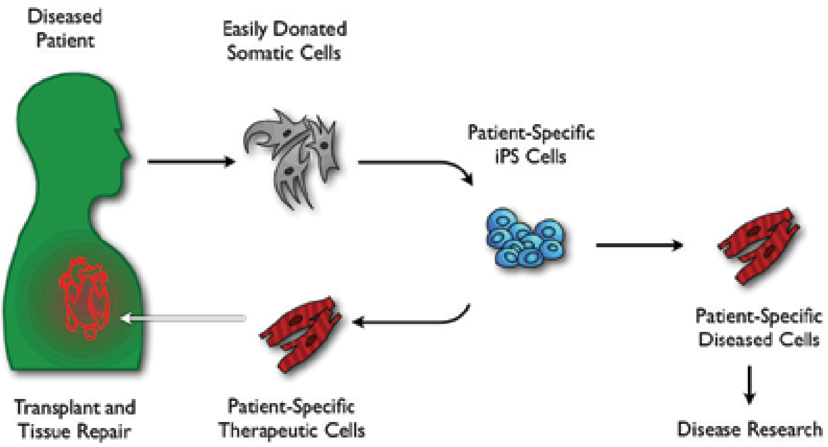

Diseased Patient
Easily Donated Somatic Cells
Patient-Specific iPS Cells
Patient-Specific Diseased Cells
Disease Research
Transplant and Tissue Repair
Patient-Specific Therapeutic Cells

Fig. 2. Biomedical applications of iPSCs. Because iPSCs are patient-specific, neither they nor any cells derived from them are subject to transplant rejection. In this paradigm, easily obtained somatic cells such as skin fibroblasts or blood lymphocytes are reprogrammed to iPSCs, and therapeutic cells are derived from them to regenerate damaged or diseased tissue. In the case of the patient with a genetic disease, patient-specific iPSCs allow the derivation of genetically diseased tissues, which then could be subject to downstream research.

4. Reprogramming techniques

4.1 SCNT and induced pluripotency

The first experiments to demonstrate that a terminally differentiated phenotype could be reprogrammed to pluripotency took place in the mid 1990s with the generation of cloned animals, the most famous of which were Dolly the sheep and her counterparts. Megan and Morag (Wilmut et al., 1997). In these experiments, donor cell nuclei were transplanted into enucleated oocytes, a technique termed somatic cell nuclear transfer (SCNT). Upon receiving a diploid genome, the re-nucleated oocytes assumed a zygote phenotype, which formed embryos and eventually whole animals. Although these experiments were critical in showing somatic cells can be epigenetically reprogrammed to pluripotency, animals cloned by SCNT were argued to begin life not with the long telomeres typical of early-stage embryonic cells, but instead with the aged and shortened telomeres of their nuclear donors- in the case of Dolly, an adult mammary epithelial cell. However, this claim remains controversial and has been refuted by a study that used cloned cattle (Lanza et al., 2000), and a recent demonstration that mice could be continuously cloned through 15 generations without any appreciable age-related issues or cloning efficiency (Thuan et al., 2010).

Yamanaka's demonstration, ten years later, that somatic cells could be directly reprogrammed into a pluripotent state by the over-expression of exogenous Oct3/4, Sox2, cMyc and Klf4, did not suffer from telomere problems because the resulting iPSCs express telomerase at high levels, quickly extending the reprogrammed cells' telomeres to ESC-like lengths. However, introducing these transgenes using integrating retroviral vectors also brought about the risk of oncogenesis, and subsequent studies have attempted to circumvent this through one of three approaches: inserting the transgenes in a transient or easily-excisable manner, reprogramming through the use of fewer transgenes, or reprogramming without the use of genome-integrating vectors.

4.2 Excision of reprogramming transgenes

Variations of the Yamanaka technique induce transient expression using non-integrating plasmids (Okita et al., 2008) or adenoviral vectors (Stadtfeld et al., 2008) and have produced isogenic pluripotent stem cells after many treatments, but with poor yields (in both cases, < 0.0005%). Poor reprogramming efficiency makes necessary large donor sample sizes and high multiplicities of vector transfection. Additionally, half of iPSCs reprogrammed by plasmids had genomic integration. Although the incidence of this is reduced by the use of Adenoviral vectors, karyotypic abnormalities were observed in 23% of iPSC lines produced. Nonetheless, these experiments demonstrate reprogramming without genome integration. Retrotransposon elements (Woltjen et al., 2009) and cre-lox excision (Soldner et al., 2009) have also been used to insert and subsequently remove integrated transgenes, however complete removal of transgenes occurred at only 2% efficiency in the retrotransposon-reprogrammed cells and 18% of excisions still left trace sequences in the genomic DNA. The efficiency of cre-lox excision was not demonstrated, but was consistently noted to leave proviral trace sequences.

4.3 Use of fewer transgenes

Reprogramming without using c-Myc has been demonstrated (Nakagawa et al., 2008) at 1.6% the efficiency of the standard technique of when all four genes are used. c-Myc is the most hazardous of the reprogramming factors, and its omission in this experiment reduced tumor formation in mouse chimeras past the observation period. The histone deacytelase inhibitor valproate has been shown to enhance reprogramming efficiency >100-fold (Huangfu et al., 2008a) and can substitute for the Klf4 transgene in reprogramming (Huangfu et al., 2008b), at reduced efficiency. The histone methyltransferase inhibitor BIX01294 and L-channel Ca2+ agonist BayK8644, when used together, can replace the Sox2 transgene (Shi et al., 2008) at reduced efficiency. No approach has yet replaced Oct3/4, which appears to be absolutely essential for reprogramming and is also the most reliable pluripotency marker. Reprogramming of different somatic cell types with fewer transgenes has yielded varying results. Neuronal stem cells (NSCs) have been reprogrammed to pluripotency using just Oct3/4 (Kim, 2009) and at improved efficiencies using Oct3/4 and Sox2 (Kim et al., 2008).

4.4 Direct treatment with pluripotency factors

Despite concerns with telomere length, SCNT remains the gold standard of induced pluripotency demonstrating that induced pluripotency can be accomplished through an entirely non-genetic approach (albeit with very low efficiency). This is echoed by recent experiments: somatic cells, when fused with pluripotent cells (either ESCs or iPSCs) always produce binucleate cells with both nuclei in a pluripotent state (Sumer et al., 2009). Furthermore, induced pluripotency has been demonstrated at extremely low (0.006%) efficiencies through the use of recombinant pluripotent factors (Kim et al., 2009) with multiple treatments over an 8-week period.

Permeablization of somatic cells using bacterial pore-forming toxins, and then treating them with whole cell extracts from the desired cell phenotype has shown some promise in epigenetic reprogramming, but so far no such experiments have generated induced pluripotent stem cells. In one such study, human 293T epithelial cells treated with T cell extracts, assume a phenotype similar to T-cells, and a neuronal-like phenotype when treated with neural precursor extracts (Hakelien et al., 2002). Numerous attempts have been made to induce pluripotency by using extracts from ESCs, ECs, iPSCs and oocytes (Taranger et al., 2005; Zhu et al., 2010), however none of these have succeeded in producing stable iPSC lines.

5. Critical thinking on induced pluripotency

Although iPSCs are highly similar to ESCs in biology and function, an increasing body of literature describes defects and subtle differences between the two pluripotent cell types. Effective characterization of these phenomena is critical and will likely give rise to new and more stringent criteria by which reprogrammed cells can be evaluated for suitability.

One of the earliest characterizations of reprogrammed cells was the demonstration of DNA demethylation on the promoters of genes involved in pluripotency, such as Oct3/4 and Nanog (Okita et al., 2007; Takahashi and Yamanaka, 2006). Although these promoters exhibited near-total demethylation, indicating an activation of gene transcription, the demethylation was rarely as complete as the pattern observed in ESCs, and this was particularly true for early-passage (newly-created) iPSCs. Although these differences were small, genome-wide analysis of methylation patterns outside of these specific genes has revealed significant errors in epigenetic reprogramming (Lister et al., 2011). On a genomic scale, iPSC and ESC methylomes are similar, but most iPSC methylomes analyzed had megabase-sized loci of aberrant DNA methylation, which persist long after reprogramming and even after differentiation. Some of these loci are shared among distinct iPSC lines, suggesting that certain regions of the methylomes are susceptible to aberrant and incomplete reprogramming. There have also been reports of differences in gene expression patterns between ESCs and iPSCs, although a recent study found that most of these differences are laboratory-specific and can be attributed to microenvironment differences in growth conditions from one laboratory to the next (Newman and Cooper, 2010).

DNA sequence defects have also been described in iPSCs. Early-passage iPSCs display a range of polymorphism in copy number variant (CNV) regions compared to their parental fibroblasts (Hussein et al., 2011). As with DNA methylation, it was also found that CNVs occurred more commonly in “fragile regions” of the genome. As CNVs arise from damaged DNA improperly repaired by homologous recombination, this phenomenon suggests DNA damage and replicative stress in cells undergoing reprogramming. However, while early-passage iPSCs contain significantly more CNVs, a vast majority of these mutations put the cells at a selective disadvantage. Mid- to late-passage iPSCs therefore lose CNVs and soon approach a genomic state highly similar to ESCs. Point mutations in specific genes have also been identified in iPSCs, however unlike CNVs these display a nonrandom pattern of enrichment, with a majority occurring in proto-oncogenes and tumor suppressors (Gore et al., 2011). Half of these mutations can be traced back to the parental fibroblasts, which harbor these mutations in low frequencies; however the other half most likely arise during reprogramming and, more importantly, the subsequent selection and propagation steps. Oncogenic mutations are generalized to give a selective advantage to pluripotent cells, and although an accumulation in oncogenic mutations has also been demonstrated in ESCs, this study still establishes the need for extensive genetic testing of iPSCs before they are to be used on a clinical scale.

6. Conclusions

Although induced pluripotency as a reprogramming technique currently brings significant concerns about carcinogenicity as well as genomic and epigenomic integrity, a significant portion of the ESC research community has jumped ship in recent years in order to study iPSCs. This is because of the exciting promise these cells hold, as well as the mainstream belief that the obstacles that come with them will be overcome. With applications in a variety of fields including regenerative therapies, disease modeling, animal cloning and genetic engineering, induced pluripotency is actively transforming the stem cell community. Given how young the field is, induced pluripotency has a surprisingly well-developed body of basic research, which has already contributed enormously to our understanding on developmental biology and epigenetics, as well as given us insights on a large number of modeled genetic diseases. Taken together, the current body of literature on induced pluripotency describes why it is a very exciting time to be a part of this field.

7. References

Aasen, T., Raya, A., Barrero, M., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilic, J., Pekarik, V., Tiscornia, G., Edel, M., Boue, S. & Izpisua-Belmonte, J. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nature Biotechnology,* Vol.26, No.11, (October 2008), pp. 1276-1284, ISSN 1087-0156

Avilion, A., Nicolis, S., Pevny, L., Perez, L., Vivian, N. & Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes and Development,* Vol.17, No.1, (January 2003), pp. 126-140, ISSN 1549-5477

Biancotti, J., Narwani, K., Buehler, N., Mandefro, B., Golan-Lev, T., Yanuka, O., Clark, A., Hill, D., Benvenisty, N. & Lavon, N. (2010). Human embryonic stem cells as models for aneuploid chromosomal syndromes. *Stem Cells,* Vol.28, No.9, (September 2010), pp. 1530-1540, ISSN 1549-4918

Bieberich, E., Silva, J., Wang, G., Krishnamurthy, K. & Condie, B. (2004). Selective apoptosis of pluripotent mouse and human stem cells by novel ceramide analogues prevents teratoma formation and enriches for neural precursors in ES cell-derived neural transplants. *The Journal of Cell Biology,* Vol.167, No.4, (November 2004), pp. 723-734, ISSN 1540-8140

Bourchis, D., Le-Bourhis, D., Patin, D., Niveleau, A., Comizzoli, P., Renard, P. & Viegas-Pequignot, E. (2001). Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. *Current Biology,* Vol.11, No.19, (October 2001), pp. 1542-1546, ISSN 0960-9822

Boyer, L., Lee, T., Cole, M., Johnstone, S., LEvine, S., Zucker, J., Guenther, M., Kumar, R., Murray, H., Jennifer, R., Gifford, D., Melton, D., Jaenisch, R. & Young, R. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell,* Vol.122, No.6, (September 2005), pp. 947-956, ISSN 0092-8674

Carvajal-Vergara, X., Sevilla, A., D'Souza, S.L., Ang, Y.S., Schaniel, C., Lee, D.F., Yang, L., Kaplan, A.D., Adler, E.D., Rozov, R., Ge, Y., Cohen, N., Edelmann, L.J., Chang, B., Waghray, A., Su, J., Pardo, S., Lichtenbelt, K.D., Tartaglia, M., Gelb, B.D. & Lemischka, I.R. (2010). Nature, Vol.465, No.7299, (June 2010), pp. 808-812, ISSN 0028-0836

Caspi, O., Huber, I., Kehat, I., Habib, M., Arbel, J., Gepstein, A., Yankelson, L., Aronson, D., Beyar, R. & Gepstein, L. (2007). Transplantation of Human Embryonic Stem Cell-Derived Cardiomyocytes Improves Myocardial Performance in Infarcted Rat Hearts. *Journal of the American College of Cardiology,* Vol.50, No.19, (October 2007), pp. 1884-1893, ISSN 1558-3597

Ciavatta, D., Ryan, T., Farmer, S. & Townes, T. (1995). Mouse model of human beta zero thalassemia: targeted deletion of the mouse beta maj- and beta min-globin genes in embryonic stem cells. *PNAS,* Vol.92, No.20, (September 1995), pp. 9259-9263, ISSN 0027-8424

Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Goland, R., Wichterle, H., Henderson, C.E. & Eggan, K. (2008). Science, Vol.321, No.5893, (August 2008), pp. 1218-1221, ISSN 1095-9203

Ebert, A.D., Yu, J., Rose, F.F. Jr., Mattis, V.B., Lorson, C.L., Thomson, J.A. & Svendsen, C.N. (2009). Induced pluripotent stem cells from a spinal muscular atrophy patient. Nature, Vol.457, No.7227, (January 2009), pp.277-280, ISSN 0028-0836

Eggan, K., Rode, A., Jentsch, I., Samuel, C., Hennek, T., Tintrup, H., Zevnik, B., Erwin, K., Loring, J., Jackson-Grusby, L., Speicher, M., Kuehn, R. & Jaenisch, R. (2002). Male and female mice derived from the same embryonic stem cell clone by tetraploid embryo complementation. *Nature Biotechnology,* Vol.20, No.5, (May 2002), pp. 455-459, ISSN 1087-0156

Eminli, S., Foudi, A., Stadtfield, M., Maherali, N., Ahfeldt, T., Mostoslavsky, G., Hock, H. & Hochedlinger, K. (2009). Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nature Genetics,* Vol.41, No.9, (September 2009), pp. 968-976, ISSN 1546-1718

Esteban, M., Xu, J., Yang, J., Peng, M., Qin, D., Li, W., Jiang, Z., Chen, J., Deng, K., Zhong, M., Cai, J., Lai, L. & Pei, D. (2009). Generation of induced pluripotent stem cell lines from Tibetan Miniature Pig. *Journal of Biological Chemistry,* Vol.284, No.26, (June 2009), pp. 17634-17640, ISSN 0021-9258

Evans, M. & Kaufman, M. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature,* Vol.292, No.5819, (July 1981), pp. 154-156, ISSN 0028-0836

Gangemi, R., Griffero, F., Marubbi, S., Perera, M., Capra, M., Malatesta, P., Ravetti, G., Zona, G., Daga, A. & Corte, G. (2009). SOX2 silencing in glioblastoma tumor-initiating cells causes stop of proliferation and loss of tumorigenicity. *Stem Cells,* Vol.27, No.1, (January 2009), pp. 40-48, ISSN 1549-4918

Gidekel, S., Pizov, G., Bergman, Y. & Pikarsky, E. (2003). Oct-3/4 is a dose-dependent oncogenic fate determinant. *Cancer Cell,* Vol.4, No.5, (November 2003), pp. 361-370, ISSN 1535-6108

Gordon, J., Scangos, G., Plotkin, D., Barbarosa, J. & Ruddle, F. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *PNAS,* Vol.77, No.12, (December 1980), pp. 7380-7384, ISSN 0027-8424

Gore, A., Li, Z., Fung, H., Young, J., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M., Kiskinis, E., Lee, J., Loh, Y., Manos, P., Montserrat, N., Panopoulos, A., Ruiz, S., Wilbert, M., Yu, J., Kirkness, E., Belmonte, J., Rossi, D., Thomson, J. & Eggan, K., Daley, G., Goldstein, L., Zhang, K. (2011). Somatic coding mutations in human induced pluripotent stem cells. *Nature,* Vol.471, No.7336, (March 2011), pp. 63-72, ISSN 1476-4687

Hakelien, A., Landsverk, H., Robl, J., Skalhegg, B. & Collas, P. (2002). Reprogramming fibroblasts to express T-cell functions using cell extracts. *Nature Biotechnology,* Vol.20, No.5, (May 2002), pp. 460-466, ISSN 1087-0156

Han, J., Sachdev, P. & Sidhu, K. (2010). A combined epigenetic and non-genetic approach for reprogramming human somatic cells. *PLoS ONE,* Vol.5, No.8, (August 2010), pp. 12297, ISSN 1932-6203

Hochedlinger, K., Yamada, Y., Beard, C. & Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell,* Vol.121, No.3, (May 2005), pp. 465-477, ISSN 0092-8674

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A. & Melton, D. (2008a). Induction of Pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature Biotehcnology,* Vol.26, No.7, (July 2008), pp. 795-797, ISSN 1546-1696

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W. & Melton, D. (2008b). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nature Biotehcnology,* Vol.26, No.11, (November 2008), pp. 1269-1275, ISSN 1546-1696

Hussein, S., Batada, N., Vuoristo, S., Ching, R., Autio, R., Narva, E., Ng, S., Sourour, M., Hamalainen, R., Olsson, C., Lundin, K., Mikkola, M., Trokovic, R., Peitz, M., Brustle, O., Bazett-Jones, D., Alitalo, K., Lahesmaa, R., Nagy, A. & Otonkoski, T. (2011). Copy number variation and selection during reprogramming to pluripotency. *Nature,* Vol.471, No.7336, (March 2011), pp., ISSN 1476-4687

Kang, L., Wang, J., Zhang, Y., Kou, Z. & Gao, S. (2009). iPS Cells Can Support Full-Term Development of Tetraploid Blastocyst-Complemented Embryos. *Cell Stem Cell,* Vol.5, No.2, (August 2009), pp. 135-138, ISSN 1875-9777

Kim, D., Gu, Y., Ishii, M., Fujimiya, M., Qi, M., Nakamura, N., Yoshikawa, T., Sumi, S. & Inoue, K. (2003). In vivo functioning and transplantable mature pancreatic islet-like cell clusters differentiated from embryonic stem cell. *Pancreas,* Vol.27, No.2, (August 2003), pp. 34-41, ISSN 1536-4828

Kim, D., Kim, C.H., Moon, J.I., Chung, Y.G., Chang, M.Y., Han, B.S., Ko, S., Yang, E., Cha, K.Y., Lanza, R. & Kim, K.S. (2009). Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell,* Vol.4, No.6, (June 2009), pp. 472-476, ISSN 1875-9777

Kim, J. (2009). Oct4-induced pluripotency in adult neural stem cells. *Cell,* Vol.136, No.3, (February 2009), pp. 411-419, ISSN 1097-4172

Kim, J., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Arauzo-Bravo, M., Ruau, D., Han, D., Zenke, M. & Scholer, H. (2008). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature,* Vol.454, No.7204, (July 2008), pp. 646-650, ISSN 1476-4687

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M., Ji, H., Ehrlich, L., Uabuuchi, A., Takeuchi, A., Cunniff, K., Hongguang, H., McKinney-Freeman, S., Naveiras, O., Yoon, T., Irizarry, R., Jung, N., Seita, J., Hanna, J., Murakami, P., Jaenisch, R., Weissleder, R., Orkin, S., Weissman, I., Feinberg, A. & Daley, G. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature,* Vol.467, No.7313, (September 2010), pp. 285-290, ISSN 1476-4687

Kleinsmith, L. & Pierce, G.J. (1964). Multipotentiality of Single Embryonal Carcinoma Cells. *Cancer Res.,* Vol.24, No.1, (October 1964), pp. 1544-1551, ISSN 0008-5472

Kopp, J., Ormsbee, B., Desler, M. & Rizzino, A. (2008). Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells,* Vol.26, No.4, (April 2008), pp. 903-911, ISSN 1549-4918

Krizhanovsky, V. & Lowe, S. (2009). Stem cells: The promises and perils of p53. *Nature,* Vol.460, No.7259, (August 2009), pp. 1085-1086, ISN 1476-4687

Lan, J., Hua, S., Zhang, H., Song, Y., Liu, J. & Zhang, Y. (2010). Methylation patterns in 5' terminal regions of pluripotency-related genes in bovine in vitro fertilized and cloned embryos. *Journal of Genetics and Genomics,* Vol.37, No.5, (May 2010), pp. 297-304, ISSN 1673-8527

Lanza, R.P., Cibelli, J.B., Blackwell, C., Cristofalo, V.J., Francis, M.K., Baerlocher, G.M., Mak, J., Schertzer, M., Chavez, E.A., Sawyer, N., Lansdrop, P.M. & West, M.D. (2000). Extension of cell life-span and telomere length in animals cloned from senescent somatic cells. *Science*, Vol.288, No.5466, (April 2000), pp.665-669, ISSN 1095-9203.

Lee, G., Papapetrou, E.P., Kim, H., Chambers, S.M., Tomishima, M.J., Fasano, C.A., Ganat, Y.M., Menon, J., Shimizu, F., Viale, A., Tabar, V., Sadelain, M. & Studer, L. (2009). Modelling parthenogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature*, Vol.461, No.7262, (September 2009), pp. 402-406, ISSN 0028-0836

Lengner, C., Camargo, F., Hochedlinger, K., Welstead, G., Zaidi, S., Gokhale, S., Scholer, H., Tomilin, A. & Jaenisch, R. (2007). Oct4 Expression Is Not Required for Mouse Somatic Stem Cell Self-Renewal. *Cell Stem Cell,* Vol.1, No.4, (October 2007), pp. 403-415, ISSN 1875-9777

Li, W., Wei, W., Zhu, S., Zhu, J., Shi, Y., Lin, T., Hao, E., Hayek, A., Deng, H. & Ding, S. (2009). Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors. *Cell Stem Cell,* Vol.4, No.1, (January 2009), pp. 16-19, ISSN 1875-9777

Liao, J., Wu, Z., Wang, Y., Cheng, L., Cui, C., Gao, Y., Chen, T., Rao, L., Chen, S., Jia, N., Dai, H., Xin, S., Kang, J., Pei, G. & Xiao, L. (2008). Enhanced efficiency of generating induced pluripotent stem (iPS) cells from human somatic cells by a combination of six transcription factors. *Cell Research.,* Vol.18, No.5, (May 2008), pp. 600-603, ISSN 1748-7838

Lin, S., Chang, D., Chang-Lin, S., Lin, C., Wu, D., Chen, D. & Ying, S. (2008). Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA,* Vol.14, No.10, (October 2008), pp. 2115-2124, ISSN 1469-9001

Lister, R., Pelizzola, M., Kida, Y., Hawkins, R., Nery, J., Hon, G., Antosiewcz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., Downes, M., Yu, R., Stewart, R., Ren, B., Thomson, J., Evans, R. & Ecker, J. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature,* Vol.471, No.7336, (March 2011), pp. 68-76, ISSN 1476-4687

Marion, R., Strati, K., Li, H., Tejera, A., Schoeftner, S., Ortega, S., Serrano, M. & Blasco, M. (2008). Telomeres Acquire Embryonic Stem Cell Characteristics in Induced Pluripotent Stem Cells. *Cell Stem Cell,* Vol.6, No.4, (February 2009), pp. 141-154, ISSN 1875-9777

Mintz, B. & Illmensee, K. (1975). Normal genetically mosaic mice produced from malignant teratocarcinoma cells. *PNAS,* Vol.72, No.9, (September 1975), pp. 3585-3589, ISSN 0027-8424

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. & Yamanaka, S. (2003). The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells. *Cell,* Vol.113, No.5, (May 2003), pp. 631-642, ISSN 0092-8674

Moretti, A., Bellin, M., Welling, A., Jung, C.B., Lam, J.T., Bott-Flügel, L., Dorn, T., Goedel, A., Höhnke, C., Hofmann, F., Seyfarth, M., Sinnecker, D., Schömig, A. & Laugwitz, K.L. (2010). *New England Journal of Medicine*. Vol.363, No.15, (October 2010), pp. 1397-1409, ISSN 0028-4793

Moustafa, L. & Brinster, R. (1972). Induced chimaerism by transplanting embryonic stem cells into mouse blastocysts. *Journal of Experimental Zoology,* Vol.181, No.2, (August 1972), pp. 193-201, ISSN 0022-104X

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N. & Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature Biotechnology,* Vol.26, No.1, (January 2008), pp. 101-106, ISSN 1546-1696

Newman, A. & Cooper, J. (2010). Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell,* Vol.7, No.2, (August 2010), pp. 258-262, ISSN 1875-9777

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H. & Smith, A. (1998). Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. *Cell,* Vol.95, No.3, (October 1998), pp. 379-391, ISSN 0092-8674

Niwa, H., Miyazaki, J. & Smith, A. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genetics,* Vol.24, No.4, (April 2000), pp. 372-376, ISSN 1061-4036

Okita, K., Ichisaka, T. & Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature,* Vol.448, No.7151, (July 2007), pp. 313-317, ISSN 1476-4687

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T. & Yamanaka, S. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. *Science,* Vol.322, No.5903, (November 2008), pp. 949-953, ISSN 1095-9203

Pandya, A., Talley, L., Frost, A., Fitzgerald, T., Trivedi, V., Chakravarthy, M., Chieng, D., Grizzle, W., Engler, J., Krontiras, H., Bland, K., LoBuglio, A., Lobo-Ruppert, S. & Ruppert, J. (2004). Nuclear Localization of KLF4 is Associated with an Aggressive Phenotype in Early-Stage Breast Cancer. *Clinical Cancer Research,* Vol.10, No.8, (April 2004), pp. 2709-2719, ISSN 1078-0432

Park, I., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M., Cowan, C., Hochedlinger, K. & Daley, G. (2008a). Disease-Specific Induced Pluripotent Stem Cells. *Cell,* Vol.134, No.5, (September 2008), pp. 877-886, ISSN 1097-4172

Park, I., Zhao, R., West, J., Yabuuchi, A., Huo, H., Ince, T., Lerou, P., Lensch, M. & Daley, G. (2008b). Reprogramming of human somatic cells to pluripotency with defined factors. *Nature,* Vol.451, No.7175, (January 2008), pp. 141-146, ISSN 1476-4687

Pechloff, K., Holch, J., Ferch, U., Schweneker, M., Brunner, K., Kremer, M. & Sparwasser, T. (2010). The fusion kinase ITK-SYK mimics a T cell receptor signal and drives oncogenesis in conditional mouse models of peripheral T cell lymphoma. *Journal of Experimental Medicine,* Vol.207, No.5, (May 2010), pp. 1031-1044, ISSN 1540-9538

Pruszak, J., Sonntag, K., Aung, M., Sanchez-Pernaute, R. & Isacson, O. (2007). Markers and methods for cell sorting of human embryonic stem cell-derived neural cell populations. *Stem Cells,* Vol.25, No.9, (September 2007), pp. 2257-2268, ISSN 1549-4918

Pryzhkova, M., Peters, A. & Zambidis, A. (2010). Erythropoietic differentiation of a human embryonic stem cell line harbouring the sickle cell anaemia mutation. *Reproductive BioMedicine Online,* Vol.21, No.2, (August 2010), pp. 196-205, ISSN 1472-6491

Pulvers, J., Bryk, J., Fish, J., Brauninger, M., Arai, Y., Schreier, D., Naumann, R., Helppi, J., Habermann, B., Vogt, J., Nitsch, R., Toth, A., Enard, W., Paabo, S. & Huttner, W. (2010). Mutations in mouse Aspm (abnormal spindle-like microcephaly associated) cause not only microcephaly but also major defects in the germline. *PNAS,* Vol.107, No.38, (September 2010), pp. 16595-16600, ISSN 1091-6490

Raya, A., Rodríguez-Pizà, I., Guenechea, G., Vassena, R., Navarro, S., Barrero, M.J., Consiglio, A., Castellà, M., Río, P., Sleep, E., González, F., Tiscornia, G., Garreta, E., Aasen, T., Veiga, A., Verma, I.M., Surrallés, J., Bueren, J. & Izpisúa Belmonte, J.C. (2009). Nature, Vol.460, No.7251, (July 2009), pp.53-59, ISSN 0028-0836

Reya, T., Morrison, S., Clarke, M. & Weissman, I. (2001). Stem cells, cancer, and cancer stem cells. *Nature,* Vol.141, No.6859, (November 2001), pp. 105-111, ISSN 0028-0836

Richt, J., Kasinathon, P., Hamir, A., Castilla, J., Sathyyaseelan, T., Vargas, F., Sathiyaseelan, J., Wu, H., Matsushita, H., Koster, J., Kato, S., Ishida, I., Soto, A., Robl, J. & Kuroiwa, Y. (2006). Production of cattle lacking prion protein. *Nature Biotechnology,* Vol.25, No.1, (January 2007), pp. 132-138, ISSN 1087-0156

Schuldiner, M., Itskovitz-Eldor, J. & Benvenisty, N. (2003). Selective ablation of human embryonic stem cells expressing a "suicide" gene. *Stem Cells,* Vol.21, No.3, (May 2003), pp. 257-265, ISSN 1066-5099

Shi, Y., Desponts, C., Do, J., Hahm, H., Scholer, H. & Ding, S. (2008). Induction of Pluripotent Stem Cells from Mouse Embryonic Fibroblasts by Oct4 and Klf4 with Small-Molecule Compounds. *Cell Stem Cell,* Vol.3, No.5, (November 2008), pp. 568-574, ISSN 1875-9777

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G., Cook, E., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., Isacson, O. & Jaenisch, R. (2009). Parkinson's disease patient-derived pluripotent stem cells free of viral reprogramming factors. *Cell,* Vol.136, No.5, (March 2009), pp. 964-977, ISSN 1097-4172

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. & Hochedlinger, K. (2008). Induced pluripotent stem cells generated without viral integration. *Science,* Vol.322, No.5903, (November 2008), pp. 945-949, ISSN 1095-9203

Sumer, H., Jones, K., Liu, J., Heffernan, C., Tat, P., Upton, K. & Verma, P. (2009). Reprogramming of somatic cells after fusion with iPS and ntES cells. *Stem Cells & Development,* Vol.19, No.2, (February 2010), pp. 239-246, ISSN 1557-8534

Takahashi, K. & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* Vol.126, No.4, (August 2006), pp. 663-676, ISSN 0092-8674

Taranger, C., Noer, A., Sorensen, A., Hakeliel, A., Boquest, A. & Collas, P. (2005). Induction of dedifferentiation, genomewide transcriptional programming, and epigenetic reprogramming by extract of carcinoma and embryonic stem cells. *Molecular Biology of the Cell,* Vol.16, No.12, (December 2005), pp. 5719-5735, ISSN 1059-1524

Telugu, B., Ezashi, T. & Roberts, R. (2010). The Promise of stem cell research in pigs and other ungulate species. *Stem Cell Reviews,* Vol.6, No.1, (March 2010), pp. 31-41, ISSN 1558-6804

Thomson, J., Itskovitz-Eldor, J., Shapiro, S., Waknitz, M., Swiergiel, K., Marshall, V. & Jones, J. (1998). Embryonic stem cell lines derived from human blastocysts. *Science,* Vol.282, No.5391, (November 1998), pp. 1145-1147, ISSN 0036-8075

Thuan, N. V., Kishigami, S. & Wakayama, T. (2010). How to improve the success rate of mouse cloning technology. *Journal of Reproduction and Development,* Vol.56, No.1, (January 2010), pp 20-30, ISSN 0916-8818.

Tomioka, I., Maeda, T., Shimada, H., Kawai, K., Okada, Y., Igarashi, H., Oiwa, R., Iwasaki, T., Aoki, M., Kimura, T., Shiozawa, S., Shinohara, H., Suemizu, H., Sasaki, E. & Okano, H. (2010). Generating induced pluripotent stem cells from common marmoset (Callithrix jacchus) fetal liver cells using defined factors, including Lin28. *Genes to Cells,* Vol.15, No.9, (September 2010), pp. 959-969, ISSN 1365-2443

Tong, C., Li, P., Wu, N., Yan, Y. & Ying, Q. (2010). Production of p53 gene knockout rats by homologous recombination in embryonic stem cells. *Nature,* Vol.467, No.7312, (September 2010), pp. 211-213, ISSN 1476-4687

Warren, L., Manos, P., Ahfeldt, T., Loh, Y., Li, H., Lau, F., Ebina, W., Mandal, P., Smith, Z., Meissner, A., Daley, G., Brack, A., Collins, J., Cowan, C., Schlaeger, T. & Rossi, D. (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell,* Vol.7, No.5, (November 2010), pp. 618-630, ISSN 1875-9777

Wilmut, I., Schnieke, A., McWhir, J., Kind, A. & Campbell, K. (1997). Viable offspring derived from fetal and adult mammalian cells. *Nature,* Vol.385, No.6619, (February 1997), pp. 810-813, ISSN 0028-0836

Woltjen, K., Michael, I., Mohseni, P., Desai, R., Mileikovsky, M., Cowling, R.H.R., Wang, W., Liu, P., Gertsenstein, M., Kaji, K., Sung, H. & Nagy, A. (2009). PiggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature,* Vol.458, No.7239, (April 2009), pp. 766-770, ISSN 1476-4687

Wu, L., Sun, C., Ryan, T., Pawlik, K., Ren, J. & Townes, T. (2006). Correction of sickle cell disease by homologous recombination in embryonic stem cells. *Blood,* Vol.108, No.4, (August 2006), pp. 1183-1188, ISSN 0006-4971

Wu, Z., Chen, J., Ren, J., Bao, L., Liao, J., Cui, C., Rao, L., Li, H., Gu, Y., Dai, H., Zhu, H., Teng, X., Cheng, L. & Xiao, L. (2009). Generation of pig induced pluripotent stem cells with a drug-inducible system. *Journal of Molecular and Cellular Biology,* Vol.1, No.1, (October 2009), pp. 46-54, ISSN 1759-4685

Yang, D., Zhang, Z., Oldenburg, M., Ayala, M. & Zhang, S. (2008). Human Embryonic Stem Cell-Derived Dopaminergic Neurons Reverse Functional Deficit in Parkinsonian Rats. *Stem Cells,* Vol.26, No.1, (January 2008), pp. 55-63, ISSN 1549-4918

Yu, J., Vodyanik, M., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J., Tian, S., Nie, J., Jonsdottir, G., Ruotti, V., Stewart, R., Slukvin, I. & Thomson, J. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science,* Vol.318, No.5858, (December 2007), pp. 1917-1920, ISSN 1095-9203

Zhong, X., Li, N., Liang, S., Huang, Q., Coukos, G. & Zhang, L. (2010). Identification of mircoRNAs regulating reprogramming factor Lin28 in embryonic stem cells and cancer cells. *Journal of Biological Chemistry,* Vol.285, No.53, (December 2010), pp. 41961-41971, ISSN 1083-351X

Zhu, X., Pan, X., Wang, W., Chen, Q., Pang, R., Cai, X., Hoffman, A. & Hu, J. (2010). Transient in vitro epigenetic reprogramming of skin fibroblasts into multipotent cells. *Biomaterials,* Vol.31, No.10, (April 2010), pp. 2779-2787, ISSN 1878-5905

29

Genetic Modification of Domestic Animals for Agriculture and Biomedical Applications

Cai-Xia Yang and Jason W. Ross
Department of Animal Science, Iowa State University, USA

1. Introduction

Transgenic technology has been applied mainly in the study of gene structure and function in model organisms and gene therapy for human diseases. Transgenic technology has potential for rapidly improving quantity and quality of agricultural products, compared to traditional selection and breeding methods in domestic animals that are time consuming when attempting to alter the desired allele frequency for specific traits. Additionally, transgenic animals can be used as biomedical research models or directly for human health, by producing recombinant pharmaceutical proteins and/or organs for xenotransplantation. Due to the advantage of bypassing the need of embryonic stem (ES) cells that are difficult to isolate in domestic animal species, cell-based method of transgenesis followed by somatic cell nuclear transfer (SCNT) is currently widely applied. However, due to the limitations in making genetic modifications and SCNT, producing genetically modified animals is still inefficient. Fortunately, the current advancement of new techniques and methods in both gene targeting (Urnov et al., 2010) and abilities to produce pluripotent stem cells (Voigt and Serikawa, 2009) holds great promises for this field.

In this chapter, we will review the recent progress and technical route of the cell-based method of transgenesis by SCNT and discuss the newly emerging methods to enrich the gene targeting frequency of somatic cells. We will also discuss factors to improve the efficiency of SCNT and our future perspectives on the promises of this field.

2. Recent progress and applications of transgenic domestic animals

2.1 Methods of creating genetically modified animals

2.1.1 Pronuclear microinjection /viral-mediated /sperm-mediated /ICSI (Intractyoplasmic sperm injection)- mediated gene transfer

Numerous methods have been successfully used to introduce genetic modifications and produce transgenic animals, including pronuclear microinjection of foreign DNA into zygotes (Hammer et al., 1985; Pursel and Rexroad, 1993), viral-mediated gene transfer (Chan et al., 1998; Cabot et al., 2001; Whitelaw et al., 2008), sperm-mediated gene transfer (Castro et al., 1991; Chang et al., 2002; Lavitrano et al., 2002 and 2006) and intracytoplasmic injection (ICSI) of a sperm head carrying foreign DNA (Perry et al., 1999; Osada et al., 2005; Moisvadi et al., 2009; García-Vázquez et al., 2010). Despite the proven successful application of these techniques, some problems, such as inefficiency and mosaicism (transgene not going into the germline) (Table 1) remain to be solved and limit the practical application of these methods.

Method	**Advantages**	**Disadvantages**	**Reference**
Pronuclear microinjection	The first method successfully used for different animal species.	Low embryo survival, low and random integration, multiple copies, high cost in domestic animals.	Hammer et al., 1985; Pursel and Rexroad, 1993
Viral-mediated DNA transfer	Infect both dividing and non-dividing cells, less damage by co-culture with zona-free zygotes or injection into the perivitelline space compared with pronuclear microinjection, high integration.	Limited DNA capacity, random integration.	Chan et al, 1998; Cabot et al., 2001; Whitelaw et al., 2008
Sperm-mediated DNA transfer	Relatively high efficiency as compared to pronuclear injection, low cost, ease of use.	No control of integration site.	Castro et al., 1991; Chang et al., 2002; Lavitrano et al., 2002; Lavitrano et al., 2006
Intracytoplasmic sperm injection -mediated DNA transfer	Allow introduction of very large DNA transgenes, relatively high efficiency as compared to pronuclear injection.	No control of integration site.	Perry et al., 1999; Osada et al., 2005; Moisvadi et al., 2009; García-Vázquez et al., 2010

Table 1. The Advantages and disadvantages of different transgenic methods.

Gene targeting by homologous recombination often offers more precise and site-specific integration, sometimes at single nucleotide level. This is particularly important since single nucleotide changes can be a common culprit for some of the human diseases, which require more precise manipulation to build biomedical models using transgenic animals.

2.1.2 Cell-based transgenesis via SCNT

The first live animal by SCNT was produced in 1997, "Dolly" (Wilmut et al., 1997), demonstrating the ability of a differentiated somatic cell to produce live offspring following nuclear remodeling and reprogramming by an oocyte. Then, the birth of "Polly" in the same year (Schnieke et al., 1997), the first transgenic sheep produced by transfer of nuclei from transfected fetal fibroblasts, demonstrated a route to create transgenic cloned animals. This cell-based method of transgenesis by SCNT can bypass the absence of ES cells, offers the reliability of germline transmission by avoiding mosaic transgene integration, and provides the only currently used strategy to knock out a gene in domestic animals (reviewed by Ross et al., 2009a). Currently, SCNT using transgenic cells cultured *in vitro* as a source of donor nuclei is becoming the most utilized technique to produce the transgenic domestic animals. However, while the advantages and success of this strategy are well documented (Table 2), the procedure is still labor-intensive and inefficient. Recently, some new techniques have been reported that may have the potential to improve the production efficiency of cloned transgenic domestic animals by increasing the efficiency of gene targeting or nuclear remodeling and reprogramming following SCNT.

2.2 Utilization of transgenic models

The potential for transgenic domestic animals to benefit humans is not only in agricultural production by providing more and better agricultural products for human consumption but also in biomedicine, such as for producing recombinant pharmaceutical proteins, making organs suitable for xenotransplantation and establishing human disease models. An overview list of the transgenic domestic animals produced via SCNT is given in Table 2 to demonstrate their applications.

2.2.1 Improved animal agriculture production

Increased utilization of domestic animals and their products requires breeding and selection strategies for specific traits. However, classical breeding and genetic selection have some disadvantages, such as the inability to control gene frequency of desired genotypes coupled with long generation intervals. The application of transgenic technology offers a powerful tool to rapidly improve agriculture production by developing domestic animals that express desired traits via genome manipulation strategies. Previous studies have demonstrated the practical application of transgenesis to improve numerous agricultural traits of domestic animals, including increased growth rate (Pursel et al., 1999), increased meat quality (Saeki et al., 2004), enhanced disease resistance (Lo et al., 1991; Clements et al., 1994), and better milk production and composition (Wheeler et al., 2001; Reh et al., 2004). Nevertheless, utilization of the cell-based method of transgenesis via SCNT has also been successfully used to alter characteristics of pork quality (Lai et al., 2006), enhance disease resistance (Denning et al., 2001a; Wall et al., 2005; Richt et al., 2007) and improve milk composition (Brophy et al., 2003).

Species	Application	Donor cells	Key Molecule	Construct	DNA delivery method	Reference
Sheep	Bioreactor	serum starved fetal fibroblasts	human clotting factor IX (FIX)	pMIX1	Lipofection	Schnieke et al., 1997
	Disease model	serum starved fetal fibroblasts	alpha1(I) procollagen (COL1A1)	COLT-1/COLT-2	Lipofection	McCreath et al., 2000
	Xenotransplantation; disease resistance	serum starved fetal fibroblasts	alpha(1,3)galactosyl transferase (GGTA1), prion protein (PrP)	GGTA1; PrP	Electroporation	Denning et al., 2001a
Cattle	Marker gene	actively dividing fetal fibroblasts	β-galactosidase–neomycin, cytomegalovirus (CMV)	pCMV/ βGEO	Not described	Cibelli et al., 1998
	Bioreactor	serum starved fetal/ear fibroblasts	bovine prochymosin coding gene, αS1-casein promoter	TFF1/TFF2	Lipofection	Zakhartchenko et al., 2001
	GFP-reporter	dividing fetal fibroblast	enhanced, humanized version of the GFP-reporter	CEEGFP	Lipofection	Bordignon et al., 2003
	Bioreactor	serum starved fetal fibroblasts and G1 fibroblasts	β-casein and Κ-casein	$CSN2^{A3}$ / $CSN2/3^{3}$	Lipofection	Brophy et al., 2003
	Bioreactor	fetal fibroblasts	bispecific single-chain variable fragment (biscFV) molecule with anti-human CD28 × anti-human melanoma specificity	Bi-scFV r28M	Lipofection	Grosse-Hovest et al., 2004
	Bioreactor	fetal fibroblasts	growth hormone (GH)	hGH vector	Not described	Salamone et al., 2006
	Bioreactor	fetal oviduct epithelial cells	human α lactalbumin	phLa4 EGFP NEO	Electroporation	Wang et al., 2008
	Bioreactor	fetal fibroblasts	human lactoferrin	hLF BAC	Microinjection/ Electroporation/ Lipofection	Yang et al., 2008
Goat	eGFP	fetal fibroblasts	green fluorescent protein reporter gene	CLeGFP	Lipofection	Keefer et al., 2001
Pig	Xenotransplantation	fetal fibroblasts	alpha -1,3-galactosyltransferase locus	pGalGT	Electroporation	Lai et al., 2002a
	Xenotransplantation	fetal fibroblasts	alpha -1,3-galactosyltransferase locus	pPL654 and pPL657	Electroporation	Dai et al., 2002
	Disease model	fetal fibroblasts	cystic fibrosis (CF); CF transmembrane conductance receptor (CFTR)	CFTR-*null and* CFTR-ΔF508	Adeno-associated virus (AAV)	Rogers et al., 2008
	Disease model	ear fibroblasts	Alzheimer's disease; neuronal variant of the human amyloid precursor protein gene with the Swedish mutation	pPDGFbEGFP	Lipofection	Kragh et al., 2009
	Xenotransplantation	fetal fibroblasts	human A20 gene	pCAGGSEhA20-IRESNEO	Electroporation	Oropeza et al., 2009
	Disease model	fetal fibroblasts	Huntington's disease	pCAG-HTT-2A-ECFP	Electroporation	Yang et al., 2010
Rabbit	eGFP	serum starved adult fibroblasts	green fluorescent protein reporter gene	pEGFP-C1	Lipofection	Li et al., 2009

Table 2. Overview on successful transgenic domestic animals via SCNT.

2.2.2 Xenotransplantation

Due to the potential application of human organ transplantation and the growing gap between the demand and availability of organs for transplantation, the pig has long been considered as an alternative source to provide organs for humans. In contrast to organs from other animal species, domestic pig organs share many similarities to those of humans including size, anatomy and physiology. However, despite these similarities, significant immunological barriers exist impeding the success of pig to human xenotransplantation (Cooper et al., 2008), in addition to concerns regarding the transmission of pig specific viruses to the human genome (Magre et al., 2003). The immunological obstacles of xenotransplantation include rapid hyperacute rejection (HAR), delayed acute vascular rejection (AVR), and the cellular immune response that occurs within weeks (Auchincloss and Sachs, 1998). Two transgenic strategies have been successfully applied to overcome HAR. One is to express human proteins that inhibit the complement cascade in transgenic pigs (Fodor et al., 1994; Cozzi and White, 1995; Diamond et al., 2001). For example, transgenic expression of human complement inhibitor CD59, CD46 and DAF (decay-accelerating factor, also named as CD55) in pigs prolongs survival rates from minutes to days and months following heart and kidney transplantation into baboons or monkeys by blocking the damage from HAR (Diamond et al., 1996 and 2001; Byrne et al., 1997; Zaidi et al., 1998; Bhatti et al., 1998; Chen et al., 1999). The second strategy to avoid HAR is to knockout the genes that induce the production of antigenic structures (α-gal-epitopes) on the surface of pig organs (Lai et al., 2002a; Dai et al., 2002; Phelps et al., 2003; Yamada et al., 2005). α-gal-epitopes on endothelial cells of porcine transplanted organs can be recognized by human xenoreactive natural antibodies (XNA) and activate the HAR cascade (Galili, 1993). Genetically engineering of pigs to lower or inhibit the expression level of XNA targets is thought to be a promise way to eliminate the HAR. Following the successful production of α-1,3-galactosyltransferase knockout pigs (Lai et al., 2002a; Dai et al., 2002; Phelps et al., 2003; Yamada et al., 2005), van Poll et al. (2010) recently showed that exposure of isolated xenogeneic pig liver sinusoidal endothelial cells (LSECs) from α-1,3-galactosyltransferase-deficient pigs to human and baboon serum reduces IgM binding and complement activation levels as compared to wild-type pig LSECs. However, Diswall et al. (2010) found a different reactivity pattern of baboon and human serum to pig glycolipid antigens isolated from α-1,3-galactosyltransferase knockout and wild-type pig hearts and kidneys, suggesting that non-human primates may not be an ideal model for modeling pig to human xenotransplantation. If HAR is controlled, the next obstacle to xenotransplantation is AVR which is due to the loss of porcine thrombomodulin in xenograft rejection or the inability of porcine thrombomodulin to activate human protein C. One of genetic engineering strategies to overcome AVR is to express human thrombomodulin in pigs. Petersen et al. (2006) showed the production of transgenic cloned pigs using CD59/DAF and human thrombomodulin triple transgenic adult donor cells.

With regard to risks associated with xenotransplantation, previous studies have shown that the risk of cross-transmission of pig endogenous retrovirus (PERV) to human patients or nonhuman primate recipients is low (Paradis et al., 1999; Switzer et al., 2001), although it has been found PERV can infect human cells in culture (Patience et al., 1997). Never-the-less, some investigators have been working to further reduce the possibility by creating pigs with suppressed expression of endogenous retroviruses (Ramsoondar et al., 2009). In addition to PERV, herpesvirus is another concern regarding biosafety in xenotransplantation (Mueller et al., 2011). Overall, dramatically increased knowledge will facilitate the clinical application of transgenic strategies of pig to human xenotransplantation.

2.2.3 Production of recombinant proteins

The mammary gland and blood of transgenic domestic animals, including sheep, goats, cows, pigs and rabbits, have been successfully used as bioreactor to produce numerous recombinant proteins, such as antibodies (Grosse-Hovest et al., 2004), growth factors (Schnieke et al., 1997) and pharmaceuticals (reviewed by Melo et al., 2007). Using various mammary gland-specific or blood-specific promoters to drive the expression of specific protein-coding genes, transgenic domestic animals can continuously produce the recombinant proteins in large quantities in their milk or blood. Recombinant proteins, including human von Willebrand factor (Lee et al., 2009), human erythropoietin (Park et al., 2006), human insulin-like growth factor-I (Monaco et al., 2005), human factor VIII (Paleyanda et al., 1997) and bovine alpha-lactalbumin (Bleck et al., 1998) have been produced in the milk of transgenic pigs. Transgenic goats, capable of synthesizing human butyrylcholinesterase (Huang et al., 2007) and human longer acting tissue plasminogen activator (Ebert et al., 1991) in their milk have also been created. Human salmon calcitonin in milk of transgenic rabbits (McKee et al., 1998); human factor IX (Schnieke et al., 1997) and alpha-1-antitrypsin (Wright et al., 1991) in milk of transgenic sheep; and human lactoferrin (van Berkel et al., 2002; Yang et al., 2008), human growth hormone (Salamone et al., 2006) and human α-lactalbumin (Wang et al., 2008) in milk of transgenic cows are all additional examples of using transgenic domestic animals and the mammary gland as a bioreactor for production of recombinant proteins. Table 2 summarizes some recombinant proteins expressed in milk or blood of cloned transgenic domestic animals. One of the major advantages of using domestic animals for this purpose is that the produced protein is thought to undergo more accurate posttranslational processing to ensure their biological activity. While this application of transgenic technology to produce recombinant protein products is rapidly developing, research efforts exploring the efficacy of these products are still needed.

2.2.4 Biomedical models of human diseases

Another important application of genetically modified domestic animals is to create better and novel biomedical models of human diseases. Pig models of different human diseases, including retinitis pigmentosa, cardiovascular disease, diabetes, Huntington's disease, cystic fibrosis and Alzheimer's disease have been well discussed by Prather et al. (2008). Many of these biomedical models created by SCNT are listed in Table 2.

3. Technical aspects of cell-based transgenesis by SCNT

The general procedure of cell-based transgenesis via SCNT is to construct a DNA vector, deliver the vector into cultured somatic cells, select transgenic cell lines, utilize SCNT and transfer cloned embryo into surrogates (Figure 1).

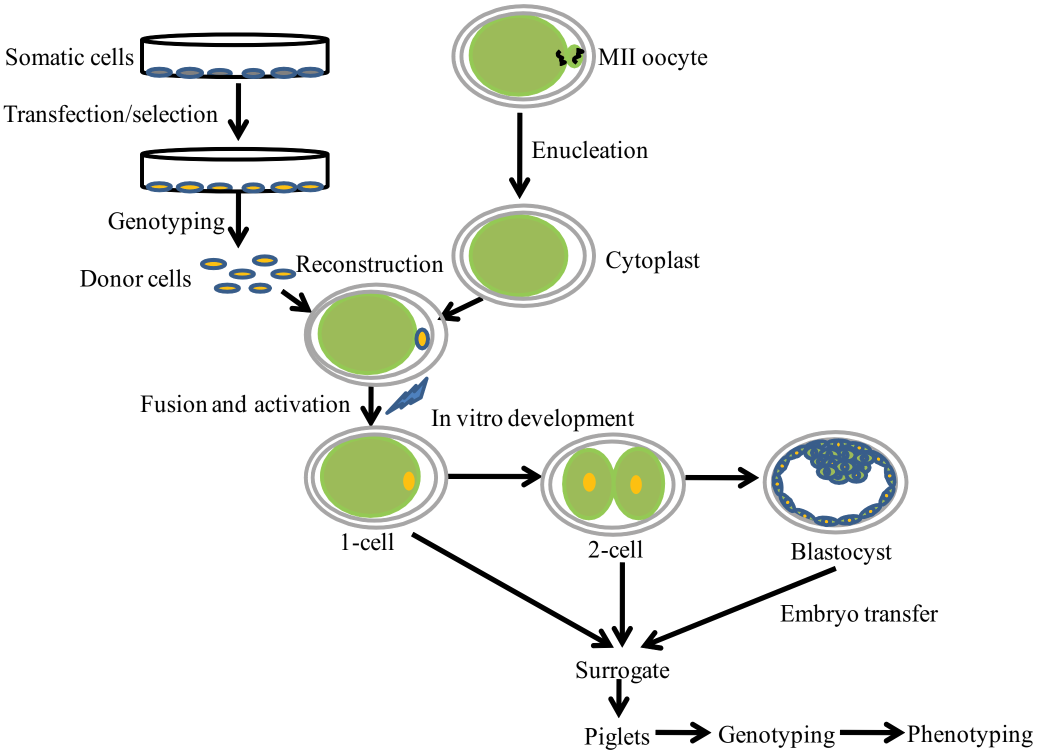

Somatic cells
Transfection/selection
Genotyping
Donor cells
Reconstruction
MII oocyte
Enucleation
Cytoplast
Fusion and activation
In vitro development
1-cell
2-cell
Blastocyst
Embryo transfer
Surrogate
Piglets
Genotyping
Phenotyping

Fig. 1. Technical diagram of cell-based transgenesis followed by SCNT in pigs.

3.1 Vector construction

Currently, the whole genome sequence of several domestic animals, including cattle, poultry, cats, dogs, horses, pigs and rabbits is available via the ensembl database (www.ensembl.org), providing researchers useful sequence information for a large number of genes useful in designing DNA constructs for transgenic genome modification. Precise design and construction of DNA constructs is critical efficiently creating transgenic domestic animals. Transgenesis involves adding a gene to a host genome (transgenic), physically deleting a specific region of the host genome making a non-functional gene (knock-out), replacing an active gene by another active gene (knock-in), or introducing a point mutation (point mutation knock-in). Depending on the objective of the transgenic modification, different strategies of vector design need to be carefully considered to ensure success.

3.1.1 Transgenic vs. Knock-out vs. Knock-in

The design strategy for the transgenic vector, which is based on random integration, essentially includes a gene ORF (open reading frame), a promoter element and the appropriate RNA processing component(s). The promoter is a major transcriptional regulatory element that normally includes regulatory elements and the transcriptional start site typically located in 5′ sequence of the gene. Utilization of specific promoter with a transgene enables tissue or cell specific expression and can significantly impact the expression efficiency of the transgene. Carefully choosing a well-characterized promoter will enable the precise control of transgene expression. The gene ORF is usually derived from the cDNA for the protein of interest, which includes translational start (ATG) and stop codons. Examination of extra sequence, if any, existing between the transcriptional start site and the translational start codon should be performed to ensure the absent of potential regulatory elements. Furthermore, consideration should be given to remove non-coding sequence of genes to avoid the introduction of the regulatory elements although this may also have consequences impacting mRNA stability. Reliable transgene expression is not only regulated by the sequence in the expression vector, but also by intrinsic factors in the host genome following transgene integration. Several additional points should be considered in transgenic vector construction. (i) How GC rich regulatory sequence are, especially CpG islands in the promoters can have significant implications for the expression of the transgene as methylation of CpG islands can inactivate the promoter and silence the transgene expression. (ii) Inclusion of specific elements that favor mRNA maturation and transfer to the cytoplasm, as it has been demonstrated that the inclusion of an intron in the transgenic vector can increase the transgene expression level (Choi et al., 1991; Duncker et al., 1997). These genomic regions may have critical sequence motifs affecting mRNA splicing and accumulation. (iii) Removal of unnecessary plasmid DNA sequence used in recombinant DNA cloning. In mice, it was found that plasmid sequence existing in the transgenic vector can decrease transgene expression (Kjer-Nielsen et al., 1992). Furthermore, the local chromatin status of the transgenic locus can also affect expression meaning that the integration of the same transgene in different genomic locations can have profound effects on the expression level of the transgene.

The primary strategy for targeted genome modifications, including knock-out and knock-in applications, is to use homologous recombination to introduce precise, site-specific genome alterations. In knock-out targeting vector systems, the primary approach is to delete DNA fragments (entire gene or partial deletion) important for the gene function by homologous recombination. Compared with the knock-out targeting vector, in addition to the targeting arms and positive selection cassette, the knock-in targeting vector includes the extra replacement cassette that will replace the target gene with a new gene (a set of genes or a point mutation). Targeting vectors relying on homologous recombination contain 5′ and 3′ homologous arms flanking a positive selection cassette (Rogers et al., 2008; Sun et al., 2008). Several principles should be considered when designing a successful targeting vector. (i) Avoid excessive repetitive DNA. In mouse ES cells it has been demonstrated that the excessive repetitive DNA within the targeting vector can significantly reduce targeting frequency (Wu et al., 2008). (ii) Use isogenic DNA as the source for producing exogenous homologous arms. While gene targeting in domestic species by using non-isogenic DNA as a source for targeting arms is possible (McCreath et al., 2000; Denning et al., 2001a and b; Kuroiwa et al., 2004; Marques et al., 2006; Richt et al., 2007), the use of isogenic DNA can largely improve the efficiency of gene targeting (te Riele et al., 1992). (iii) Increase the length of continuous exogenous homologous arms. The efficiency of gene targeting was generally found to be increased with the length of targeting arms in mouse ES cells (Hasty et al., 1991a). (iv) Use multiple cell lines; when targeting the CFTR (cystic fibrosis transmembrane conductance receptor) gene in pig fetal fibroblast cell lines, Rogers et al. (2008) demonstrated drastic differences in the targeting efficiencies between cell lines derived from littermate pig fetuses.

3.1.2 Tissue specificity and inducible promoters

As mentioned, the promoter used in a transgenic DNA construct determines when, where and to what extent the transgene is expressed. Hundreds of promoters can be isolated for the expression of transgenes, and are generally classified as constitutive promoters, tissue/cell-specific or Developmental stage-specific promoters and inducible promoters. Constitutive promoters, such as commonly used SV40 (simian virus 40 promoter), CMV (cytomegalovirus immediate-early promoter), PGK (mouse phosphoglycerate kinase 1 promoter) and CAGG (chicken β-Actin promoter coupled with CMV early enhancer) promoters for mammalian systems (Qin et al., 2010), can continuously drive transgene expression in all tissues and species. Tissue/cell-specific or developmental stage-specific promoters can restrict transgene expression to specific tissue(s) or only during certain developmental stages. For example, the 5.5-Kb osteopontin (OPN) promoter has been used to drive GFP expression in transgenic mice in the same cell-specific and developmental stage-specific manner as endogenous OPN expression (Higashibata et al., 2004). Inducible promoters, as their name suggests, may be activated by the presence of endogenous or exogenous factors. Exogenous factors include chemical compounds such as antibiotics or physical factors such as heat and light. For example, TRE promoter (tetracycline-responsive element promoter) can be activated by the rtTA (reverse tetracycline-controlled transcriptional activator) in a doxycycline-inducible manner (Qin et al., 2010). Antibiotic-induced promoters are the most commonly used in animal genetic modification because of easy manipulation. Inducible promoters provide a very useful tool in animal genetic engineering to turn on or off transgene expression in a particular tissue or at certain developmental stages.

3.1.3 Positive/negative selection strategies

Considering the rarity of a homologous recombination event relative to random integration of the targeting vector, an efficient targeting vector design should incorporate a good selection strategy, providing a powerful tool to improve the frequency of targeted colonies and reduce screening cell lines that result from random integration. Promoter-less gene targeting vector, also referred to the promoter-trapping method, has been used to enrich the gene targeting events in somatic cells in pigs and sheep (McCreath et al., 2000; Denning et al., 2001a and b). In a promoter-trapping vector, the selectable gene lacks its own promoter but it becomes activated from the target gene promoter after correctly integrating into the genome. In a fibroblast cell line, promoter-less vectors can enrich targeting frequency 5,000- to 10,000-fold (Hanson and Sedivy, 1995). Despite this potential improvement, the major limitation of the promoter-trapping method is that it requires active transcription of the targeted gene to drive expression of the selectable marker used in the targeting vector. Thus, if the target gene is only active in cell types that are difficult to culture, it is nearly impossible to target the gene locus using this method. Compared with promoter-trapping method, a more widely used strategy includes utilization of both positive and negative selection (PNS) (Jin et al., 2003; Kuroiwa et al., 2004; Richt et al., 2007). A positive selection gene (antibiotic-resistant gene such as neo) in the targeting vector is needed to select cells with an integrated construct a negative selection gene (cytotoxic genes such as the thymidine kinase (TK) gene or the diphtheria toxin A-chain (DT-A) gene) to further select against random integration event. The negative selection marker in the targeting vector usually is placed downstream of homologous arms and recombined away during the process of homologous recombination. The enrichment of the PNS selection is the ratio of clones recovered with the positive selection only (PS) versus the positive and negative dual selections. The PNS strategy can be used to target both active and inactive gene loci. It has been shown that the targeting efficiency at the COL1A1 locus in fibroblasts by the promoter-trapping strategy is 15.7-fold higher than by PS only (Marques et al., 2006).

3.2 Delivery of DNA vector into cultured somatic cells

Following vector construction and preparation of the DNA construct for delivery into a somatic cell, it is important to identify an efficient method of DNA delivery into somatic cells as efficiency between delivery methods varies greatly. There are numerous methods to introduce an exogenous DNA construct into somatic cells, which can be categorized as liposome-mediated DNA transfer (Hyun et al., 2003a; Lee et al., 2005), electroporation (Dai et al., 2002; Ramsoondar et al., 2003; Watanabe et al., 2005) and viral-mediated delivery (Lai et al., 2002b; Rogers et al., 2008).

3.2.1 Liposome-mediated DNA delivery

Liposome-mediated DNA transfer can easily transfect a large number of somatic cells without the need of specialized equipment and expertise, compared with other methods. Lee et al. (2005) demonstrated the efficiency of gene transfection with a plasmid containing the enhanced green fluorescence protein gene into fetal-derived bovine fibroblast cells by lipids was significantly higher than that obtained by electroporation. They also validated that transfection efficiency in fetal-derived bovine fibroblast cells, regardless of the delivery methods, was significantly higher than delivering DNA into cumulus-derived fibroblast cells and adult ear skin-derived fibroblast cells, establishing that both delivery method and cell line origin affect the efficiency of gene transfection. Using liposome-mediated DNA delivery followed by SCNT, genetically modified pigs (Hyun et al., 2003a) and sheep (McCreath et al., 2000) have successfully been created.

3.2.2 Electroporation

In contrast to liposome-mediated DNA delivery, electroporation has been widely used for the delivery of exogenous DNA into the cytoplasm of somatic cells to generate genetically modified cell lines for nuclear transfer. Electroporation has been utilized to successfully provide genetically modified donor cells for SCNT to create transgenic cloned domestic animals, including cattle (Kuroiwa et al., 2004), goat (Yu et al., 2006; Zhu et al., 2009), pig (Dai et al., 2002; Lai et al., 2002a; Ramsoondar et al., 2003; Watanabe et al., 2005) and sheep (Denning et al., 2001a). Ross et al. (2010a) indentified optimal electroporation conditions (three 1 ms pulses of 300 V to 200 µL of 1x10^6 cells/mL in the presence of 12.5µg DNA/mL), which can consistently deliver DNA vector into the 65-80% surviving porcine fetal fibroblasts and have been used to produce healthy, viable transgenic piglets (Ross et al., 2009b). In adult rhesus macaque fibroblasts, it has been demonstrated that electroporation can generate more transfected cells than liposome-mediated methods (Meehan et al., 2008), which is consistent with other similar comparisons (Yáñez and Porter, 1999). Of the numerous delivery methods, electroporation was demonstrated to have the greatest efficiency in generating targeted cell lines via homologous recombination (Vasquez et al., 2001). Targeting the hypoxanthine guanine phosphoribosyl transferase (HPRT) locus, Mir and Piedrahita (2004) demonstrated that the electroporation of a DNA construct with a nuclear localization signal into s-phase synchronized cells can increase targeting efficiency sevenfold and decrease random integration events 54-fold in primary fetal bovine fibroblasts. Later, Meehan et al. (2008) confirmed this method by successful gene targeting of the HPRT locus in adult rhesus macaque fibroblasts achieved by electroporation of S-phase synchronized cells with a construct containing a SV40 enhancer.

3.2.3 Viral-mediated delivery

Because viruses have the natural ability to stably transfect somatic cells with high efficiency, utilizing viral particles to delivery exogenous DNA into somatic cells has been widely successful. In contrast to liposome-mediated delivery and electroporation, which deliver linear, double-strand DNA into the cytoplasm of somatic cells, viral delivery of exogenous DNA delivers a high number of the linear, intact single-strand DNA molecules into the nucleus of cells. One report has demonstrated the efficient targeting of the PRNP gene encoding the prion protein PrP in bovine fetal fibroblasts by adeno-associated virus (AAV) vectors (Hirata et al., 2004). Also, the same transfection method followed by SCNT was used to successfully produce the CFTR-null and CFTR-DeltaF508 heterozygous pigs and CFTR-deficient ferrets (Rogers et al., 2008; Sun et al., 2008). Despite high transfection efficiency and production of transgenic cloned animals using viral-mediated delivery method, the need to produce high concentrations of virus particles in addition to limitations of the size of DNA capable of being delivered via virus limits the application of this method.

3.3 Selection and characterization of transgenic cell lines

3.3.1 Selection by marker

Following DNA construct delivery into somatic cells, transfected cells are cultured 24-48h in the absence of selection, followed by selection. Selection agents are chosen according to the DNA construct and added in cell-type specific concentrations into the cell culture medium. Typically, G418 (Geneticin) is used when the neomycin resistance gene is the positive selection marker and gangciclovir when using the TK gene for negative selection. Due to the limited lifespan of many of the somatic cell lines that are typically used for nuclear transfer and the significant amount of time to produce clonal colonies, it is important to perform the genetic screening by PCR or southern blotting as early as possible, prior to cell senescence.

3.3.2 Genotyping by PCR or southern blotting

Screening selected colonies by PCR for a gene targeting event typically involves using one PCR primer specific to the host genome sequence and the other primer specific to the selectable marker between the arms of homology in the targeting vector. This approach provides a simple, rapid and highly sensitive method to identify the gene targeting event in the transfected cells (Gómez-Rodríguez et al., 2008). Furthermore, PCR can be performed using the lysate from the small amount of cells, and also can be facilitated by pooled analysis of multiple cell lines. These types of PCR analysis sometimes require optimization as the amplicon is typically several thousand bp depending on the length of the targeting arms. Following the initial screening via PCR, the targeted colonies can be further expanded for analysis by southern blotting. A strategy for identifying targeted cells by southern blotting should be incorporated in the vector design. Two DNA probes on each side of the wild type gene, but outside the targeting vector, are designed to detect a change in fragment size resulting from either introduction or elimination. It is necessary to verify that the probes for southern blotting work well on wild-type genomic DNA from the somatic cells before the targeting experiment. Also, it is important to perform southern blotting analysis using both probes to confirm the double-crossover event since homologous recombination can occur at only one end of the targeting vector (Hasty et al., 1991b; Moens et al., 1992). Of the methods of screening for gene targeting in somatic cells, southern blotting is considered to be the golden approach to identify correctly targeted colonies despite being time-consuming and relatively expensive compared to PCR approaches.

3.3.3 Fluorescence In situ hybridization

Fluorescence in situ hybridization (FISH) is most commonly used to identify the number of integration sites following random integration of a transgene. FISH has been successfully used to detect the number of bacterial artificial chromosome (BAC) sequences integrated and chromosomal location(s) in mouse ES cells (Yang and Seed, 2003). Although the chromosomal location can be shown directly by FISH, Gómez-Rodríguez et al. (2008) found that screening of BAC-based constructs by FISH can be prone to false positives because of small pieces of integrated DNA that are below the limits of detection by fluorescent hybridization.

3.4 SCNT using genetically modified somatic cells

Once the cell line containing the appropriate genetic modification has been identified, those cells can then be used for SCNT as described in Figure 1. SCNT is a process in which the nucleus of a somatic cell is transferred into the cytoplasm of an enucleated oocyte. The hereby reconstructed SCNT embryos can be activated to initiate development and then are transferred into a synchronized surrogate mother immediately or after short-term *in vitro* culture. In addition to the technical factors involved in the SCNT process, the status of the donor cells and the quality of unfertilized oocytes are considered to largely affect the overall efficiency of the SCNT. This point will be discussed more below.

4. Improvement of the efficiency of cell-based transgenesis via SCNT

Since cell-based transgenesis via SCNT relies extensively on both transgenic and SCNT techniques, the efficiency could be improved by increasing SCNT efficiency and/or gene transfer efficiency.

4.1 Factors affecting SCNT efficiency

SCNT is a technique that requires precise skills in micromanipulation. In addition to technical skills, numerous other factors also impact SCNT efficiency, including donor cell types, quality of the recipient oocytes, cell cycle synchronization of both donor cell and recipient cytoplasm, the epigenetic status of the donor cell, method of reconstructed zygote activation, epigenetic reprogramming following activation, in vitro culture conditions of reconstructed embryos and embryo transfer into surrogate mothers. Here, we will focus on the cell cycle synchronization and methods being used to improve nuclear remodeling and reprogramming during SCNT.

4.1.1 Type of cell and synchronization of cell and recipient oocyte

Different types of donor cells and the different origins of recipient oocyte have been successfully used to produce the cloned animals in domestic species (Table 2). Generally, nuclei from less differentiated donor cells demonstrate greater SCNT efficiency than those from differentiated donor cells (Hill et al., 2000; Lee et al., 2003). SCNT embryos created from *in vivo*-derived MII oocytes have better developmental potential than *in vitro*-maturated MII oocytes; and *in vitro*-maturated oocytes from sexually mature animals have been demonstrated to more efficiently produce cloned embryos than oocytes *in vitro*-matured from prepubertal animals (Lai et al., 2002a; Lee et al., 2003; Hyun et al., 2003b).

Since the first SCNT animal was produced by transferring the nuclei from quiescent cells (G0), synchronization of the donor nuclei into G0/G1 is thought to be crucial to the success of SCNT (Wilmut et al., 1997). The G0 cells, with lower transcription activity and different chromatin configuration in contrast to the cells at other stages of the cell cycle, may be more responsive to factors inside the recipient oocyte cytoplasm that impact the nuclear remodeling and reprogramming process following SCNT. The cell cycle of cultured somatic cells can be synchronized by serum starvation, contact inhibition and chemical treatments (Cho et al., 2005; Gibbons et al., 2002). Serum starvation to induce the donor cells into G0 is widely used for the production of cloned transgenic animals (Table 2). However, recent studies have demonstrated that synchronizing the cell cycle of donor cells by serum starvation can cause apoptosis (Dalman et al., 2010) and reduce blastocyst production in cattle (Miranda Mdos et al., 2009). The production of the cloned transgenic calves from non-quiescent fetal fibroblasts demonstrated that the synchronization of the donor cells is not an absolute requirement for SCNT success (Cibelli et al., 1998). However, the relative cell cycle combination of both donor cells and recipient oocyte, not just donor cells, is thought to be important for the maintenance of correct ploidy and the subsequent development of reconstructed embryos (reviewed by Campbell, 1999).

4.1.2 Methods to improve nuclear remodeling and reprogramming

Following transfer of a donor nucleus into a recipient oocyte cytoplasm and subsequent activation, nuclear remodeling events of the chromatin structure, such as changes of DNA methylation patterns and histone modifications, result in the reprogramming of gene expression to recapitulate developmental patterns observed in a normal fertilized embryo (Whitworth and Prather, 2010). In contrast to less differentiated donor nuclei, the relatively high level of DNA methylation and low histone acetylation exist in the chromatin of the highly differentiated nuclei. These epigenetic modifications are used to maintain the temporal and spatial patterns of gene expression specific to the cell type or developmental stage. When differentiated nuclei are transferred into the enucleated oocyte cytoplasm, correctly establishing normal patterns of zygotic gene expression is crucial to the full term development of SCNT animals. However, numerous studies have demonstrated improper reprogramming of genes in embryos and tissues of domestic animals following SCNT (Wrenzycki et al., 2001; Pfister-Genskow et al., 2005; Aston et al., 2010; Ross et al., 2010b). Thus, various strategies are under development to facilitate and promote appropriate nuclear reprogramming of the transferred nucleus following nuclear transfer or to pre-program the genome of the donor nucleus prior to SCNT. Of these strategies, one widely used is to treat reconstructed SCNT embryos, but not the donor cells, with histone deacetylase inhibitors (HDACi), such as trichostatin A (TSA), 6-(1,3-dioxo-1H, 3H-benzo [de] isoquinolin-2-yl)-hexanoic acid hydroxyamide (Scriptaid), sodium butyrate and valproic acid. It has been demonstrated that histone deacetylase inhibitor treatment after SCNT can improve both *in vitro* development of SCNT embryos to the blastocyst stage and *in vivo* development to term following embryo transfer (Li et al., 2008; Cervera et al., 2009; Zhao et al., 2009a, 2010; Das et al., 2010; Himaki et al., 2010a; Miyoshi et al., 2010). However, these treatments may be with some level of toxicity, as one group has reported that offspring from TSA-treated rabbit embryos did not survive to adulthood (Meng et al., 2009). The exact mechanism by which the HDACi treatment significantly improves the cloning efficiency remains largely unknown, although it has been shown to increase levels of global histone acetylation after HDACi treatment which may subsequently change the structure of chromatin and improve nuclear reprogramming (Shi et al., 2008; Iager et al., 2008; Zhao et al., 2010; Das et al., 2010).

Accruing investigations have demonstrated that abnormal DNA methylation patterns contribute to the lower developmental competency of SCNT derived embryos (Kang et al., 2001, 2002; Bourc'his et al., 2001; Santos et al., 2003; Wrenzycki et al., 2006), suggesting the inability of oocyte to fully restore the DNA methylation pattern of differentiated donor nuclei to that of normal totipotent 1-cell stage embryos. Thus, the second method used to assist epigenetic reprogramming is to reduce the DNA methylation level in donor cells or reconstructed embryos by treating them with the DNA methyl-transferase inhibitor such as 5-aza-2'-deoxycytidine (5-aza-dC). Unfortunately, previous studies have demonstrated that 5-aza-dC treatment of donor cells or cloned embryos does not improve the *in vitro* and *in vivo* development of SCNT derived embryos (Enright et al., 2005; Tsuji et al., 2009). However, combining the treatment of donor cells and embryos with both TSA and 5-aza-dC resulted in improved blastocyst development (Ding et al., 2008). Furthermore, an additional study has demonstrated enhanced gene targeting frequency in ES cells with low genomic methylation levels, suggesting the epigenetic status of targeted loci may influence the efficiency of gene targeting by affecting the accessibility for the homologous recombination machinery (Domínguez-Bendala and McWhir, 2004).

An additional strategy to promote developmental reprogramming of cloned embryos is to treat them with latrunculin A (LatA), an actin polymerization inhibitor. One group has reported that post-activation treatment with LatA is effective to improve *in vitro* developmental capacity of gene-modified cloned miniature pig embryos and embryos treated with LatA have the ability to develop into fetuses (Himaki et al., 2010b). Pre-reprogramming donor nuclei prior to SCNT has also been attempted. Rathbone et al (2010) reported that pretreatment of permeabilized ovine fetal fibroblasts with a cytoplasmic extract produced from germinal vesicle (GV) stage *Xenopus laevis* oocytes improves the live birth rate, but not development to blastocyst stage or pregnancy rate following embryo transfer.

4.2 Methods to improve gene transfer efficiency

Production of transgenic domestic animals has been widely accomplished, however, several limitations remain. Typically, plasmid based DNA constructs are limiting in the size of exogenous DNA to transfer and creating animals with targeted genetic modifications has been significantly more challenging. Thus, it is important to continue development of new strategies that broaden application and increase the efficiency of creating targeted genetic modifications in domestic animals.

4.2.1 Artificial chromosomes as DNA transfer vector

In contrast to plasmid based vector, artificial chromosomes have the capacity to carry Megabase-sized pieces of DNA that are maintained as autonomous, replicating chromosomes. Artificial chromosome vectors include a centromere, two telomeres and origins of replication (Robl et al., 2007). A 10 Mb human artificial chromosome (HAC) vector containing the entire unarranged sequences of the human immunoglobulin heavy and light chain loci (1-1.5 Megabase for each locus) have been transferred into bovine fibroblasts using a microcell-mediated chromosome transfer approach. Following SCNT using selected cells, trans-chromosomal cloned bovine offspring were produced that expressed human immunoglobulin proteins in the blood. HAC was retained as an independent chromosome with the proportion of cells ranged from 78 to 100% in most animals and HAC retention rate has not changed over several years. This system provides a useful tool to produce human therapeutic polyclonal antibodies using trans-chromosomal cloned domestic animals (Kuroiwa et al., 2002; Robl et al., 2007). In domestic animals, the swine artificial chromosome (SAC) (about 310 kb) containing pig centromeric DNA and the neomycin resistance gene was constructed and introduced into pig cell lines, and one positive clone was characterized, showing the possibility for producing transgenic pigs for xenotransplantation and other purposes (Poggiali et al., 2002).

4.2.2 Zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALNs)

The low frequency of homologous recombination hampers rapid progress and wide application of gene targeting in domestic animals. The generation of a site-specific double-stranded DNA break (DSB) within the desired locus can facilitate gene targeting by resulting in additions and deletions causing inactivation of gene function. Naturally occurring DNA-binding proteins, including zinc finger proteins (ZFPs) and meganucleases, have been engineered to bind site-specific DNA sequence. Zinc finger nucleases (ZFNs) combine the specific DNA-binding domain (ZFP) with the non-specific cleavage domain of the restriction endonuclease *FokI* and offer powerful tools to create a site-specific DSB to facilitate local homologous recombination. The ZFN induced DSB can lead to incorporating exogenous DNA in a site specific manner by utilization of a homologous recombination targeting vector that overlaps with the DSB region. Additionally, DSBs repaired by non-homologous end-joining (NHEJ) can result in loss of a single nucleotide, multiple nucleotides, or small regions; all capable of rendering a targeted gene dysfunctional. For increasing the specificity of DNA binding, multiple zinc fingers, each recognizing and binding to a 3-bp sequence of DNA nucleotides can be linked in tandem to recognize a unique genomic locus. In human somatic cells, custom-designed ZFNs yielded more than 18% targeting efficiency at the X-linked interleukin-2 receptor gamma gene locus and about 7% of the cells possessed a bi-allelic gene modification (Urnov et al., 2005). ZFNs can also promote the addition of novel DNA sequence into a targeted endogenous locus of human cells at a frequency ranging from 5% to 15% depending on the size of extra-chromosomal DNA (Moehle et al., 2007). Thus, the efficiency of gene targeting using ZFNs offers the possibility of gene therapy for human genetic disease and an approach for improvement of genetic engineering in domestic animals. Recently, targeted gene disruption of exogenous EGFP gene was achieved in porcine somatic cells using ZFNs (Watanabe et al., 2010) and transgenic EGFP knockout pigs were produced using ZFNs followed by SCNT (Whyte et al., 2011).

The appropriate design of a site-specific ZFN is critical to successfully introduce ZFN-mediated genetic modifications. Of the available zinc-finger engineering methods, modular assembly is the most easily performed method, but is also associated with high failure rates to yield a functional three–zinc finger array for the majority of potentially targetable sites (Ramirez et al, 2008). The second method is to combine selection-based methods with ZFNs made using modular assembly. While effective, this method is labor intensive and requires additional expertise. The third method, referred to as OPEN (Oligomerized Pool ENgineering) is based on bacterial 2-hybrid (B2H) selection and has been proven to be a rapid platform for plant and human cells with high targeting efficiency (ranging from 1 to 50% at different loci) and less toxicity compared to modular assembly system. However, utilization of OPEN requires an archive of pre-selected zinc-finger pools and *E. coli* selection (Maeder et al., 2008). Furthermore, due to the challenge of engineering the endonucleases, orthophenanthroline (OP, a DNA cleaving molecule) was conjugated with triplex-forming oligonucleotides (TFOs, sequence-specific binding capacity) to induce targeted DSBs and stimulate mutations at the target site in approximately 10% of treated human cells (Cannata et al., 2008). TFO conjugating to OP or other DNA cleaving molecules may provide a useful tool to induce targeted gene modification because triplex-forming sequences are frequent in mammalian genes. While ZFN-driven gene targeting can be much more efficient than homologous recombination-based methods, the design and development of highly specific ZFNs remain difficult because of the lack of a simple correspondence between amino acid sequence and DNA recognition sequence.

Recently, several groups have shown that transcription activator–like effectors (TALEs) from the bacterial genus *Xanthomonas* contain a central domain of tandem repeats that can be readily engineered to bind virtually any DNA sequence (Boch et al., 2009; Christian et al., 2010; Morbitzer et al., 2010). The structure of the central protein domain, highly conserved in all the known TALEs, includes 17.5 tandem repeats with 34 amino acids per repeat. In each repeat monomer of a TALE, only amino acid positions 12 and 13 are hypervariable (repeat variable diresidues) (Boch et al., 2011), which can specifically recognize a single nucleotide in the target site (Boch et al., 2009; Moscou et al., 2009). Thus, the correspondence between each repeat variable diresidues and the binding nucleotide in DNA sequence opens the possibility to create novel sequence-specific DNA binding proteins by rearrangement of TALE repeats. The engineered hybrid TALE nucleases (TALNs), produced by fusion of the *FokI* endonuclease domain with the high-specificity DNA-binding domains of TALEs, can bind and create targeted DSBs in tobacco and yeast (Mahfouz et al., 2011; Li et al., 2011), showing the feasibility of engineering TALE-based hybrid nucleases capable of generating site-specific genome modification. Recently, Miller et al. (2011) reported the generation of discrete edits or small deletions within endogenous human NTF3 and CCR5 genes and the insertion of 46-bp sequence at CCR5 locus into the genome of human K562 cells using designed TALNs, demonstrating the effective application of TALNs to modify endogenous genes. While the simple DNA-binding code of TALEs enables easier design strategies as compared to ZFPs, the repetitive nature of TALE DNA-binding domains results in difficulty to efficiently synthesize new TALEs by currently used vector construction methods. To overcome this problem, Zhang et al. (2011) recently developed a new strategy to construct repeat domains of TALEs by hierarchical ligation.

4.3 Induced pluripotent stem (iPS) cells

Owing to the lack of ES cells in domestic animals, it is difficult to replicate strategies routinely used to create genetically modified mice. As an alternative, cell-based transgenesis via SCNT is currently used to produce genetically modified domestic animals. Recent advancements in the ability to generate induced pluripotent stem (iPS) cells may open another potential strategy to improve the efficiency of SCNT in domestic animals. Induced pluripotent stem cells in mice and human were successfully generated by reprogramming somatic cells with viral delivery of a combination of four defined transcription factors including Sox2, Oct4, Klf4, and c-Myc or Sox2, Oct4, Nanog, and Lin28 (Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Yu et al., 2007). The iPS cells are similar to ES cells in morphology, biochemistry, gene expression and the ability to differentiate into many cell types and self renew (Wernig et al., 2007; Lowry et al., 2008). Furthermore, subsequent studies have optimized existing procedures and discovered novel reprogramming protocols to generate the iPS cells, including non-integrating viruses (Stadtfeld et al., 2008), non-viral vectors (Okita et al., 2008), non-integration episomal vectors (Yu et al., 2009) and RNA-induced reprogramming (Warren et al., 2010), which greatly decrease the biosafety concerns associated with the application of iPS cells. The use of iPS cells can benefit animal transgenesis in several aspects. (i) Stable genetic modification in iPS cells may be more efficient compared to somatic cells as is the case with ES cells. (ii) Genetically modified iPS cells can be used to produce chimeric animals since it has been reported that iPS cells produce viable, live-born and fertile mice offspring through tetraploid complementation (Zhao et al., 2009b; Boland et al., 2009). (iii) The use of genetically modified iPS cells as donors may increase the efficiency of cell-based transgenesis via SCNT owing to the pluripotent status of iPS cells. (iv) In contrast to the limited lifespan of somatic cells, true iPS cells are immortalized. These advantages, coupled to successful derivation of iPS cells from domestic animals (Esteban et al., 2009; Ezashi et al., 2009; Wu et al., 2009) presents a new opportunity to produce transgenic animals using iPS cells.

5. Summary

While the potential opportunities of transgenic domestic animals in biomedicine and agriculture are significant, current procedures, including cell-based transgenesis via SCNT, to produce genetically modified domestic animals are not without limitations. The combination of new technologies, including ZFNs/TALNs to enhance targeted genome modification and iPS cells and other strategies to improve epigenetic remodeling of SCNT embryos, represent pathways for improving the success rates of current genome manipulation strategies resulting in transgenic domestic animals. These modern approaches may have limitations of their own, such as the difficulty and high cost to design, produce and validate the target-specific ZFPs, constructing custom-designed TALEs, and maintenance of iPS cells. However, despite these limitations, we expect to see these strategies become widely utilized as a result of the potential opportunities that utilization of these strategies offers to the field of targeted genome manipulation in domestic animals.

6. References

Aston, K.I., Li, G.P., Hicks, B.A., Sessions, B.R., Davis, A.P., Rickords, L.F., Stevens, J.R., White, K.L. 2010. Abnormal levels of transcript abundance of developmentally important genes in various stages of preimplantation bovine somatic cell nuclear transfer embryos. Cell Reprogram. 12(1): 23-32.

Auchincloss, H. Jr., Sachs, D.H. 1998. Xenogeneic transplantation. Annu Rev Immunol 16: 433-470.

Bhatti, F.N., Zaidi, A., Schmoeckel, M., Cozzi, E., Chavez, G., Wallwork, J., White, D.J., Friend, P.J. 1998. Survival of life-supporting HDAF transgenic kidneys in primates is enhanced by splenectomy. Transplant Proc 30(5): 2467.

Bleck, G.T., White, B.R., Miller, D.J., Wheeler, M.B. 1998. Production of bovine alpha-lactalbumin in the milk of transgenic pigs. J Anim Sci 76(12):3072-3078.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., Bonas, U. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326(5959): 1509-1512.

Boch, J. 2011. TALEs of genome targeting. Nat Biotechnol 29(2): 135-136.

Boland, M. J., Hazen, J. L., Nazor, K. L., Rodriguez, A.R., Gifford, W., Martin, G., Kupriyanov, S., Baldwin, K.K. 2009. Adult mice generated from induced pluripotent stem cells. Nature 461 (7260): 91-94.

Bordignon, V., Keyston, R., Lazaris, A., Bilodeau, A.S., Pontes, J.H., Arnold, D., Fecteau, G., Keefer, C., Smith, L.C. 2003. Transgene expression of green fluorescent protein and germ line transmission in cloned calves derived from in vitro-transfected somatic cells. Biol Reprod 68: 2013-2023.

Bourc'his, D., Le Bourhis, D., Patin, D., Niveleau, A., Comizzoli, P., Renard, J.P., Viegas-Pequignot, E. 2001. Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. Curr Biol 11: 1542-1546.

Brophy, B., Smolenski, G., Wheeler, T., Wells, D., L'Huillier, P., Laible, G. 2003. Cloned transgenic cattle produce milk with higher levels of beta-casein and kappa-casein. Nat Biotechnol 21(2):157-162.

Byrne, G.W., McCurry, K.R., Martin, M.J., McClellan, S.M., Platt, J.L., Logan, J.S. 1997. Transgenic pigs expressing human CD59 and decay-accelerating factor produce an intrinsic barrier to complement mediated damage. Transplantation 63: 149.

Cabot, R.A., Kühholzer, B., Chan, A.W., Lai, L., Park, K.W., Chong, K.Y., Schatten, G., Murphy, C.N., Abeydeera, L.R., Day, B.N., Prather, R.S. 2001. Transgenic pigs produced using in vitro matured oocytes infected with a retroviral vector. Anim Biotechnol 12(2):205-214.

Campbell, K.H. 1999. Nuclear transfer in farm animal species. Semin Cell Dev Biol. 10(3):245-252.

Cannata, F., Brunet, E., Perrouault, L., Roig, V., Ait-Si-Ali, S., Asseline, U., Concordet, J.P., Giovannangeli, C. 2008. Triplex-forming oligonucleotide-orthophenanthroline conjugates for efficient targeted genome modification. Proc Natl Acad Sci USA 105(28): 9576-9581.

Castro, F.O., Hernandez, O., Uliver, C., Solano, R., Milanes, C., Aguilar, A., Perez, R., De Armas, R., Herrera, N., Fuente, J.D.L. 1991. Introduction of foreign DNA into the spermatozoa of farm animals. Theriogenology 34: 1099-1110.

Cervera, R.P., Martí-Gutiérrez, N., Escorihuela, E., Moreno, R., Stojkovic, M. 2009. Trichostatin A affects histone acetylation and gene expression in porcine somatic cell nucleus transfer embryos. Theriogenology 72(8): 1097-1110.

Chan, A.W., Homan, E.J., Ballou, L.U., Burns, J.C., Bremel, R.D. 1998. Transgenic cattle produced by reverse-transcribed gene transfer in oocytes. Proc Natl Acad Sci USA 95(24): 14028-14033.

Chang, K., Qian, J., Jiang, M., Liu, Y.H., Wu, M.C., Chen, C.D., Lai, C.K., Lo, H.L., Hsiao, C.T., Brown, L., Bolen, J. Jr., Huang, H.I., Ho, P.Y., Shih, P.Y., Yao, C.W., Lin, W.J., Chen, C.H., Wu, F.Y., Lin, Y.J., Xu, J., Wang, K. 2002. Effective generation of transgenic pigs and mice by linker based sperm-mediated gene transfer. BMC Biotechnol 2: 5.

Chen, R.H., Naficy, S., Logan, J.S., Diamond, L.E., Adams, D.H. 1999. Hearts from transgenic pigs constructed with CD59/DAF genomic clones demonstrate improved survival in primates. Xenotransplantation 6: 194.

Cho, S.R., Ock, S.A., Yoo, J.G., Mohana kumar, B., Choe, S.Y., Rho, G.J. 2005. Effect of confluent, Roscovitine treatment and serum starvation on the cell-cycle synchronization of bovine foetal fibroblasts. Reprod Domest Anim 40: 171-176.

Choi, T., Huang, M., Gorman, C., Jaenisch, R. 1991. A generic intron increases gene expression in transgenic mice. Molecular Cellular Biology 11: 3070-3074.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., Voytas, D.F. 2010. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186(2): 757-761.

Cibelli, J.B., Stice, S.L., Golueke, P.J., Kane, J.J., Jerry, J., Blackwell, C., Ponce de Leon, F. A., Robl. J.M. 1998. Cloned transgenic calves produced from nonquiescent fetal fibroblasts. Science 280: 1256-1258.

Clements, J.E., Wall, R.J., Narayan, O., Hauer, D., Schoborg, R., Sheffer, D., Powell, A., Carruth, L.M., Zink, M.C., Rexroad, C.E. 1994. Development of transgenic sheep that express the visna virus envelope gene. Virology 200: 370-380.

Cooper, D.K., Ezzelarab, M., Hara, H., Ayares, D. 2008. Recent advances in pig-to-human organ and cell transplantation. Expert Opin Biol Ther 8(1): 1-4.

Cozzi, E., White, D.J.G. 1995. The generation of transgenic pigs as potential organ donors for humans. Nat Med 1: 964-966.

Dai, Y., Vaught, T.D., Boone, J., Chen, S.H., Phelps, C.J., Ball, S., Monahan, J.A., Jobst, P.M., McCreath, K.J., Lamborn, A.E., Cowell-Lucero, J.L., Wells, K.D., Colman, A., Polejaeva, I.A., Ayares, D.L. 2002. Targeted disruption of the alpha1,3-galactosyltransferase gene in cloned pigs. Nat Biotechnol 20: 251-255.

Dalman, A., Eftekhari-Yazdi, P., Valojerdi, M., Shahverdi, A., Gourabi, H., Janzamin, E., Fakheri, R., Sadeghian, F., Hasani, F. 2010. Synchronizing Cell Cycle of Goat Fibroblasts by Serum Starvation Causes Apoptosis. Reprod Domest Anim. 45: e46–e53.

Das, Z.C., Gupta, M.K., Uhm, S.J., Lee, H.T. 2010. Increasing histone acetylation of cloned embryos, but not donor cells, by sodium butyrate improves their in vitro development in pigs. Cell Reprogram 12(1): 95-104.

Denning, C., Burl, S., Ainslie, A., Bracken, J., Dinnyes, A., Fletcher, J., King, T., Ritchie, M., Ritchie, W.A., Rollo, M., de Sousa, P., Travers, A., Wilmut, I., Clark, A.J. 2001a. Deletion of the lpha(1,3)galactosyl transferase (GGTA1) gene and the prion protein (PrP) gene in sheep. Nat Biotechnol 19: 559-562.

Denning, C., Dickinson, P., Burl, S., Wylie, D., Fletcher, J. Clark, A.J. 2001b. Gene targeting in primary fetal fibroblasts from sheep and pig. Cloning Stem Cells 3(4): 221-231.

Diamond LE, McCurry KR, Martin MJ, McClellan SB, Oldham ER, Platt JL, Logan JS. 1996. Characterization of transgenic pigs expressing functionally active human CD59 on cardiac endothelium. Transplantation 61(8): 1241-1249.

Diamond, L.E., Quinn, C.M., Martin, M.J., Lawson, J., Platt, J.L., Logan, J.S. 2001. A human CD46 transgenic pig model system for the study of discordant xenotransplantation. Transplantation 71(1): 132-142.

Ding X, Wang Y, Zhang D, Wang Y, Guo Z, Zhang Y. 2008. Increased pre-implantation development of cloned bovine embryos treated with 5-aza-2'-deoxycytidine and trichostatin A. Theriogenology 70(4): 622-630.

Diswall, M., Angström, J., Karlsson, H., Phelps, C.J., Ayares, D., Teneberg, S., Breimer, M.E. 2010. Structural characterization of alpha1,3-galactosyltransferase knockout pig heart and kidney glycolipids and their reactivity with human and baboon antibodies. Xenotransplantation 17(1): 48-60.

Domínguez-Bendala, J., McWhir, J. 2004. Enhanced gene targeting frequency in ES cells with low genomic methylation levels. Transgenic Res 13(1): 69-74.

Duncker, B.P., Davies, P.L., Walker, V.K. 1997. Introns boost transgene expression in Drosophila melanogaster. Mol Gen Genet 254: 291-296.

Ebert, K.M., Selgrath, J.P., DiTullio, P., Denman, J., Smith, T.E., Memon, M.A., Schindler, J.E., Monastersky, G.M., Vitale, J.A., Gordon, K. 1991. Transgenic production of a variant of human tissue-type plasminogen activator in goat milk: generation of transgenic goats and analysis of expression. Biotechnology (N Y) 9(9): 835-838.

Enright, B.P., Sung, L.Y., Chang, C.C., Yang, X., Tian, X.C. 2005. Methylation and acetylation characteristics of cloned bovine embryos from donor cells treated with 5-aza-2'-deoxycytidine. Biol Reprod 72(4): 944-948.

Esteban, M. A., Xu, J., Yang, J., Peng, M., Qin, D., Li, W., Jiang, Z., Chen, J., Deng, K., Zhong, M., Cai, J., Lai, L., Pei, D. 2009. Generation of induced pluripotent stem cell lines from Tibetan miniature pig. J Biol Chem 284(26): 17634-17640.

Ezashi, T., Telugu, B. P., Alexenko, A. P., Sachdev, S., Sinha, S., Roberts, R. M. 2009. Derivation of induced pluripotent stem cells from pig somatic cells. Proc Natl Acad Sci USA 106(27): 10993-10998.

Fodor, W.L., Williams, B.L., Matis, L.A., Madri, J.A., Rollins, S.A., Knight, J.W., Velander, W., Squinto, S.P. 1994. Expression of a functional human complement inhibitor in a transgenic pig as a model for the prevention of xenogeneic hyperacute organ rejection. Proc Natl Acad Sci USA 91(23): 11153-11157.

Galili, U. 1993. Interaction of the natural anti-Gal antibody with alpha-galactosyl epitopes: a major obstacle for xenotransplantation in humans (see comments). Immunol Today 14: 480-482.

García-Vázquez, F.A., Ruiz, S., Matás, C., Izquierdo-Rico, M.J., Grullón, L.A., De Ondiz, A., Vieira, L., Avilés-López, K., Gutiérrez-Adán, A., Gadea, J. 2010. Production of transgenic piglets using ICSI-sperm-mediated gene transfer in combination with recombinase RecA. Reproduction 140(2): 259-272.

Gibbons, J., Arat, S., Rzucidlo, J., Miyoshi, K., Waltenburg, R., Respess, D., Venable, A., Stice, S. 2002. Enhanced survivability of cloned calves derived from roscovitine-treated adult somatic cells. Biol Reprod 66: 895–900.

Gómez-Rodríguez, J., Washington, V., Cheng, J., Dutra, A., Pak, E., Liu, P., McVicar, D.W., Schwartzberg, P.L. 2008. Advantages of q-PCR as a method of screening for gene targeting in mammalian cells using conventional and whole BAC-based constructs. Nucleic Acids Res. 36(18): e117.

Grosse-Hovest, L., Müller, S., Minoia, R., Wolf, E., Zakhartchenko, V., Wenigerkind, H., Lassnig, C., Besenfelder, U., Müller, M., Lytton, S.D., Jung, G., Brem, G. 2004. Cloned transgenic farm animals produce a bispecific antibody for T cell-mediated tumor cell killing. Proc Natl Acad Sci USA 101(18): 6858-6863.

Hammer, R.E., Pursel, V.G., Rexroad, C.E. Jr, Wall, R.J., Bolt, D.J., Ebert, K.M., Palmiter, R.D., Brinster, R.L. 1985. Production of transgenic rabbits, sheep and pigs by microinjection. Nature 315(6021): 680-683.

Hanson, K.D., Sedivy, J.M. 1995. Analysis of biological selections for high-efficiency gene targeting. Mol Cell Biol. 15(1): 45-51.

Hasty, P., Rivera-Pérez, J., Bradley, A. 1991a. The length of homology required for gene targeting in embryonic stem cells. Mol Cell Biol. 11(11): 5586-5591.

Hasty, P., Rivera-Perez, J., Chang, C., Bradley, A. 1991b. Target frequency and integration pattern for insertion and replacement vectors in embryonic stem cells. Mol Cell Biol 11: 4509-4517.

Higashibata, Y., Sakuma, T., Kawahata, H., Fujihara, S., Moriyama, K., Okada, A., Yasui, T., Kohri, K., Kitamura, Y., Nomura, S. 2004. Identification of promoter regions involved in cell- and developmental stage-specific osteopontin expression in bone, kidney, placenta, and mammary gland: an analysis of transgenic mice. J Bone Miner Res 19(1): 78-88.

Hill, J.R., Winger, Q.A., Long, C.R., Looney, C.R., Thompson, J.A., Westhusin, M.E. 2000. Development rates of male bovine nuclear transfer embryos derived from adult and fetal cells. Biol Reprod 62(5): 1135-1140.

Himaki, T., Yokomine, T.A., Sato, M., Takao, S., Miyoshi, K., Yoshida, M. 2010a. Effects of trichostatin A on in vitro development and transgene function in somatic cell nuclear transfer embryos derived from transgenic Clawn miniature pig cells. Anim Sci J 81(5): 558-563.

Himaki, T., Mori, H., Mizobe, Y., Miyoshi, K., Sato, M., Takao, S., Yoshida, M. 2010b. Latrunculin A dramatically improves the developmental capacity of nuclear transfer embryos derived from gene-modified clawn miniature pig cells. Cell Reprogram 12(2): 127-131.

Hirata, R.K., Xu, C., Dong, R., Miller, D.G., Ferguson, S., Russell, D.W. 2004. Efficient PRNP gene targeting in bovine fibroblasts by adeno-associated virus vectors. Cloning Stem Cells 6(1): 31-36.

Huang, Y.J., Huang, Y., Baldassarre, H., Wang, B., Lazaris, A., Leduc, M., Bilodeau, A.S., Bellemare, A., Côté, M., Herskovits, P., Touati, M., Turcotte, C., Valeanu, L., Lemée, N., Wilgus, H., Bégin, I., Bhatia, B., Rao, K., Neveu, N., Brochu, E., Pierson, J., Hockley, D.K., Cerasoli, D.M., Lenz, D.E., Karatzas, C.N., Langermann, S. 2007. Recombinant human butyrylcholinesterase from milk of transgenic animals to protect against organophosphate poisoning. Proc Natl Acad Sci USA 104(34): 13603-13608.

Hyun, S., Lee, G., Kim, D., Kim, H., Lee, S., Nam, D., Jeong, Y., Kim, S., Yeom, S., Kang, S., Han, J., Lee, B., Hwang, W. 2003a. Production of nuclear transfer-derived piglets using porcine fetal fibroblasts transfected with the enhanced green fluorescent protein. Biol Reprod 69(3): 1060–1068.

Hyun, S.H., Lee, G.S., Kim, D.Y., Kim, H.S., Lee, S.H., Kim, S., Lee, E.S., Lim, J.M., Kang, S.K., Lee, B.C., Hwang, W.S. 2003b. Effect of maturation media and oocytes derived from sows or gilts on the development of cloned pig embryos. Theriogenology 59(7): 1641-1649.

Iager, A.E., Ragina, N.P., Ross, P.J., Beyhan, Z., Cunniff, K., Rodriguez, R.M., Cibelli, J.B. 2008. Trichostatin A improves histone acetylation in bovine somatic cell nuclear transfer early embryos. Cloning Stem Cells 10(3): 371-379.

Jin, D.I., Lee, S.H., Choi, J.H., Lee, J.S., Lee, J.E., Park, K.W., Seo, J.S. 2003. Targeting efficiency of a-1,3-galactosyl transferase gene in pig fetal fibroblast cells. Exp Mol Med. 35(6): 572-577.

Kang, Y.K., Koo, D.B., Park, J.S., Choi, Y.H., Chung, A.S., Lee, K.K., Han, Y.M. 2001. Aberrant methylation of donor genome in cloned bovine embryos. Nat Genet 28: 173–177.

Kang, Y.K., Park, J.S., Koo, D.B., Choi, Y.H., Kim, S.U., Lee, K.K., Han, Y.M. 2002. Limited demethylation leaves mosaic-type methylation states in cloned bovine pre-implantation embryos. EMBO J 21: 1092–1100.

Keefer, C. L., Baldassarre, H., Keyston, R., Wang, B., Bhatia, B., Bilodeau, A.S., Zhou, J.F., Leduc, M., Downey, B.R., Lazaris, A., Karatzas, C.N. 2001. Generation of dwarf goat (Capra hircus) clones following nuclear transfer with transfected and nontransfected fetal fibroblasts and in vitro-matured oocytes. Biol Reprod 64: 849-856.

Kjer-Nielsen, L., Holmberg, K., Perera, J.D., McCluskey, J. 1992. Impaired expression of chimaeric major histocompatibility complex transgenes associated with plasmid sequences. Transgenic Research 1: 182-187.

Kragh, P.M., Nielsen, A.L., Li, J., Du, Y., Lin, L., Schmidt, M., Bogh, I.B., Holm, I.E., Jakobsen, J.E., Johansen, M.G., Purup, S., Bolund, L., Vajta, G., Jorgensen, A.L. 2009. Hemizygous minipigs produced by random gene insertion and handmade cloning express the Alzheimer's disease-causing dominant mutation APPsw. Transgenic Res 18: 545-558.

Kuroiwa, Y., Kasinathan, P., Choi, Y.J., Naeem, R., Tomizuka, K., Sullivan, E.J., Knott, J.G., Duteau, A., Goldsby, R.A., Osborne, B.A., Ishida, I., Robl, J.M. 2002. Cloned transchromosomic calves producing human immunoglobulin. Nat Biotechnol 20: 889-894.

Kuroiwa, Y., Kasinathan, P., Matsushita, H., Sathiyaselan, J., Sullivan, E.J., Kakitani, M., Tomizuka, K., Ishida, I., Robl, J.M. 2004. Sequential targeting of the genes encoding immunoglobulin-mu and prion protein in cattle. Nat Genet 36(7): 775-780.

Lai, L., Kolber-Simonds, D., Park, K.W., Cheong, H.T., Greenstein, J.L., Im, G.S., Samuel, M., Bonk, A., Rieke, A., Day, B.N., Murphy, C.N., Carter, D.B., Hawley, R.J., Prather, R.S. 2002a. Production of a-1, 3-galactosyltransferase knockout pigs by nuclear transfer cloning. Science 295: 1089-1092.

Lai, L., Park, K.W., Cheong, H.T., Kuhholzer, B., Samuel, M., Bonk, A., Im, G.S., Rieke, A., Day, B.N., Murphy, C.N., Carter, D.B., Prather, R.S. 2002b. Transgenic pig expressing the enhanced green fluorescent protein produced by nuclear transfer using colchicine-treated fibroblasts as donor cells. Mol Reprod Dev 62(3): 300-306.

Lai, L., Kang, J.X., Li, R., Wang, J., Witt, W.T., Yong, H.Y., Hao, Y., Wax, D.M., Murphy, C.N., Rieke, A., Samuel, M., Linville, M.L., Korte, S.W., Evans, R.W., Starzl, T.E., Prather, R.S., Dai, Y. 2006. Generation of cloned transgenic pigs rich in omega-3 fatty acids. Nat. Biotechnol 24: 435-436.

Lavitrano, M., Bacci, M.L., Forni, M., Lazzereschi, D., Di Stefano, C., Fioretti, D., Giancotti, P., Marfé, G., Pucci, L., Renzi, L., Wang, H., Stoppacciaro, A., Stassi, G., Sargiacomo, M., Sinibaldi, P., Turchi, V., Giovannoni, R., Della Casa, G., Seren, E., Rossi, G. 2002. Efficient production by sperm-mediated gene transfer of human decay accelerating factor (hDAF) transgenic pigs for xenotransplantation. Proc Natl Acad Sci USA 99(22): 14230-14235.

Lavitrano, M., Busnelli, M., Cerrito, M.G., Giovannoni, R., Manzini, S., Vargiolu, A. 2006. Sperm-mediated gene transfer. Reprod Fertil Dev 18(1-2): 19-23.

Lee, G.S., Hyun, S.H., Kim, H.S., Kim, D.Y., Lee, S.H., Lim, J.M., Lee, E.S., Kang, S.K., Lee, B.C, Hwang, W.S. 2003. Improvement of a porcine somatic cell nuclear transfer technique by optimizing donor cell and recipient oocyte preparations. Theriogenology 59(9): 1949-1957.

Lee, H.G., Lee, H.C., Kim, S.W., Lee, P., Chung, H.J., Lee, Y.K., Han, J.H., Hwang, I.S., Yoo, J.I., Kim, Y.K., Kim, H.T., Lee, H.T., Chang, W.K., Park, J.K. 2009. Production of recombinant human von Willebrand factor in the milk of transgenic pigs. J Reprod Dev 55(5): 484-490.

Lee, S.L., Ock, S.A., Yoo, J.G., Kumar, B.M., Choe, S.Y., Rho, G.J. 2005. Efficiency of gene transfection into donor cells for nuclear transfer of bovine embryos. Mol Reprod Dev 72(2): 191-200.

Li, J., Svarcova, O., Villemoes, K., Kragh, P.M., Schmidt, M., Bøgh, I.B., Zhang, Y., Du, Y., Lin, L., Purup, S., Xue, Q., Bolund, L., Yang, H., Maddox-Hyttel, P., Vajta, G. 2008. High in vitro development after somatic cell nuclear transfer and trichostatin A treatment of reconstructed porcine embryos. Theriogenology 70(5): 800-808.

Li ,S., Guo, Y., Shi, J., Yin, C., Xing, F., Xu, L., Zhang, C., Liu, T., Li, Y., Li, H., Du, L., Chen, X. 2009. Transgene expression of enhanced green fluorescent protein in cloned rabbits generated from in vitro-transfected adult fibroblasts. Transgenic Res 18: 227-235.

Li. T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P., Yang, B. 2011. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res 39(1): 359-372.

Lo, D., Pursel, V., Linton, P.J., Sandgren, E., Behringer, R., Rexroad, C., Palmiter, R.D., Brinster, R.L. 1991. Expression of mouse IgA by transgenic mice, pigs and sheep. Eur. J. Immunol 21: 1001–1006.

Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., Plath, K. 2008. Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci USA 105(8): 2883-2888.

Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichtinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., Unger-Wallace, E., Sander, J.D., Müller-Lerch, F., Fu, F., Pearlberg, J., Göbel, C., Dassie, J.P., Pruett-Miller, S.M., Porteus, M.H., Sgroi, D.C., Iafrate, A.J., Dobbs, D., McCray, P.B. Jr., Cathomen, T., Voytas, D.F., Joung, J.K. 2008. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31(2): 294-301.

Magre, S., Takeuchi, Y., Bartosch, B. 2003. Xenotransplantation and pig endogenous retroviruses. Rev Med Virol 13(5): 311-329.

Mahfouz, M.M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X., Zhu, J.K. 2011. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci USA 108(6): 2623-2628.

Marques, M.M., Thomson, A.J., McCreath, K.J., McWhir, J. 2006. Conventional gene targeting protocols lead to loss of targeted cells when applied to a silent gene locus in primary fibroblasts. J Biotechnol 125(2): 185-193.

McCreath, K. J., Howcroft, J., Campbell, K. H., Colman, A., Schnieke, A. E., Kind, A. J. 2000. Production of gene-targeted sheep by nuclear transfer from cultured somatic cells. Nature 405: 1066-1069.

McKee, C., Gibson, A., Dalrymple, M., Emslie, L., Garner, I., Cottingham, I. 1998. Production of biologically active salmon calcitonin in the milk of transgenic rabbits. Nat Biotechnol16(7): 647-651.

Meehan, D.T., Zink, M.A., Mahlen, M., Nelson, M., Sanger, W.G., Mitalipov, S.M., Wolf, D.P., Ouellette, M.M., Norgren Jr, R.B. 2008. Gene targeting in adult rhesus macaque fibroblasts. BMC Biotechnol 8: 31.

Melo, E.O., Canavessi, A.M., Franco, M.M., Rumpf, R. 2007. Animal transgenesis: state of the art and applications. J Appl Genet 48(1): 47-61.

Meng, Q., Polgar, Z., Liu, J., Dinnyes, A. 2009. Live birth of somatic cell-cloned rabbits following trichostatin A treatment and cotransfer of parthenogenetic embryos. Cloning Stem Cells 11(1): 203-208.

Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., Dulay, G.P., Hua, K.L., Ankoudinova, I., Cost, G.J., Urnov, F.D., Zhang, H.S., Holmes, M.C., Zhang, L., Gregory, P.D., Rebar, E.J. 2011. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29(2): 143-148.

Mir, B., Piedrahita, J.A. 2004. Nuclear localization signal and cell synchrony enhance gene targeting efficiency in primary fetal fibroblasts. Nucleic Acids Res 32(3): e25.

Miranda Mdos, S., Bressan, F.F., Zecchin, K.G., Vercesi, A.E., Mesquita, L.G., Merighe, G.K., King, W.A., Ohashi, O.M., Pimentel, J.R., Perecin, F., Meirelles, F.V. 2009. Serum-starved apoptotic fibroblasts reduce blastocyst production but enable development to term after SCNT in cattle. Cloning Stem Cells 11(4): 565-573.

Miyoshi, K., Mori, H., Mizobe, Y., Asasaka, E., Ozawa, A., Yoshida, M., Sato, M. 2010. Valproic acid enhances in vitro development and oct-3/4 expression of miniature pig somatic cell nuclear transfer embryos. Cell Reprog 12: 67-74.

Moehle, E.A., Rock, J.M., Lee, Y.L., Jouvenot, Y., DeKelver, R.C., Gregory, P.D., Urnov, F.D., Holmes, M.C. 2007. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci USA 104(9): 3055-3060.

Moens, C. B., Auerbach, A. B., Conlon, R. A., Joyner, A. L., Rossant, J. 1992. A targeted mutation reveals a role for N-myc in branching morphogenesis in the embryonic mouse lung. Genes Dev 6: 691-704.

Moisvadi, S., Kaminski, J.M., Yanagimachi, R. 2009. Use of intracytoplasmic sperm injection (ICSI) to generate transgenic animals. Comp Immunol Microbiol Infect Dis 32(2): 47-60.

Monaco, M.H., Gronlund, D.E., Bleck, G.T., Hurley, W.L., Wheeler, M.B., Donovan, S.M. 2005. Mammary specific transgenic over-expression of insulin-like growth factor-I (IGF-I) increases pig milk IGF-I and IGF binding proteins, with no effect on milk composition or yield. Transgenic Res 14(5): 761-773.

Morbitzer, R., Römer, P., Boch, J., Lahaye, T. 2010. Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. Proc Natl Acad Sci USA 107(50): 21617-21622.

Moscou, M.J., Bogdanove, A.J. 2009. A simple cipher governs DNA recognition by TAL effectors. Science 326(5959): 1501.

Mueller, N.J., Takeuchi, Y., Mattiuzzo, G., Scobie, L. 2011. Microbial safety in xenotransplantation. Curr Opin Organ Transplant 16(2): 201-206.

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., Yamanaka, S. 2008. Generation of mouse induced pluripotent stem cells without viral vectors. Science 322(5903): 949-953.

Oropeza, M., Petersen, B., Carnwath, J.W., Lucas-Hahn, A., Lemme, E., Hassel, P., Herrmann, D., Barg-Kues, B., Holler, S., Queisser, A-L., Schwinzer, R., Hinkel, R., Kupatt, C., Niemann, H. 2009. Transgenic expression of the human A20 gene in cloned pigs provides protection against apoptotic and inflammatory stimuli. Xenotransplantation 16: 522-534.

Osada, T., Toyoda, A., Moisyadi, S., Akutsu, H., Hattori, M., Sakaki, Y., Yanagimachi, R. 2005. Production of inbred and hybrid transgenic mice carrying large (>200 kb) foreign DNA fragments by intracytoplasmic sperm injection, Mol Reprod Dev 72: 329-335.

Paleyanda, R.K., Velander, W.H., Lee, T.K., Scandella, D.H., Gwazdauskas, F.C., Knight, J.W., Hoyer, L.W., Drohan, W.N., Lubon, H. 1997. Transgenic pigs produce functional human factor VIII in milk. Nat Biotechnol 15(10): 971-975.

Paradis, K., Langford, G., Long, Z., Heneine, W., Sandstrom, P., Switzer, W.M., Chapman, L.E., Lockey, C., Onions, D., Otto, E. 1999. Search for cross-species transmission of porcine endogenous retrovirus in patients treated with living pig tissue. Science 285(5431): 1236-1241.

Park, J.K., Lee, Y.K., Lee, P., Chung, H.J., Kim, S., Lee, H.G., Seo, M.K., Han, J.H., Park, C.G., Kim, H.T., Kim, Y.K., Min, K.S., Kim, J.H., Lee, H.T., Chang, W.K. 2006. Recombinant human erythropoietin produced in milk of transgenic pigs. J Biotechnol 122(3): 362-371.

Patience, C., Takeuchi, Y., Weiss, R.A. 1997. Infection of human cells by an endogenous retrovirus of pigs. Nat Med 3(3): 282-286.

Perry, A.C.F., Wakayama, T., Kishikawa, H., Kasai, T., Okabe, M., Toyoda, Y., Yanagimachi, R. 1999. Mammalian transgenesis by intracytoplasmic sperm injection. Science 284:1180-1183.

Petersen, B., Kues, W., Lucas-Hahn, A., Queisser, A.L., Lemme, E., Hoelker, M., Carnwath J.W., Niemann, H. 2006. Generation of pigs transgenic for hCD59/DAF and human thrombomodulin by somatic nuclear transfer. Reprod Fertil Dev 18: 142.

Pfister-Genskow, M., Myers, C., Childs, L.A., Lacson, J.C., Patterson, T., Betthauser, J.M., Goueleke, P.J., Koppang, R.W., Lange, G., Fisher, P., Watt, S.R., Forsberg, E.J., Zheng, Y., Leno, G.H., Schultz, R.M., Liu, B., Chetia, C., Yang, X., Hoeschele, I., Eilertsen, K.J. 2005. Identification of differentially expressed genes in individual bovine preimplantation embryos produced by nuclear transfer: improper reprogramming of genes required for development. Biol Reprod 72(3): 546-555.

Phelps, C.J., Koike, C., Vaught, T.D., Boone, J., Wells, K.D., Chen, S.H., Ball, S., Specht, S.M., Polejaeva, I.A., Monahan, J.A., Jobst, P.M., Sharma, S.B., Lamborn, A.E., Garst, A.S., Moore, M., Demetris, A.J., Rudert, W.A., Bottino, R., Bertera, S., Trucco, M., Starzl, T.E., Dai, Y., Ayares, D.L. 2003. Production of alpha 1,3-galactosyltransferase-deficient pigs. Science 299(5605): 411-414.

Poggiali, P., Scoarughi, G.L., Lavitrano, M., Donini, P., Cimmino, C. 2002. Construction of a swine artificial chromosome: a novel vector for transgenesis in the pig. Biochimie 84(11): 1143-1150.

Prather, R.S., Shen, M., Dai, Y. 2008. Genetically modified pigs for medicine and agriculture. Biotechnology and Genetic Engineering Reviews 25: 245-266.

Pursel, V.G., Rexroad, C.E. Jr. 1993. Recent progress in the transgenic modification of swine and sheep. Mol Reprod Dev 36(2): 251-254.

Pursel, V.G., Wall, R.J., Mitchell, A.D., Elsasser, T.H., Solomon, M.B., Coleman, M.E., Mayo, F., Schwartz, R. J. 1999. Expression of insulin-like growth factor-1 in skeletal muscle of transgenic pigs. In 'Transgenic Animals in Agriculture'. (Eds J. D. Murray, G. B. Anderson, A. M. Oberbauer and M. M. McGloughlin.) pp. 131-144. (CABI Publishing: New York.)

Qin, J.Y., Zhang, L., Clift, K.L., Hulur, I., Xiang, A.P., Ren, B.Z., Lahn, B.T. 2010. Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter. PLoS ONE 5(5): e10611.

Ramirez, C.L., Foley, J.E., Wright, D.A., Müller-Lerch, F., Rahman, S.H., Cornu, T.I., Winfrey, R.J., Sander, J.D., Fu, F., Townsend, J.A., Cathomen, T., Voytas, D.F., Joung, J.K. 2008. Unexpected failure rates for modular assembly of engineered zinc fingers. Nat Methods 5(5): 374-375.

Ramsoondar, J.J., Machaty, Z., Costa, C., Williams, B.L., Fodor, W.L., Bondioli, K.R. 2003. Production of alpha 1, 3-galactosyltransferase-knockout cloned pigs expressing human alpha 1, 2-fucosylosyltransferase. Biol Reprod 69(2): 437-445.

Ramsoondar, J., Vaught, T., Ball, S., Mendicino, M., Monahan, J., Jobst, P., Vance, A., Duncan, J., Wells, K., Ayares, D. 2009. Production of transgenic pigs that express porcine endogenous retrovirus small interfering RNAs. Xenotransplantation 16(3): 164-180.

Rathbone, A.J., Fisher, P.A., Lee, J.H., Craigon, J., Campbell, K.H. 2010. Reprogramming of ovine somatic cells with Xenopus laevis oocyte extract prior to SCNT improves live birth rate. Cell Reprogram. 12(5): 609-616.

Reh, W.A., Maga, E.A., Collette, N.M., Moyer, A., Conrad-Brink, J.S., Taylor, S.J., DePeters, E.J., Oppenheim, S., Rowe, J.D., BonDurant, R.H., Anderson, G.B., Murray, J.D. 2004. Hot topic: using a stearoyl-CoA desaturase transgene to alter milk fatty acid composition. J Dairy Sci 87: 3510-3514.

Richt, J.A., Kasinathan, P., Hamir, A.N., Castilla, J., Sathiyaseelan, T., Vargas, F., Sathiyaseelan, J., Wu, H., Matsushita, H., Koster, J., Kato, S., Ishida, I., Soto, C., Robl, J.M., Kuroiwa, Y. 2007. Production of cattle lacking prion protein. Nat Biotechnol 25(1): 132-138.

Robl, J.M., Wang, Z., Kasinathan, P., Kuroiwa, Y. 2007. Transgenic animal production and animal biotechnology. Theriogenology 67(1): 127-133.

Rogers, C.S., Hao, Y., Rokhlina, T., Samuel, M., Stoltz, D.A., Li, Y., Petroff, E., Vermeer, D.W., Kabel, A.C., Yan, Z., Spate, L., Wax, D., Murphy, C.N., Rieke, A., Whitworth, K., Linville, M.L., Korte, S.W., Schnieke, A.E., Kind, A.J., Ritchie, W.A., Mycock, K., Scott, A.R., Ritchie, M., Wilmut, I., Colman, A., Campbell, K.H. 1997. Human factor

IX transgenic sheep produced by transfer of nuclei from transfected fetal fibroblasts. Science 278: 2130-2133.

Rogers, C.S., Hao, Y., Roklina, T., Samuel, M., Stoltz, D.A., Li, Y., Petroff, E., Vermeer, D.W., Kabel, A.C., Yan, Z., Spate, L., Wax, D., Murphy, C.N., Rieke, A., Whitworth, K., Linville, M.L., Korte, S.W., Engelhardt, J.F., Welsh, M.J., Prather, R.S. 2008. Production of CFTR-null and CFTR-deltaF508 heterozygous pigs by adeno-associated virus-mediated gene targeting and somatic cell nuclear transfer. J Clin Invest 118(4): 1571-1577.

Ross, J.W., Prather, R.S., Whyte, J.J., Zhao, J. 2009a. Cloning and transgenics: progress and new approaches in domestic animals. CAB reviews: perspectives in agriculture, veterinary science. Nutr Nat Resour 4(38): 13.

Ross, J.W., Zhao, J., Walters, E.M., Samuel, M., Narfstrom, K., Jeong, M., DeMarco, M.A., McCall, M.A., Kaplan, H.J., Prather, R.S. 2009b. Somatic cell nuclear transfer to create a miniature swine model of retinitis pigmentosa, April 18-22, New Orleans, LA

Ross, J.W., Whyte, J.J., Zhao, J., Samuel, M., Wells, K.D., Prather, R.S. 2010a. Optimization of square-wave electroporation for transfection of porcine fetal fibroblasts. Transgenic Res 19(4): 611-620.

Ross, P.J., Wang, K., Kocabas, A., Cibelli, J.B. 2010b. Housekeeping gene transcript abundance in bovine fertilized and cloned embryos. Cell Reprogram 12(6): 709-717.

Saeki, K., Matsumoto, K., Kinoshita, M., Suzuki, I., Tasaka, Y., Kano, K., Tagachui, Y., Mikami, K., Hirabayashi, M., Kashiwazaki, N., Hosoi, Y., Murata, N., Iritani, A. 2004. Functional expression of a Delta12 fatty acid desaturase gene from spinach in transgenic pigs. Proc. Natl. Acad. Sci. USA 101: 6361-6366.

Salamone, D., Barañao, L., Santos, C., Bussmann, L., Artuso, J., Werning, C., Prync, A., Carbonetto, C., Dabsys, S., Munar, C., Salaberry, R., Berra, G., Berra, I., Fernández, N., Papouchado, M., Foti, M., Judewicz, N., Mujica, I., Muñoz, L., Alvarez, S.F., González, E., Zimmermann, J., Criscuolo, M., Melo, C. 2006. High level expression of bioactive recombinant human growth hormone in the milk of a cloned transgenic cow. J Biotechnol. 124(2): 469-472.

Santos, F., Zakhartchenko, V., Stojkovic, M., Peters, A., Jenuwein, T., Wolf, E., Reik, W., Dean, W. 2003. Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. Curr Biol 13: 1116-1121.

Schnieke, A.E., Kind, A.J., Ritchie, W.A., Mycock, K., Scott, A.R., Ritchie, M., Wilmut, I., Colman, A., Campbell, K.H. 1997. Human factor IX transgenic sheep produced by transfer of nuclei from transfected fetal fibroblasts. Science 278(5346): 2130-2133.

Shi, L.H., Ai, J.S., Ouyang, Y.C., Huang, J.C., Lei, Z.L., Wang, Q., Yin, S., Han, Z.M., Sun, Q.Y., Chen, D.Y. 2008. Trichostatin A and nuclear reprogramming of cloned rabbit embryos. J Anim Sci 86(5): 1106-1113.

Sun, X., Yan, Z., Yi, Y., Li, Z., Lei, D., Rogers, C.S., Chen, J., Zhang, Y., Welsh, M.J., Leno, G.H., Engelhardt, J.F. 2008. Adeno-associated virus-targeted disruption of the CFTR gene in cloned ferrets. J Clin Invest 118(4): 1578-1583.

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G., Hochedlinger, K. 2008. Induced pluripotent stem cells generated without viral integration. Science 322(5903): 945-949.

Switzer, W.M., Michler, R.E., Shangmugam, V., Matthews, A., Hussain, A.I., Wright A, Sandstrom, P., Chapman, L.E., Weber, C., Safley, S., Denny, R.R., Navarro, A., Evans, V., Norin, A.J., Kwiatkowski, P., Heneine, W. 2001. Lack of cross-species transmission of porcine endogenous retrovirus infection to nonhuman primate recipients of porcine cells, tissues and organs. Transplantation 71: 959-965.

Takahashi, K., Yamanaka, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126 (4): 663-676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131(5): 861-872.

te Riele, H., Maandag, E.R., Berns, A. 1992. Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. Proc Natl Acad Sci USA. 89(11): 5128-5132.

Tsuji, Y., Kato, Y., Tsunoda, Y. 2009. The developmental potential of mouse somatic cell nuclear-transferred oocytes treated with trichostatin A and 5-aza-2'-deoxycytidine. Zygote 17(2): 109-115.

Urnov, F.D., Miller, J.C., Lee, Y.L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., Holmes, M.C. 2005. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435(7042): 646-651.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., Gregory, P.D. 2010. Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11(9): 636-646.

van Berkel, P.H., Welling, M.M., Geerts, M., van Veen, H.A., Ravensbergen, B., Salaheddine, M., Pauwels, E.K., Pieper, F., Nuijens, J.H., Nibbering, P.H. 2002. Large scale production of recombinant human lactoferrin in the milk of transgenic cows. Nat Biotechnol 20(5): 484-487.

van Poll, D., Nahmias, Y., Soto-Gutierrez, A., Ghasemi, M., Yagi, H., Kobayashi, N., Yarmush, M.L., Hertl, M. 2010. Human immune reactivity against liver sinusoidal endothelial cells from GalTα(1,3)GalT-deficient pigs. Cell Transplant 19(6): 783-789.

Vasquez, K.M., Marburger, K., Intody, Z., Wilson, J.H. 2001. Manipulating the mammalian genome by homologous recombination. Proc Natl Acad Sci USA. 98(15): 8403-8410.

Voigt, B., Serikawa, T. 2009. Pluripotent stem cells and other technologies will eventually open the door for straightforward gene targeting in the rat. Dis Model Mech 2(7-8): 341-343.

Wall, R.J., Powell, A., Paape, M.J., Kerr, D.E., Bannermann, D.D., Pursel, V.G., Wells, K.D., Talbot, N., Hawk, H. 2005. Genetically enhanced cows resist intramammary *Staphylococcus aureus* infection. Nat Biotechnol 23: 445-451.

Wang, J., Yang, P., Tang, B., Sun, X., Zhang, R., Guo, C., Gong, G., Liu, Y., Li, R., Zhang, L., Dai, Y., Li, N. 2008. Expression and characterization of bioactive recombinant human α-lactalbumin in the milk of transgenic cloned cows. J Dairy Sci 91(12): 4466-4476.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A., Daley, G.Q., Brack, A.S., Collins, J.J., Cowan, C., Schlaeger, T.M., Rossi, D.J. 2010. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell 7(5): 618-630.

Watanabe, S., Iwamoto, M., Suzuki, S., Fuchimoto, D., Honma, D., Nagai, T., Hashimoto, M., Yazaki, S., Sato, M., Onishi, A. 2005. A novel method for the production of transgenic cloned pigs: electroporation-mediated gene transfer to non-cultured cells and subsequent selection with puromycin. Biol Reprod 72(2): 309-315.

Watanabe, M., Umeyama, K, Matsunari, H., Takayanagi, S., Haruyama, E., Nakano, K., Fujiwara, T., Ikezawa, Y., Nakauchi, H., Nagashima, H. 2010. Knockout of exogenous EGFP gene in porcine somatic cells using zinc-finger nucleases. Bio chem Biophys Res Commun 402(1): 14-18.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., Jaenisch, R. 2007. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 448(7151): 318-324.

Wheeler, M.B., Bleck, G.T., Donovan, S.M. 2001. Transgenic alteration of sow milk to improve piglet growth and health. Reproduction 58 (Suppl.): 313-324.

Whitelaw, C.B., Lillico, S.G., King, T. 2008. Production of transgenic farm animals by viral vector-mediated gene transfer. Reprod Domest Anim 43 Suppl 2: 355-358.

Whitworth, K.M., Prather, R.S. 2010. Somatic cell nuclear transfer efficiency: how can it be improved through nuclear remodeling and reprogramming? Mol Reprod Dev 77(12): 1001-1015.

Whyte, J.J., Zhao, J., Wells, K.D., Samuel, M.S., Whitworth, K.M., Walters, E.M., Laughlin, M.H., Prather, R.S. 2011. Gene targeting with zinc finger nucleases to produce cloned eGFP knockout pigs. Mol Reprod Dev 78(1): 2.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., Campbell, K.H. 1997. Viable offspring derived from fetal and adult mammalian cells. Nature 385(6619): 810-813.

Wrenzycki, C., Wells, D., Herrmann, D., Miller, A., Oliver, J., Tervit, R., Niemann, H. 2001. Nuclear transfer protocol affects messenger RNA expression patterns in cloned bovine blastocysts. Biol Reprod 65(1): 309-317.

Wrenzycki, C., Herrmann, D., Gebert, C., Carnwath, J.W., Niemann, H. 2006. Gene expression and methylation patterns in cloned embryos. Methods Mol Biol 348: 285-304.

Wright, G., Carver, A., Cottom, D., Reeves, D., Scott, A., Simons, P., Wilmut, I., Garner, I., Colman, A. 1991. High level expression of active human alpha-1-antitrypsin in the milk of transgenic sheep. Biotechnology (N Y) 9(9): 830-834.

Wu, S., Ying, G.X., Wu, Q., Capecchi, M.R. 2008. A protocol for constructing gene targeting vectors: generating knockout mice for the cadherin family and beyond. Nature Protocols 3: 1056-1076.

Wu, Z., Chen, J., Ren, J., Bao, L., Liao, J., Cui, C., Rao, L., Li, H., Gu, Y., Dai, H., Zhu, H., Teng, X., Cheng, L., Xiao, L. 2009. Generation of pig induced pluripotent stem cells with a drug inducible system. J Mol Cell Biol 1(1): 46-54.

Yamada, K., Yazawa, K., Shimizu, A., Iwanaga, T., Hisashi, Y., Nuhn, M., O'Malley, P., Nobori, S., Vagefi, P.A., Patience, C., Fishman, J., Cooper, D.K., Hawley, R.J., Greenstein, J., Schuurman, H.J., Awwad, M., Sykes, M., Sachs, D.H. 2005. Marked prolongation of porcine renal xenograft survival in baboons through the use of α1,3-galactosyltransferase gene-knockout donors and the cotransplantation of vascularized thymic tissue. Nat Med 11: 32-34.

Yáñez, R.J., Porter, A.C. 1999. Influence of DNA delivery method on gene targeting frequencies in human cells. Somat Cell Mol Genet 25(1): 27-31.

Yang, Y., Seed, B. 2003. Site-specific gene targeting in mouse embryonic stem cells with intact bacterial artificial chromosomes. Nat Biotechnol 21(4): 447-451.

Yang, P., Wang, J., Gong, G., Sun, X., Zhang, R., Du, Z., Liu, Y., Li, R., Ding, F., Tang, B., Dai, Y., Li, N. 2008. Cattle mammary bioreactor generated by a novel procedure of transgenic cloning for large-scale production of functional human lactoferrin. PloS One 3(10): e3453.

Yang, D., Wang, C.E., Zhao, B., Li, W., Ouyang, Z., Liu, Z., Yang, H., Fan, P., O'Neill, A., Gu, W., Yi, H., Li, S., Lai, L., Li, X.J. 2010. Expression of Huntington's disease protein results in apoptotic neurons in the brains of cloned transgenic pigs. Hum Mol Genet 19(20): 3983-3994.

Yu, G., Chen, J., Yu, H., Liu, S., Chen, J., Xu, X., Sha, H., Zhang, X., Wu, G., Xu, S., Cheng, G. 2006. Functional disruption of the prion protein gene in cloned goats. J Gen Virol 87: 1019-1027.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., Slukvin, I.I., Thomson, J.A. 2007. Induced pluripotent stem cell lines derived from human somatic cells. Science 318(5858): 1917-1920.

Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I., Thomson, J.A. 2009. Human induced pluripotent stem cells free of vector and transgene sequences. Science 324(5928): 797-801.

Zaidi, A., Schmoeckel, M., Bhatti, F., Waterworth, P., Tolan, M., Cozzi, E., Chavez, G., Langford, G., Thiru, S., Wallwork, J., White, D., Friend, P. 1998. Life-supporting pig-to primate renal xenotransplantation using genetically modified donors. Transplantation 65: 1584.

Zakhartchenko, V., Mueller, S., Alberio, R., Schernthaner, W., Stojkovic, M., Wenigerkind, H., Wanke, R., Lassnig, C., Mueller, M., Wolf, E., Brem, G. 2001. Nuclear transfer in cattle with non-transfected and transfected fetal or cloned transgenic fetal and postnatal fibroblasts. Mol Reprod Dev 60: 362-369.

Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M., Arlotta, P. 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29(2): 149-153.

Zhao, J., Ross, J.W., Hao, Y., Spate, L.D., Walters, E.M., Samuel, M.S., Rieke, A., Murphy, C.N., Prather, R.S. 2009a. Significant improvement in cloning efficiency of an inbred miniature pig by histone deacetylase inhibitor treatment after somatic cell nuclear transfer. Biol Reprod 81(3): 525-530.

Zhao, J., Hao, Y., Ross, J.W., Spate, L.D., Walters, E.M., Samuel, M.S., Rieke, A., Murphy, C.N., Prather, R.S. 2010. Histone deacetylase inhibitors improve in vitro and in vivo developmental competence of somatic cell nuclear transfer porcine embryos. Cell Reprogram 12(1): 75-83.

Zhao, X. Y., Li, W., Lv, Z., Liu, L., Tong, M., Hai, T., Hao, J., Guo, C.L., Ma, Q.W., Wang, L., Zeng, F., Zhou, Q. 2009b. iPS cells produce viable mice through tetraploid complementation. Nature 461(7260): 86-90.

Zhu, C., Li, B., Yu, G., Chen, J., Yu, H., Chen, J., Xu, X., Wu, Y., Zhang, A., Cheng, G. 2009. Production of Prnp -/- goats by gene targeting in adult fibroblasts. Transgenic Res 18(2):163-171.

30

Animal Models of Angiogenesis and Lymphangiogenesis

L. D. Jensen[1,2] et al.*
[1]*Department of Microbiology, Tumor and Cell Biology,*
The Karolinska Institute, Stockholm,
[2]*Institution of Medicine and Health, Linköping University, Linköping,*
Sweden

1. Introduction

Blood and lymphatic vessels are present in all tissues, and play important roles for their function, homeostasis and maintenance. Angiogenesis, the growth of new blood vessels, is therefore highly important during development, but is largely not observed in the adult, except for during the female reproduction cycle and during wound healing. In pathological situations, however, angiogenesis may be turned on, and in this case contribute to the onset and progression of most severe human pathologies characterized by high mortality, including cancer, diabetes, obesity and retinopathies (Carmeliet, 2003) or is insufficiently activated such as in the case of myocardial infarction and stroke (Y Cao et al, 2005). Thus, angiogenesis is one of the largest and fastest evolving areas of research today. Angiogenesis is a highly complicated process, involving many different cell types, and it is therefore highly recommended that researchers use *in vivo* animal models for their studies. Accordingly, today there are many *in vivo* models available.

The aim of this chapter is to give insights into the most commonly used *in vivo* angiogenesis models in both mice and zebrafish. We will provide detailed descriptions and discussions of the adipose tissue-, tumor-, ischemic hind limb- wound healing- and corneal micropocket angiogenesis models in mice and developmental-, tumor-, hypoxia-induced retinal- and regenerating tail fin angiogenesis models in zebrafish. We will provide a base for comparison between the different assays to quickly identify which model is best suited for a particular research focus.

1.1 Basic mechanisms of angiogenesis

Angiogenesis is a multistep process which is tightly regulated by an intimate balance between pro- and anti-angiogenic factors. Angiogenesis is stimulated by angiogenic factors the most commonly studied being members of the vascular endothelial growth factor (VEGF), fibroblast growth factor (FGF), transforming growth factor (TGF) or platelet derived growth factor (PDGF) families in the tissue. These factors either act locally, or are released into the blood stream and activate angiogenesis through binding to their corresponding receptors on pre-existing endothelial cells and/or perivascular cells (Y Cao, 2008). Upon ligand binding to the receptors, a cascade of intracellular signal transduction is triggered resulting in activation of the endothelial cell and initiation of the cellular mechanisms required for angiogenesis including basal membrane degradation, endothelial cell (EC) proliferation and migration, formation of tube like structures, and finally vascular maturation by coverage with smooth muscle cells or pericytes to ensure stability (Davis & Senger, 2005; Jensen et al, 2007).

---

* J. Honek[1], K. Hosaka[1], P. Rouhi[1], S. Lim[1], H. Ji[1], Z. Cao[2], E. M. Hedlund[1], J. Zhang[1] and Y. Cao[1,2]
1 Department of Microbiology, Tumor and Cell Biology, Karolinska Institute, Stockholm, Sweden,
2 Institution of Medicine and Health, Linköping University, Linköping, Sweden.

The dynamics of angiogenesis as well as the morphology of the resulting vasculature differs depending on which factor induces the process (R Cao, 2004). In the case of the most studied angiogenic factor, VEGF, a mature vessel may be exposed to a gradient of VEGF induced by hypoxia in a nearby tissue such as a tumor. VEGF activates VEGF receptor 2 (VEGFR2) on endothelial cells leading to production of matrix-metalloproteases (MMPs) which disrupts EC-pericyte contacts, degrade the basal membrane and induce a tip-cell behavior in one cell, which is allowed to form cell processes important for cell migration such as filopodia and lammelopodia. The tip cell will via lateral inhibition signal to surrounding cells that they should not adopt this phenotype thus leading to an ordered sprouting of just one or a few neovessels from the original mother-vessel. The tip cell will start migrating up the angiogenic factor gradient, but remain in contact with the underlying and proliferating endothelial cells (stalk cells) and thus retain the connection to the original vessel. At a certain point, the tip cell will meet other tip cells, anastomose – i.e. fuse – and thus form a circulation loop. The endothelial chords will lumenize, allowing blood to flow through the vessel and the vessels will mature by recruiting perivascular mural cells which are tightly associated with the endothelial cells thus providing stability and trophic factors for the endothelial cells, leading to their maturation and re-entry into quiescence (Risau 1997).

1.2 Angiogenesis in physiology and pathology

During development the first vessels – the aorta and the cardinal vein – are formed by vasculogenesis - i.e. the *de novo* formation of blood vessels by aggregation and vascular morphogenesis of single endothelial progenitor cells. Following this process, the majority of the blood vessels in the body are formed by angiogenic expansion of this initial, primitive circulation loop (Risau 1997). Thus angiogenesis is very important during development. In adults however, most tissues and therefore also the vasculature have stopped growing and instead adopted a quiescent state. In some cases such as during the female reproductive cycle and during wound healing, angiogenesis is needed to regenerate the tissue, but otherwise the erroneous or insufficient induction of angiogenesis in adults is usually associated with pathology. For example, as a tumor is a growing tissue and therefore need blood vessels for supplying energy, angiogenesis is induced after the tumor has reached a certain size (approximately 1 cubic centimetre). Similarly in growing fat tissue new blood vessels are needed when the distance from the existing blood vessels are too large for efficient oxygenation. By blocking such tumor- or adipose-angiogenesis, it is possible to inhibit the growth of these tissues and therefore anti-angiogenic drugs have potential as treatment for cancer and obesity. In patients with retinopathy retinal hypoxia is a consequence of blocked or disrupted blood vessels and also here lead to induction of retinal angiogenesis. The newly formed blood vessels are however often immature and prone to rupture which leads to micro-hemorrhages and thereby drives the pathological progression to advanced states of the disease (Carmeliet, 2003; Folkman, 1995). On the other hand, in other hypoxic tissues such as the hypoxic heart tissue downstream of a blocked coronary artery, it is important to speed up the angiogenic revascularization of the tissue in order to secure sufficient oxygen for the highly metabolically active myocardium such that it may sustain its function (Y Cao, 2010).

These aspects of physiological and pathological angiogenesis are accurately modeled by the angiogenesis models we will discuss in this chapter.

1.3 Mouse models of angiogenesis

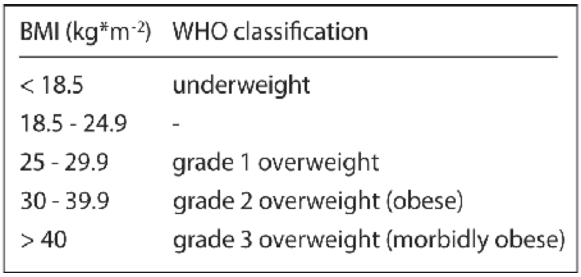
Mouse models are the primary *in vivo* tools used in biomedical research. Most research institutions have their own mouse or rodent facilities, and today several strains have been genetically purified through serial in-breeding, which allow sophisticated investigations such as transplantation and grafting experiments to be done in a wildtype animal with a fully functional immune system. Furthermore, genetically manipulated mice have been generated for most of the genes important for angiogenesis and vascular biology. Such mice include conditional or global knock-outs or mice over-expressing angiogenic factors in defined tissues. As the murine and human vasculatures are highly similar, these tools have been very valuable in defining the role of angiogenesis and distinct angiogenic factors in the onset and progression of various human diseases.

Here we present five mouse assays of pathological angiogenesis that have shown tremendous power in the study of mechanisms behind human diseases. These are the adipose-angiogenesis models used to study obesity and diabetes, the xenograft models used to study cancer biology, the ischemia models used to study myocardial infarction and stroke, the wound healing models and the cornea micropocket models used to study basic mechanisms of angiogenesis and lymphangiogenesis.

2. Angiogenesis in adipose tissue

2.1 Introduction

According to a report of the World Health Organization in 2009, there are currently more than one billion people overweight and more than 300 million individuals considered clinically obese. The escalating number of obese individuals is no longer a problem faced only in high income countries. This adverse trend has also been adopted by low and middle income countries. Body Mass Index (BMI) is a commonly used method to categorize overweight ($BMI \geq 25\ kg*m^{-2}$) and obese ($BMI > 30\ kg*m^{-2}$) individuals (Table 1) (Kopelman, 2000). However, rather than just being a cosmetic problem or being associated with social issues, obesity is a serious pre-disease that frequently leads to the development of severe complications and metabolic disorders that often have a fatal outcome. Amongst these diseases are dyslipidemias, fatty liver, sleep apnea, cardiovascular complications, stroke, type 2 diabetes as well as certain types of cancer such as prostate, breast and colon cancer (Y Cao, 2010; Y Cao, 2007). In the affected individuals, development of metabolic disorders is frequently related to endothelial dysfunction.



BMI ($kg*m^{-2}$)	WHO classification
< 18.5	underweight
18.5 - 24.9	-
25 - 29.9	grade 1 overweight
30 - 39.9	grade 2 overweight (obese)
> 40	grade 3 overweight (morbidly obese)

Table 1. Association of Body Mass Index (BMI) with overweight and obesity. A BMI value greater than 30 indicates obesity. Adapted from Kopelman, 2000.

The adipose tissue constantly experiences expansion and regression during growth and repair throughout adulthood. Interestingly, it is known that adipose tissue growth relies on angiogenesis. Already in embryos, the formation of the primitive fat organ is preceded by angiogenesis. Moreover, anti-angiogenic therapy can prevent the expansion and even induce regression of adipose tissue (Brakenhielm et al, 2004).

The adipose tissue vasculature plays an important role in supplying oxygen and nutrients, as well as plasma containing cytokines and growth factors. Furthermore, the vasculature provides the adipose tissue with bone marrow-derived stem cells, able to differentiate into pre-adipocytes and adipocytes, endothelial cells as well as pericytes. The blood vessels also facilitate infiltration of inflammatory cells, such as monocytes and neutrophiles, and play a role in the removal of metabolic waste products. Due to the close interaction of blood vessels and adipocytes, studying angiogenesis in the adipose tissue is a promising strategy to identify potential novel targets for anti-obesity/anti-diabetic therapy and might open new avenues in the prevention and treatment of metabolic disorders in the future.

There are two types of adipose tissue in the human adult, the white adipose tissue (WAT) and the brown adipose tissue (BAT). WAT is frequently regarded as the ‘bad’ fat which stores excess energy in the form of triglycerides. Indeed, this notorious WAT usually amasses in undesirable parts of the body such as in the intra-abdominal area resulting in a so called ‘apple-shaped’ body type or the thighs or hips leading to the so called ‘pear-shaped’ body type. WAT is essential to serve as heat insulation, mechanical cushion and to provide energy for the body. WATs are complex tissues consisting of different cell types such as endothelial cells, pericytes, macrophages and mesenchymal cells having close interaction with one another and collaboratively regulating processes in the adipose tissue. Being an endocrine organ, the WAT produces a myriad of cytokines and angiogenic factors including vascular endothelial growth factor (VEGF), leptin, adiponectin, resistin, interleukin-6 (IL-6), IL-8, hepatocyte growth factor (HGF), angiopoietin (Ang)-1 and -2 FGF-2, estrogen, TGF-α and -β as well as MMP-2 and -9 (Brakenhielm & Y Cao, 2008). These factors interact to regulate the survival, proliferation and differentiation of pre-adipocytes to adipocytes.

Remarkably, adipose tissue is one of the most highly vascularized tissues in the human body (Fig. 1 and 5). In BAT, blood vessel density is several folds higher than in WAT. This reflects the higher metabolic activity of BAT. Adipocytes in WAT are characterized by a large diameter, a spherical morphology and a large unilocular lipid droplet which is surrounded by a thin layer of cytoplasm whereas adipocytes originating from BAT are smaller, contain multilocular lipid droplets and higher cytoplasm content. Furthermore, brown adipocytes express uncoupling protein 1 (UCP1) which has an important role in energy metabolism.

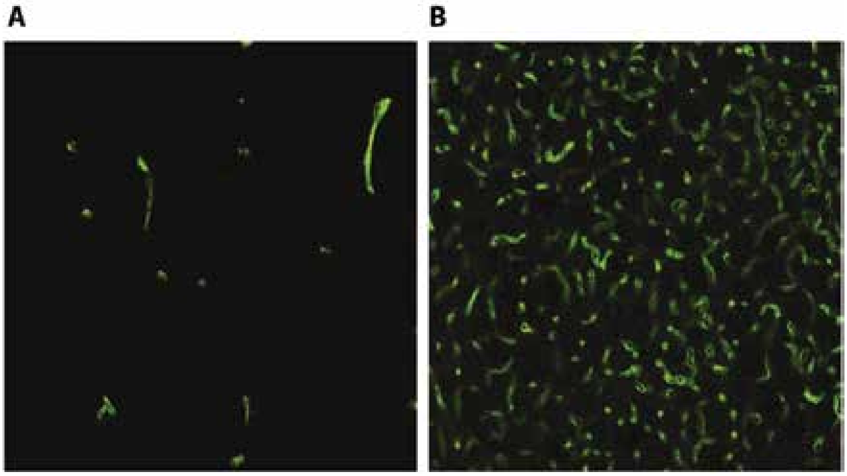


Fig. 1. Visualization of blood vessels in adipose tissue for quantitative analysis. Fluorescence-based visualization of CD31+ endothelial cells (green) in white (A) and brown (B) adipose tissue.

Many decades ago, BAT was discovered as a thermogenic (heat generating) tissue, active in new born human babies, but it was thought to disappear or rather to be inactivated in adults (Cannon & Nedergaard, 2004). BAT is densely packed with mitochondria which explains the high metabolic status compared to white adipose tissue. In rodents, BAT is primarily located in the interscapular region, on the dorsal side between the front limbs, with minor amounts being found in the thymus, thorax and abdomen. The highly metabolically active, thermogenic BAT requires a high density of blood vessels to supply oxygen and substrates to the mitochondria and for waste removal.

In recent years, there has been accumulating evidence demonstrating the presence of active BAT in adult humans (Cypess et al, 2009; van Marken Lichtenbelt et al, 2009; Virtanen et al, 2009). Several clinical observations have shown the presence of BAT in patients, with tumors such as pheochromocytoma, following exposure to high levels of catecholamines or exposure to cold. Most research utilizes the uptake of 18F-fluoro-2-deoxygucose as a tracer in positron emission tomography (PET) and computer tomography (CT) to detect active BAT depots in adult humans. PET-CT reveals that the distribution of the BAT depot is located in the fascial plane in the ventral neck and thorax bilaterally instead of the interscapular region as seen in rodents and children. The human adult possesses approximately 10 g of BAT. If all the brown fat in the adult body was fully activated, it would be able to burn around 4.1 kg of white fat in a year. Histological studies of human BAT depots show high capillary density. Here, we address the possibilities of driving the activity of BAT, of conversion of WAT to brown-like adipose tissue and of using angiogenesis modulators to treat obesity.

2.2 Models/methods to study adipose tissue angiogenesis

To study adipose tissue angiogenesis, several mouse models including models in genetically manipulated mice, are currently available (Xue Y et al, 2010, Nat. Protoc). These models usually provide highly reproducible and robust results as the mice are inbred and therefore share a highly similar genetic background. However, in humans, the cause of developing obesity is most frequently not genetic but rather due to overeating and a lifestyle based on high caloric intake and little physical exercise. Therefore, high-fat diet fed mouse models provide a powerful tool to study non-genetically related obesity.

2.3 Genetic models

2.3.1 Ob/ob mice

In 1950, obese mice carrying the mutation obese (ob) were described for the first time (Ingalls et al, 1950). The ob mutation was later shown to be located in the gene coding for a hormone known as leptin. Leptin is important in the regulation of appetite and food intake. Leptin signaling is mediated via binding to the leptin receptor (Ob-R) and subsequent signaling to the hypothalamus. Via this pathway, food uptake, energy expenditure as well as fat and glucose metabolism are regulated (Friedman & Halaas, 1998) (Fig. 2).

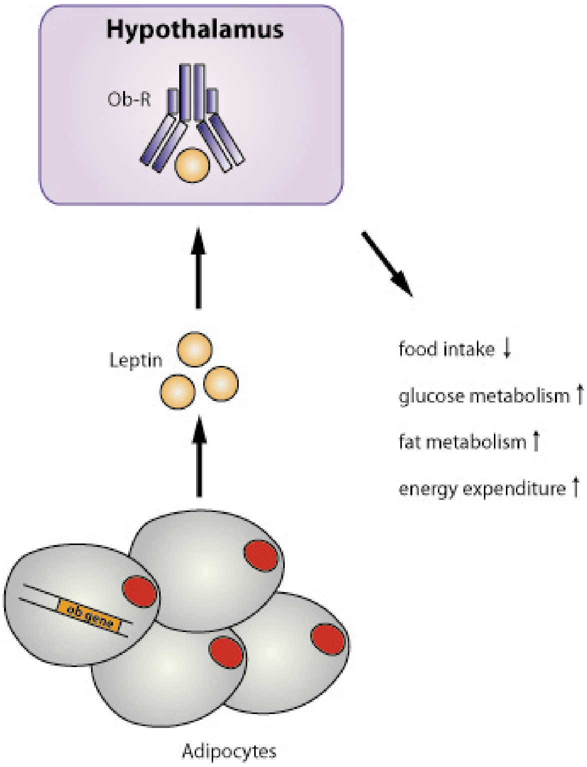


Fig. 2. Regulation of appetite and metabolism by Leptin. Leptin is the product of the *ob* gene and is expressed primarily in adipocytes. In the hypothalamus, Leptin binds to its receptor (Ob-R). It decreases food intake and increases energy expenditure, glucose as well as fat metabolism. Thereby, Leptin contributes to a satiety effect and a lean phenotype.

Due to a lack of leptin, these mice exhibit uncontrolled food intake. Constant overeating therefore results in a gain of body weight. Consequently, ob/ob mice can reach a weight that is three times higher compared to wild type littermates, and their body fat content can be elevated up to fivefold. Apart from this obvious phenotype, ob/ob mice also show decreased physical activity and energy expenditure, infertility and immune deficiencies.

Heterozygotes on the other hand do not display any phenotype as the mutation is recessive. The leptin deficient mouse can be used as an excellent model to study the role of angiogenesis in adipose tissue expansion. Obesity in these mice can be prevented by treatment with anti-angiogenic drugs (Brakenhielm et al, 2004) (this will be discussed more in depth in the Treatment section). Since these mice are comparable to morbidly obese humans regarding the obesity phenotype, using this model might be helpful to identify potential novel targets to treat obesity and obesity-related metabolic disorders in the future.

2.3.2 Db/db mice

The autosomal recessive mutation diabetes (db) was first described in 1966 in the mouse strain C57BL/KsJ (Hummel et al, 1966). These mice are deficient for the leptin receptor. Animals which are homozygous for this mutation, exhibit a phenotype that resembles human diabetes mellitus. This mutant strain is also characterized by an obese phenotype. Furthermore, homozygous mutants are infertile and hyperglycemic while heterozygotes are phenotypically indistinguishable from wild type littermates. These mice are excellent models for studying mechanisms of obesity-related diabetes and insulin insensitivity, and the role of angiogenesis in this regard.

2.4 Other mouse models of obesity

By injecting 3T3 preadipocyte cells subcutaneously into nude mice, researchers are allowed to study the close spatial and temporal correlation between neovascularization and adipose tissue development (Neels et al, 2004). 3T3 preadipocytes will differentiate into adipocytes, and start forming mature adipose tissue *in vivo*. The developing fat pad can be removed at different time points ranging from 1 to 21 days and stained for endothelial cell markers by immunohistochemistry (Fig. 3). Furthermore, the gene expression profile can be analyzed focusing on angiogenesis- as well as adipogenesis-specific genes. Due to the controlled onset and development of the fat pad in adult mice, this model also allowed for studying the origin of the cells that contribute to the formation of new blood vessels during vascularization of the growing adipose tissue.

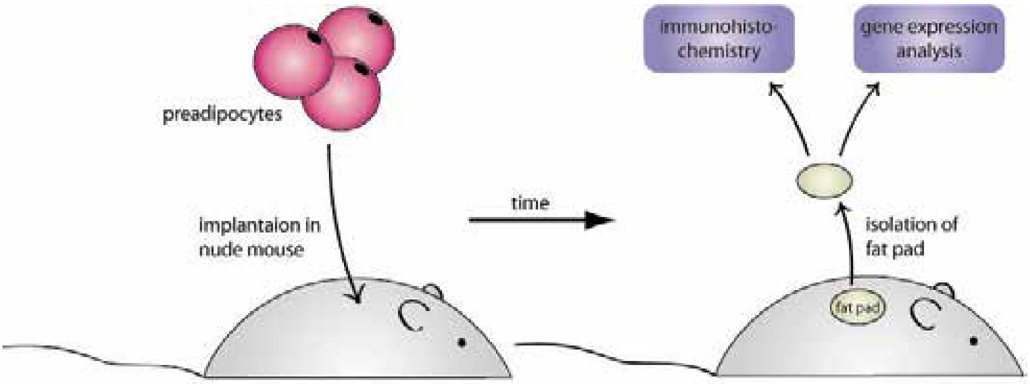


Fig. 3. A mouse model to study adipose tissue angiogenesis. In this model, 3T3 pre-adipocytes are injected into nude mice. Blood vessel formation and gene expression can be studied in the developing fat pad.

2.5 High-fat diet models

One of the first methods to obtain diet-induced obesity in rodent models was based on the so-called cafeteria diet. The cafeteria diet was first described in 1976 as a method to produce obesity in rats (Sclafani et al, 1976). Besides a nutritionally balanced diet, the rats were offered a variety of palatable food items such as cookies and chocolate, candy, salami or cheese and other foods containing high amounts of salt, sugar and fat. Interestingly, the animals ignored the nutritionally adequate chow in favor of the cafeteria food. Although this diet mimics the modern Western diet as consumed by millions of people, a major drawback of this diet is that the animals are allowed to choose independently from the available food. This self-selection may result in substantial differences in the choice of food and thereby nutrients and calories as well as the composition of protein, fat and carbohydrates consumed. Therefore, the cafeteria diet does not guarantee robust and reproducible results.

A more controlled setting to study diet-induced obesity in rodents can be achieved by providing diets in a pellet-form with a high content of fat. The typical standard chow for a laboratory mouse consists of 11.4 % of calories derived from fat, 62.8 % from carbohydrates and 25.8 % from protein, resulting in a nutritional value of 12.6 kJ/g food. A typical high-fat diet may be composed of 58 % fat, 25.6 % carbohydrate and 16.4 % protein. The high content of fat leads to a nutritional value of 23.4 kJ/g food (Winzell AM and Ahren B, 2004). Animals fed with this extreme diet elicit rapid weight gain and are prone to developing obesity.

2.6 Cold induced angiogenesis in adipose tissue

Some types of WAT in mice can acquire a BAT-like phenotype after exposure to cold temperature (4°C). Cold exposure leads to activation of the sympathetic nervous system which increases the capacity of non-shivering thermogenesis (NST) leading to increased heat production that is independent from non-productive muscle activity (shivering) (Xue et al, 2009). Acclimation of rodents to cold should be performed gradually. Mice should be adapted at 18°C for at least one week before transferring them to 4°C. The duration of adaptation is dependent on strains; genetically manipulated strains such as thermogenically incompetent UCP-1 knock-out mice require longer adaptation. Exposure of rodents to cold results in the transition of inguinal WAT to a BAT-like phenotype. Surprisingly, short term exposure (1 week) to cold is sufficient to regulate many genes involved in adipose tissue functions. For example, cold exposure results in up-regulation of BAT-related markers such as (UCP-1) and PGC-1α in the inguinal WAT. The density of blood vessels is highly correlated with the metabolic demand in the different adipose depots. During the transition from WAT to BAT-like, blood vessel density increases dramatically already after one week of cold exposure. After five weeks the WAT exhibits an even higher increase in blood vessel density. Blood vessels are constantly remodeling depending on the metabolic status of the adipose tissues. The transition of WAT to a BAT-like phenotype upon cold exposure is accompanied by the increase in pro-angiogenic factors such as VEGF. This example further demonstrates the importance of tight regulation of blood vessels and angiogenesis in adipose tissue remodeling and function.

2.7 Treatments

The major focus of biomedical research should be on interfering with physiological or pathological processes through treatment. Thus, while the models mentioned in this chapter are valuable in studying the role of the vasculature in initiation and progression of diseases, they are also suitable for evaluation of potentially clinical benefits of novel or previously uncharacterized drugs.

2.7.1 Anti-angiogenic blockade of adipose tissue angiogenesis

Following the example of cold-induced transition from WAT to BAT-like depots described above, this change occurs in parallel with increased expression of VEGF receptors 1 and 2 coupled to increased density of blood vessels in the adipose tissue. Since VEGF-A activates signaling through binding to VEGFR1 and VEGFR2, treatment with neutralizing antibodies against VEGFR1 (MF1) or VEGFR2 (DC101) would provide mechanistic insights into the roles of these receptors. Inhibition with VEGFR2 abolishes the cold-induced vascularization, demonstrating that the VEGFR2 signaling pathway is involved in the regulation of the angiogenic switch in cold. On the other hand, inhibition with VEGFR1 (MF1) resulted in further increased angiogenesis in both WAT and BAT. This suggests that VEGFR1 could be involved in the negative regulation of angiogenesis in adipose tissues (Xue, 2009).

2.7.2 Angiogenesis inhibitors to counteract obesity

As mentioned, the expansion and regression of WAT is highly dependent on angiogenesis. Indeed, treatment with angiogenesis inhibitors, angiostatin, endostatin or thalidomide, results in the reduction of body weight in obese mice (Arbiser et al, 1999). Treatment of leptin deficient obese mice and high fat diet-fed wt C57Bl mice with TNP-470, a selective angiogenic inhibitor, has resulted in reduction of body weight and adipose tissue depot masses. The vascular density in the adipose tissue in TNP-470 treated animals was significantly lower indicating that TNP-470 exerts a direct anti-angiogenic effect on adipose tissues (Brakenhielm, 2004).

Leptin is considered to be a stimulator of angiogenesis. The secretion of leptin is proportionate to the size of adipocytes and regulated by the level of oxygen in adipose tissue. In hypoxic situations such as during hyperplasia or hypertrophy of adipocytes, VEGF protein as well as leptin levels are up-regulated thereby stimulating angiogenesis to provide adequate delivery of oxygen and nutrients to the adipocytes. Co-implantation of leptin with VEGF and FGF-2 in the avascular mouse cornea revealed a remarkable synergistic angiogenic stimulation. However, the treatment of obese individuals with leptin remains controversial. Despite that leptin stimulates angiogenesis, administration of leptin to individuals with a homozygous mutation in leptin genes confer beneficial outcome in terms of body weight reduction (Frederich et al, 1995).

2.8 Methods and assays to study blood vessels in tissues

The models mentioned in this chapter are further strengthened by (usually post mortem) analysis of the tissue in which the blood vessels are growing. In pathological situations the blood vessels are usually of poor quality and functionality, which greatly contributes to progression of the disease. In this section we will discuss several histological and functional tests that can be done to gain additional insight into the structure and quality of the blood vessels.

2.8.1 Hematoxylin/Eosin staining

For Hematoxylin/Eosin (HE) staining of tissues, it is recommended that sections with paraffin-embedded tissue of a thickness of 3 – 5 μm for interscapular BAT (iBAT) and inguinal WAT (iWAT), respectively, are used. For non-adipose tissue types such as tumor, cornea, muscle or dermal tissues, a thickness 5 µm is recommended. Following de-paraffinization with xylene and rehydration using 99.7 %, 95 % and 70 % solutions of ethanol, the slides are stained for 3-5 min. with hematoxylin. This results in a clear blue/purple staining of the nuclei of the cells. Eosin is then used to stain the cytoplasmic contents of the cells pink/red. Depending on the different compartments within the cell, different shades of blue to pink can be observed (Fig. 4). With this method, the adipose tissue can be studied with regard to the size of adipocytes. Blood vessels however, cannot be visualized using this method.

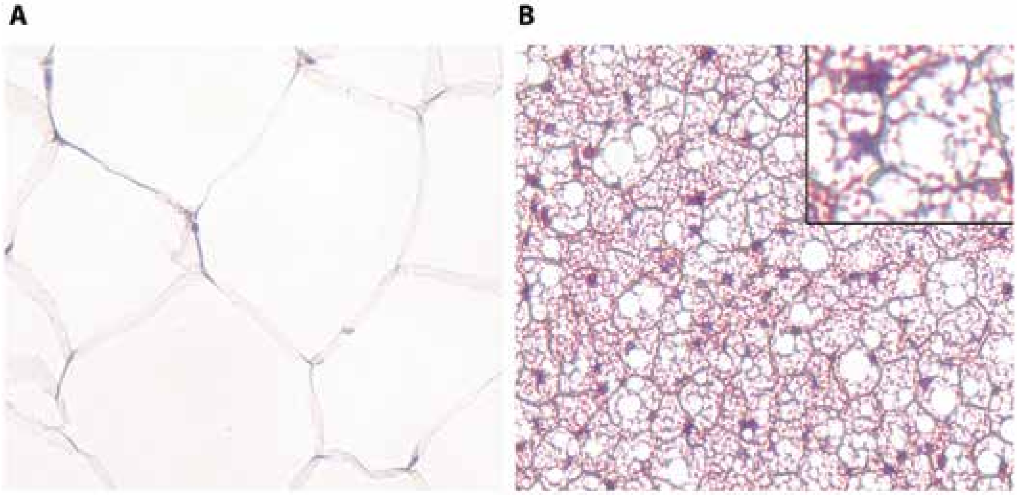


Fig. 4. Analysis of adipocytes by hematoxylin and eosin staining in white (A) and brown (B) adipose tissue. Nuclei are stained with hematoxylin (purple) and cytoplasmic components with eosin (pink).

2.8.2 Whole mount immunohistochemistry

The use of whole mount immunohistochemistry allows investigations of the vasculature in adipose and other tissues, especially in regard to its structure and functionality.

Different primary antibodies can be used to visualize endothelial cells. These antibodies target for example CD31, CD34 or isolectin all of which are expressed on endothelial cells. This provides a general overview of the vasculature in the tissue and its structure, i.e. whether the vessels are organized or disorganized, if their diameter is normal or if they are dilated and also gives information on the presence of microvessels and capillaries that might have been newly formed (Fig. 5). The tissue vascularity can be assessed by calculating the area of stained vessel signals per field. However, it should be noted that an increase of adipocyte size might imply that the number of vessels decreased even if that is not the case. This is due to the fact that with increasing adipocyte size the area per field that can be covered by vessels decreases. However, the ratio of vessels per adipocyte might not have been changed. In order to take this issue into account, it is recommended to calculate vessel number per adipocyte.

To study pericyte coverage, using an anti-NG2 antibody can provide insights into vessel maturation. Early premature vessels show a low pericyte coverage index and are therefore prone to leakiness and can be considered less functional. To gain even further knowledge regarding vessel functionality, perfusion and leakiness, tetramethyl-rhodamine dextran can be used. Perfusion is studied by injecting 2000 kDa dextran, leakiness can be investigated by injecting 70 kDa dextran via the tail vein. The tissue is then fixed in 4 % paraformaldehyde (PFA) and the vasculature is counter-stained with anti-CD31 antibody using a non-red color secondary antibody (Fig. 6).

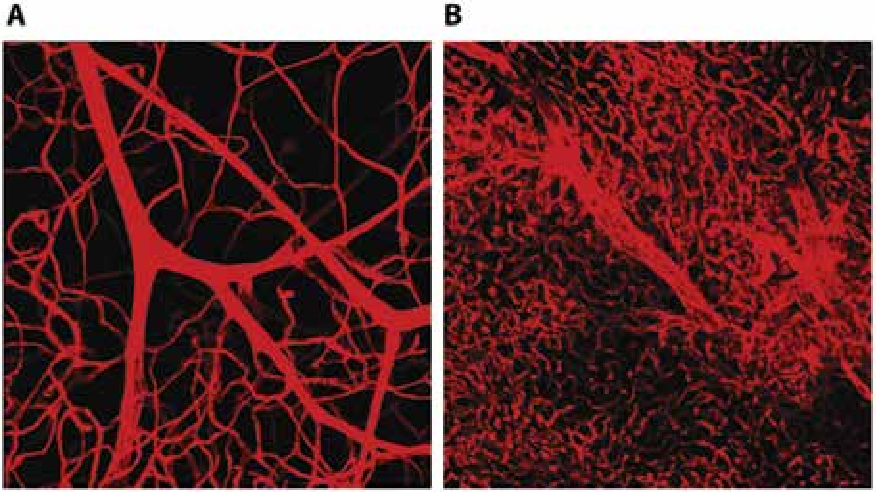


Fig. 5. Visualization of endothelial cells with the marker CD31 in whole mount immunohistochemistry. CD31+ cells (red color) are detected with a fluorochrome-linked secondary antibody in white (A) and brown (B) adipose tissue.

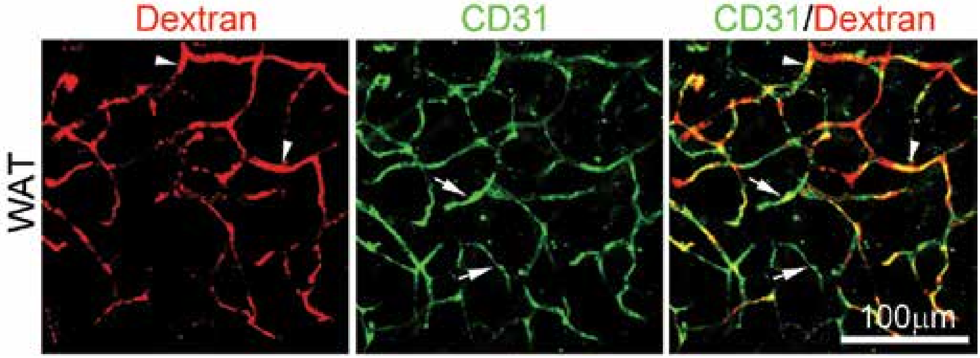


Fig. 6. Functionality test of the vasculature in adipose tissue. Dextran (red) was injected via the tail vein and perfused white adipose tissue (WAT) of a C57Bl mouse was further stained with anti-CD31 antibody (green). Perfused vessels are indicated with arrowheads and non-perfused vessels are indicated with arrows. Scale bar = 100 µm.

Whole mount staining also allows the investigation of lymphatic vessels in adipose and other tissues. Lymphatic vessel endothelial hyaluronan receptor-1 (LYVE-1), podoplanin or VEGFR-3 are common markers which can be targeted by specific antibodies to visualize lymphatic vessels (Fig. 13).

2.8.3 Immunohistochemistry on paraffin sections

While whole mount immunohistochemistry is a valuable tool to study vascular structure and functionality, different methods are advantageous to draw quantitative conclusions regarding the number of blood vessels per adipocyte. For this purpose, the use of thin sections of paraffin-embedded adipose tissue is recommended. In this approach, in contrast to whole mount immunohistochemistry where the thickness of the tissue varies as it is cut by hand, rather uniform tissue sections are prepared using a microtome. Immunohistochemistry on paraffin sections is based on the same principle as whole mount immunohistochemistry: the tissue is incubated with a primary antibody against a protein of interest. A fluorochrome or horse radish peroxidase (HRP)-labeled secondary antibody is then used to visualize the protein (Fig. 7 and 8).

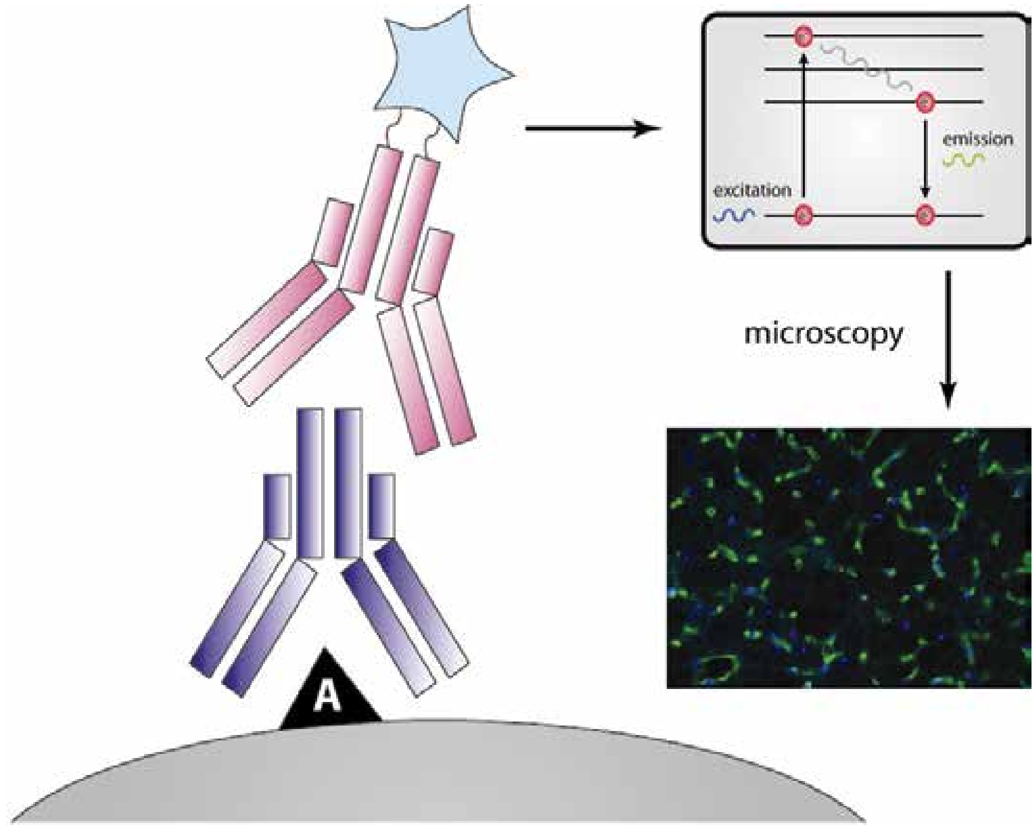


Fig. 7. Schematic representation of immunohistochemistry on paraffin sections using fluorescence-based detection of signals.

It is helpful to counterstain the tissue with DAPI, propidium iodide or Hoechst (for fluorescent stainings) or hematoxylin (for chromogenic HRP stainings) to visualize the nuclei of cells and thereby provide additional information on the structure of the tissue.

To study active angiogenesis, proliferating endothelial cells can be visualized by double staining with antibodies against an endothelial cell marker, such as CD31, and a proliferation marker, such as PCNA or Ki67. Another option is injection of 5-bromo-2'-

deoxyuridine (BrdU) into the tail vain of a mouse 1 min prior to sacrificing the animal. This synthetic thymidine analogue is incorporated into the DNA of proliferating cells. Using an anti-BrdU antibody thereby allows for detection of actively replicating cells.

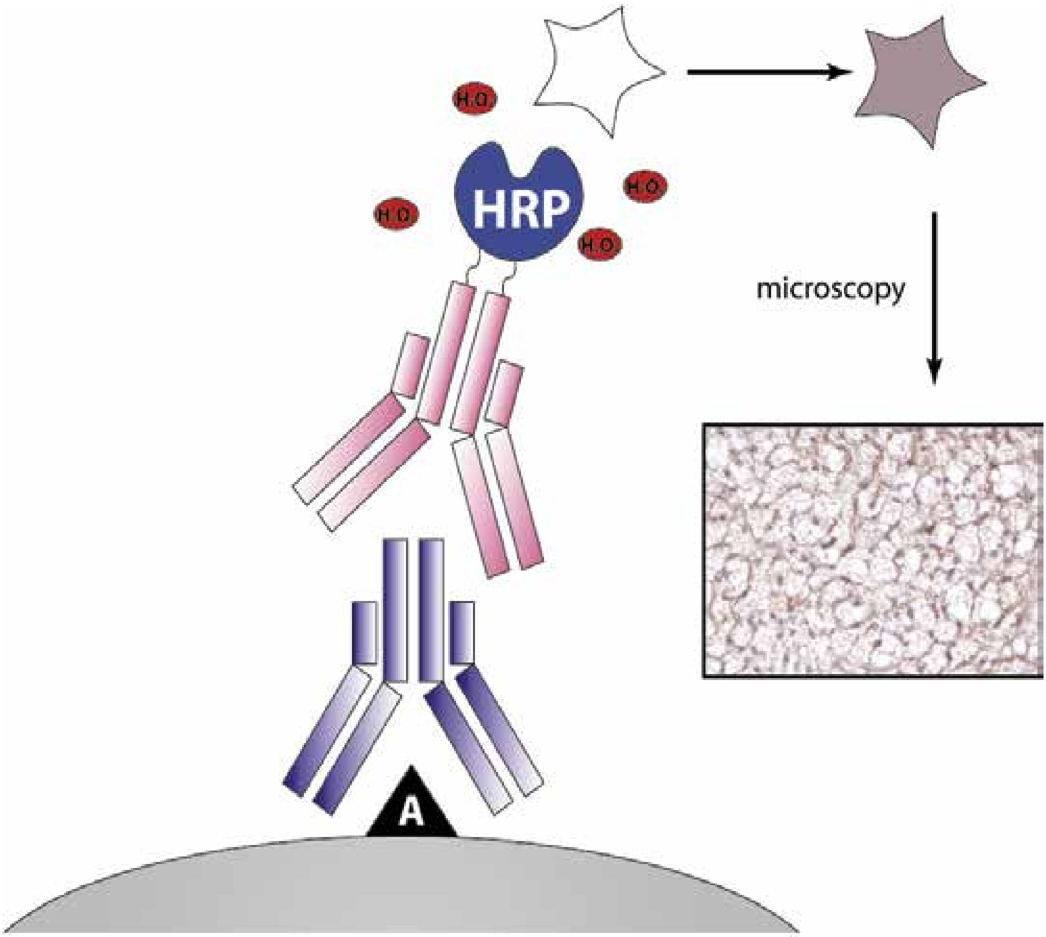


Fig. 8. Schematic representation of immunohistochemistry on paraffin sections using chromogenic detection of signals. The organic compound 3,3'-Diaminobenzidine is oxidized by HRP resulting in a brown color.

2.8.4 Immunohistochemisty on cryosections

Some antigens are masked by paraffin fixation, and in these cases staining has to be done on non-fixed, frozen tissues instead. Dissected adipose or other tissues should be embedded immediately in a plastic cryomold, snap frozen on dry ice and stored at -80°C until staining. Sectioning of cryo-embedded adipose tissue is slightly more challenging than other tissues due to its rather soft integrity. Hence, is it critical to lower the temperature of the cryotome to -30°C before sectioning cryo-embedded adipose tissue samples. Thin sections of 15 µm should be adhered on Superfrost Plus microscope slides, subsequently fixed with cold acetone and stained with specific primary antibodies. However, it is important to note, when staining for lymphatic vessels in the adipose tissue that LYVE-1 is not as widespread on the lymphatic endothelium as podoplanin in this particular tissue (Fig. 9). Here, unlike other tissues, LYVE-1 staining mainly detects inflammatory cells.

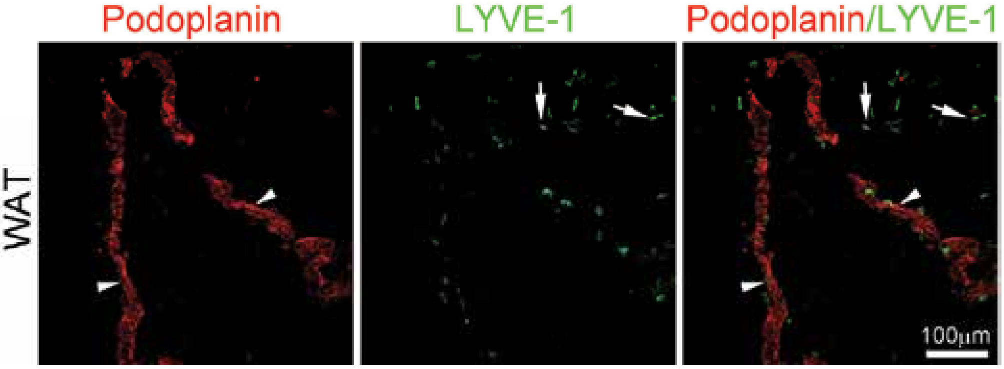


Fig. 9. Double cryosection immunostaining of the WAT of a mouse with podoplanin (red) and LYVE-1 (green). Podoplanin-positive lymphatic staining is indicated with arrowheads and LYVE-1-positive non- lymphatic staining is indicated with arrows.
Scale bar = 100 µm.

2.8.5 Hypoxia staining

Since hypoxia is one of the driving forces of angiogenesis, detection of the presence of hypoxia in tissues could also be performed. Hypoxyprobe-1 (pimonidazole hydrochloride) staining can be used to detect cell and tissue hypoxia. Pimonidazole hydrochloride has a molecular weight of 290.8 kD, ultraviolet absorbance at 324 nm and a plasma half-life of approximately 25 minutes in mice. The detection of hypoxia in tissues can be performed by intravenous or intraperitoneal injection or oral ingestion of pimonidazole hydrochloride at a dosage of 60 mg/kg, 15-90 minutes before sacrificing mice. Dissected tissues should be fixed in 4% PFA followed by paraffin embedding. Paraffin-embedded tissue sections of 3-5 µm is further stained using a peroxidase conjugated anti-pimonidazole antibodies and counterstained with nuclei staining.

3. Tumor models

3.1 Tumor models in general

Mouse tumor-angiogenesis models include xenograft/transplant models and spontaneous models in genetically engineered mice. Xenograft/transplant models are the most widely used, due to the homogeneous and fast onset and progression of the disease (Y Cao, 2005). In these models, human tumor cells are transplanted into athymic nude mice or severely compromised immunodeficient (SCID) mice (xenograft model), or murine tumor cells are transplanted into mice of the same genetic background (homologous transplantation model). Tumor cells can be injected under the skin (heterotypic transplantation, except skin cancer) or into the same organ from which the tumors originate (orthotropic transplantation). It is important to remember when studying tumor angiogenesis that the tumor microvasculature might differ depending on the implantation sites. Orthotropic transplantation is considered to show a similar tumor vessel phenotype as the primary lesion, and thus more closely recapitulate the clinical situation. However, it is often easier and more accurate to follow the tumor growth in heterotypic models such as when the tumor grows under the skin, where it is easily visualized. Furthermore, depending on the tumor growth rate and size, the vessel structure in tumors can be different. Fast growing and big tumors (≥ 1.0 cm^3) usually have extensive necrosis in the center, which affects the vascular network in these areas. The appropriate tumor size for evaluation of tumor angiogenesis is considered to be 0.5-0.8 cm^3. Tumor cells can be modified to express high levels of angiogenic factors, such as VEGF and FGF, or reporter genes, such as green fluorescent protein (GFP), red fluorescent protein (RFP) or luciferase. These modifications allow us to study the function of specific angiogenic factors on tumor microenvironment in association with invasion and metastasis (R Cao et al, 2004). Recent technological developments have permitted us to use genetically manipulated animal models for tumor studies, in which the mice are over-expressing or deficient in angiogenesis related genes. The majority of these mice are however immunocompetent, requiring the use of murine tumor cells in such animals.

Another type of tumor-angiogenesis models are the spontaneous models. Such models are based on genetically engineered mice, where a tumor suppressor has been deleted, an oncogene is being over-expressed or both, often in a particular cell type such as the pancreatic beta-cell of the islets of Langerhans (Hanahan, 1985). These models provide a disease history that is more closely recapitulating the one in human patients, as tumors are generated from one single hyperplastic cell, and progress through steps which are highly similar to pathological progression of pre-malignant to malignant lesions in humans. In later stages of tumor development in these models, angiogenesis also becomes important and contributes to growth and metastasis of these primary lesions (Koh et al, 2010).

To study tumor angiogenesis using mouse models, control animals or tumor conditions must be incorporated in the study. If tumor cells are genetically altered, the proper control would be tumor cells which have an empty expression vector inserted in their genome instead of one coding for a particular angiogenic factor. For evaluation of treatment efficacy of anti-angiogenic drugs, the control group must be given the vehicle (solvent) in which the drug was prepared, because some vehicles can themselves affect the vascular structure in tumors. Age and genetic background of the mice must also be standardized for the different groups. Standard age for tumor experiments is 6-10 weeks. Sex is usually not important, unless the tumor cells under investigation are gender-specific such as in the case of breast or prostate cancer, but should nonetheless be the same in for example treatment and control groups.

3.2 Tumor models: Assessment of angiogenesis

Dr. Judah Folkman proposed in 1971 that all tumor growth is angiogenesis-dependent (Folkman, 1971). In general, because a limiting factor of tumor growth is the supply of sufficient levels of oxygen and nutrition, fast growing tumors are characterized by a more aggressive angiogenic phenotype which arises from the ability of the tumor cells to secrete angiogenic growth factors to support and change their microenvironments. For instance, tumor cells transduced with the gene coding for VEGF grow much faster than the empty-vector transduced control cells, due to much higher intra-tumoral blood vessel density (Eriksson et al, 2002). In these tumors however, the vasculature consists of highly irregular and immature vessels and many vascular plexuses, which does not support efficient perfusion of blood compared to a less chaotic vasculature in VEGF non-transduced tumor tissue (Fig. 10). To assess angiogenesis in mouse models, there are thus three essential points which should be addressed.

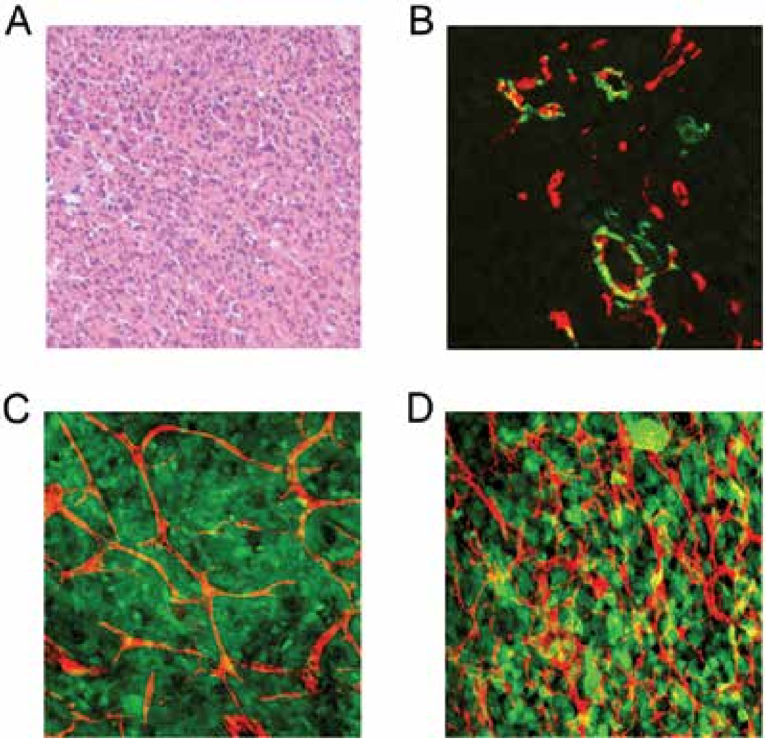


Fig. 10. Histological examinations of tumor tissue. (A) H&E staining of T241 fibrosarcoma. (B) Immunohistchemical staining of T241 fibrosarcoma visualized by fluorescence microscopy. Anti-endomucin labeling of endothelial cells is shown in red and anti-α-smooth muscle actin (SMA) labeling of pericytes/smooth muscle cells is shown in green. (C) Confocal imaging of T241 fibrosarcoma. GFP transfected tumors are shown in green, anti-CD31 labeling of endothelial cells is shown in red. (D) Confocal imaging of VEGF overexpresing T241 fibrosarcoma. GFP transfected tumors are shown in green, anti-CD31 labeling of endothelial cells is shown in red.

1. Tumor size (volume). Tumor growth depends on angiogenesis. Thus, tumor size is usually proportional to the degree of which tumor-angiogenesis is induced by the tumor. If the xenograft/transplant models are used and especially if tumors are transplanted dorsally under the skin, the size is easily measured and an accurate tumor growth curve can be generated by daily measurements of tumor width (W) and length (L) using calipers and calculating the volume (V) as $V=0.52*W^2*L$. If the tumor is not

visible because it is growing in a location inside the animal, special imaging systems, such as CT, magnetic resonance imaging (MRI) or bioluminesence imaging analysis (if tumors express luciferase) are needed to estimate the tumor size. The effect of anti-angiogenic drugs can be inferred from the reduction in tumor growth relative to vehicle treated controls.

2. Histology. Tumor tissues can be stained following the procedures listed in the section on adipose tissue angiogenesis. HE staining will give a general, structural view of the tumor, while immunohistochemical methods are used to visualize the tumor vasculature specifically. In tumors good antigens for visualizing the vasculature are for example von Willebrand Factor (vWF), CD31, CD34, endomucin, isolectin or VE-cadherin. These methods enable analysis of morphological changes and quantitative assessment of the tumor micro environment (Fig. 10). Immunohistochemical staining for multiple markers in the same sample may provide information on the localization of cells relative to each other, cell phenotypes, cell numbers, cell size, and cell conditions (i.e. apoptotic, proliferative).
3. Functional assay of tumor vessels. Tumor vasculatures are irregular and disorganized as well as leaky and poorly perfused. The stained vessels are not all functional; therefore evaluation of vessel functionality such as vessel permeability and perfusion, by using the dextran-injection method described in the adipose angiogenesis section, gives important qualitative information in all tumor models. Upon evaluation under the microscope, abluminal localization of dye indicates extravasation whereas functional vessels are characterized by retaining the dye within the vessel (Hedlund et al, 2009).

3.3 Lymphangiogenesis and lymphatic metastasis

It is well known that blood vessels can support tumor growth by providing oxygen and nutrients, and removing waste products, but the function of tumor lymphatic vessels remains poorly understood (Y Cao, 2008).

Lymphatic networks consist of lymphatic capillaries, collecting lymphatic vessels, and lymph nodes. Unlike blood vessels, lymphatic capillaries consist of one layer of lymphatic endothelial cells (LECs), discontinuous basement membrane and few vascular smooth muscle cells (VSMCs). Lymphatic vessels lack tight junctions between endothelial cells, but are instead equipped with one-way lymphatic valves which give these vessels the ability to collect fluids and macromolecules from the tissue and transport it back to the circulation. On the other hand, these features of the lymphatic endothelium also mean that malignant cells can easily enter into and disseminate via the lymphatic system, leading to lymphatic metastasis. Tumor lymphangiogenesis is therefore associated with cancer metastasis. In some common cancers, such as lung and breast cancer, lymphatic metastasis is the dominant route for tumor metastasis. Invasion of intra- or peri-tumoral lymphatics may result in dissemination of malignant cells to the lymphatic system, leading to lymphatic metastasis in regional lymph nodes (Y Cao Y, 2005; R Cao et al, 2004).

Similar to blood vessels, lymphatic vessels are quiescent in healthy individuals. The formation of lymphatic vessels in tumors is a multistep process that involves LEC proliferation, migration, tube formation and remodeling, which require up-regulation of lymphangiogenic stimulators and down-regulation of lymphangiogenic inhibitors (Y Cao, 2005). Lymphatic vessel growth may represent the imbalanced consequence between positive and negative regulators tipping toward positive regulation. Understanding the molecular mechanisms that control lymphangiogenesis is therefore an important step in the development of therapeutic agents in the prevention and treatment of cancer metastasis.

Among the list of lymphangiogenic factors, members of the VEGF family are the best characterized. VEGF-A, which binds to VEGFR-2 and VEGFR-1, PlGF and VEGF-B which bind to VEGFR-1 and especially VEGF-C and VEGF-D which bind to VEGFR-3 are lymphangiogenic factors, the -C and -D isoforms being the most potent, which regulate both physiological and pathological lymphangiogenesis. VEGF-C/VEGF-D-VEGFR-3-mediated signals are also critical for the sprouting of the first lymphatic vessel from the developing veins in the embryo. This signaling pathway is essential for differentiation of endothelial progenitor cells into the lymphatic lineage (Kukk et al, 1996; Alitalo et al, 2005).

Primary tumors produce several lymphangiogenic factors, including VEGF-A, VEGF-C, VEGF-D, Insulin-like growth factor (IGF), hepatocyte growth factor (HGF) and PDGF-B. These factors induce angiogenesis and lymphangiogenesis both in the local environment and in the regional lymph nodes. Furthermore, they play an important role in establishing the pre-metastatic niche. The term pre-metastatic niche describes the adaptations of for example the lymph nodes which are needed to allow disseminating tumor cells, arriving at a later stage, to meet optimal conditions for growth in that particular site. In addition to signals produced by the tumor cells themselves, inflammatory cells such as macrophages are recruited to tumors by a wide range of tumor cell-derived cytokines and growth factors. At the tumor site, inflammatory cells play a critical role in mediating lymphangiogenesis most likely through the secretion of several lymphangiogenic cytokines.

To detect lymphatic vessels in or around tumors, immunohistochemical staining with lymphatic specific markers, such as VEGFR-3, LYVE-1 and podoplanin is recommended (see Fig. 13). Lymphangiogenesis can be evaluated using all of the models described in this chapter, by using antibodies against one or a combination of these factors either in whole mount staining or on thin sections of frozen or paraffin embedded tissue (Fig. 13).

4. Non-tumor, xenograft models

As tumor cells often produce a multitude of factors which may cooperate in inducing angiogenesis and lymphangiogenesis, these models are considered to be relatively “dirty” and unsuitable for studying the effects of just one or a few factors. In order to overcome this problem, a xenograft model for non-tumor angiogenesis has been developed. In this model, a matrigel plug is grafted onto the mouse which can be mixed with recombinant angiogenic factors prior to grafting.

4.1 Matrigel plug assay

Matrigel consist of purified basement membrane components (collagens, proteoglycans and laminin) and, while it is liquid at temperatures just above 0 degrees, it forms a gel when it is warmed to 37 °C. Thus, the material can be cooled and then injected in the mice, where it will form a three dimensional gel, in which host blood vessels can invade. Matrigel itself is a poor inducer of angiogenesis, but it can be mixed with angiogenic growth factors and/or cells prior to injection leading to a controllable induction of blood vessel growth into the plug. Plug vessels are usually evaluated 7-21 days after implantation by gross examination/photography as well as immunohistochemical staining as described above (Akhtar et al, 2002). If the plug contains functional vessels, the blood (red) vessels can be identified from the photograph. Alternatively, by using mice which express GFP in the endothelium, immunohistochemical staining can be avoided. Also in this model, intravenous dye injection can be performed to evaluate vessel perfusion and leakiness.

5. Rat ischemic hind limb model

Most of the models mentioned above are very useful to study pathological angiogenesis, and therefore used to find novel anti-angiogenic treatment options including novel anti-angiogenic drugs, but other assays are needed to study diseases where therapy would consist of accelerated blood vessel growth, such as in the treatment of myocardial infarction (MI), stroke and wound healing/regeneration.

In MI an occluded coronary artery leads to a blockade of blood flow to a part of the cardiac muscle tissue, and thus leads to severe hypoxia (ischemia). The cardiac musculature is working constantly, has a very high metabolism and is therefore particularly sensitive to reduced oxygen (and sugar) levels. Therefore, unless new, so called collateral, arteries can be formed quickly, the affected cardiac tissue will perish, and usually so will the patient (Y Cao, 2010). Effective therapy would therefore be able to induce growth of highly functional arteries in the response to tissue hypoxia/ischemia.

An excellent model to study and manipulate the growth of blood vessels and in particular arteries in response to tissue hypoxia is the hind limb ischemia model, which can be performed in either mice or rats. In this model, all arteries supplying highly oxygenated blood to one of the back limbs of the animal are ligated in two steps, resulting in near-zero blood flow in the entire limb (Lundberg et al, 2002.) (Fig. 11). This leads to tissue

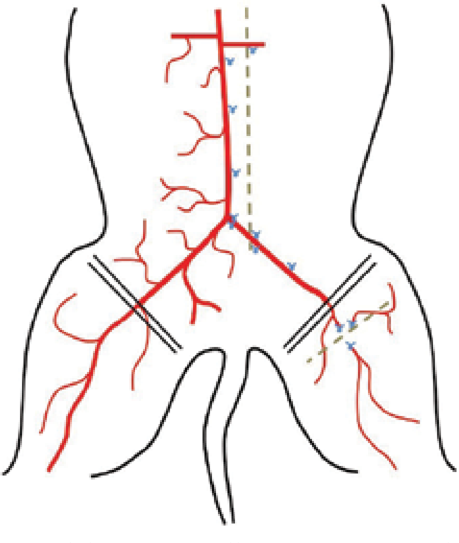


Fig. 11. Hind limb ischemia model. First, a midline incision is made (top dashed line) exposing the vessels. All branches originating from the aorta distal to the renal arteries and all branches from the left iliac artery are ligated by a small suture (blue knots). After a week a second inguinal incision is made (bottom dashed line) and the femoral artery and superficial gastric artery are ligated by a small suture (blue knots). Following this operation the left hind limb does not receive any blood and is considered ischemic.

ischemia and the induction of arteriogenesis from collateral arteries. Angiogenic factors under investigation can be injected into the limb musculature or antibodies against angiogenic or anti-angiogenic factors can be injected into circulation, thus modulating the arteriogenic response. As the hind limb is easily accessible, the circulation in the limb can be observed by Doppler-angiography and thus the same animal can be subjected to repeated investigations on how the blood flow improves over time. After euthanasia, the tissue can be excised and stained as mentioned for the adipose tissue, and the morphology of the blood vessels can therefore be studied post mortem (R Cao et al, 2003). This assay is the most commonly used assay to study therapeutic angiogenesis, and has been used in many seminal discoveries on how to therapeutically generate highly functional and stable arteries (Y Cao, 2010). However, the surgery needed to induce ischemia in the hind limb is very complicated and the assay therefore requires highly skilled and experienced surgeons. Also, there may be slight inter-individual differences in the residual blood flow in the limb after surgery – and therefore the degree of tissue hypoxia - within each experimental group.

6. Wound healing assay

The process of wound healing is divided into three stages; inflammation, new tissue formation and remodeling. Angiogenesis is critical for the formation of new tissue and vascular remodeling occurs in association with tissue remodeling at later stages. Therefore, the wound healing assay is a useful model to evaluate both angiogenesis and vascular maturation/remodelling. Wound healing models are usually performed on the skin as other accessible tissues such as the ears or the tail do not regenerate well. Usually, two circular, trans-dermal wounds are created on the back of anesthetized mice (Fig. 12) allowing one wound to serve as control while topical treatment can be administered on the other. Wound size, scar formation and re-epithelization of the wounds should be recorded daily by photography and by measuring the wound area with calipers, similar to how it is done in the tumor models. In this model, treatment given either systemically by oral administration or injection, or preferably topically on just one of the two wounds, can consist of pro- or anti-angiogenic compounds, and their effects on both the regenerative angiogenesis as well as on vessel morphology and function can be determined post mortem after the regenerated tissue has been excised, fixed and stained as mentioned for the adipose tissue. Also, transgenic or knock-out mice can be used, when available, to study the specific effects of particular genes (Xue et al, 2008).

The surgery required to create the wounds is very simple and the wounds are highly homogeneous in size and location for all animals used in the experiments. This assay is therefore very easy and robust, requires little practice and few animals as the experimental variation is relatively low. However, angiogenesis in this model occurs only in the skin, and in association with inflammation, blood clotting cascades and other highly complex biological processes, which encompass multiple cell types and a plethora of angiogenic and anti-angiogenic factors. Furthermore, regeneration of the skin is known to be quite different from regeneration of other tissues of higher clinical relevance such as the heart and nervous system, and can therefore not be expected to give much information on the role of angiogenesis in regeneration of other organs. Another drawback of this assay is that regeneration occurs by making new tissue rather than repairing/replacing damaged/dead tissue. This difference is thought to be important, as the latter is almost always the case in clinical situations where for example ischemic insults result in large patches of dead tissue that is hard to replace.

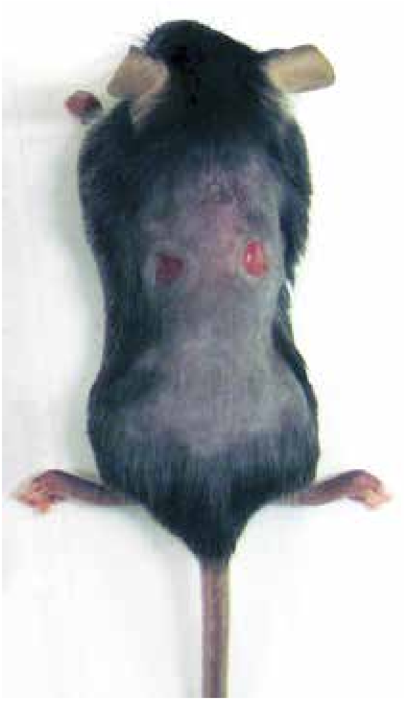


Fig. 12. Wound healing model. Two circular holes of approximately 5 mm in diameter are punched with a tissue puncher through the dorsal skin of an anesthetized wildtype C57Bl6 mouse. No bandages or cover is needed, as the skin in this region has no major blood vessels, and the wound formation leads to very little if any bleeding. Wounds close within 2 weeks and would heal completely within a month.

7. Cornea models

The cornea is an avascular tissue consisting of two, thin, transparent layers in rodents. Thus it is possible to gently cut a tiny pocket between the two layers, and in this pocket insert a pellet containing factors which are to be investigated for their angiogenic or anti-angiogenic activities *in vivo*. All vessels detected in the cornea following a few days stimulation with the implanted factors can be considered newly formed vessels, and due to the transparent nature of the cornea and the strong red color of perfused blood vessels, the angiogenic response can be followed kinetically by simply taking photographs of the eye at different time points. This makes it possible to study the effects of angiogenic factors, either alone or in combination on different processes of angiogenesis such as initial angiogenic expansion, vascular remodeling, maturation and stability in the same animal over time.

Usually angiogenic or anti-angiogenic factors are prepared in slow release pellets consisting of hydron or alumni sucrose octo-sulfate, which are dried on a nylon mesh to ensure equal size and thereby amount of factor in each pellet which is then implanted in the corneal micropocket. Using this delivery system, factors are released constantly at low amounts for weeks, thus giving rise to a continuous angiogenic stimulation or inhibition depending on the factor in question (R Cao et al, 2003). Alternatively, small pieces of tumor tissue can be implanted to study the angiogenic capacities of different tumors from different organs. Even primary human tumor samples can be investigated in this model by using immune-deficient nude or SCID mice (Jensen et al, 2009).

While blood vessels can be macroscopically studied by simply taking photographs of the cornea, more in-depth examinations on blood vessel structure and function including pericyte coverage, tip cell formation and vascular permeability require that the cornea is excised post mortem and subjected to immunohistochemical analysis such as those described for the fat tissue (Fig. 13).

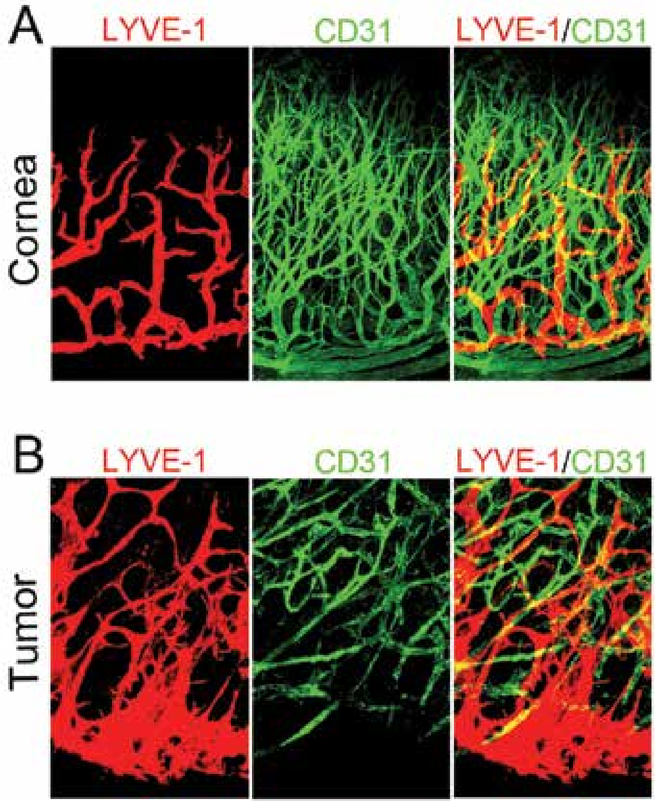


Fig. 13. Visualization of lymphangiogenesis and angiogenesis. A: Confocal analysis of mouse FGF2-induced corneal lymphatic vessels (red) stained with an anti-LYVE-1 antibody and blood vessels (green) stained with an anti-CD31 antibody. B: T241-VEGF-C tumor lymphatic vessels (red) stained with an anti-LYVE-1 antibody and blood vessels (green) stained with an anti-CD31 antibody.

Following staining with antibodies against markers specific for lymphatic endothelial cells, this assay is furthermore one of the strongest assays for investigating lymphangiogenesis.

Furthermore, by treating the mice with drugs or neutralizing antibodies against particular receptors, this model can be used to parse out the specific contributions of one or a few receptors for a particular angiogenic factor in inducing angiogenesis and lymphangiogenesis.

This model is considered to be "clean" compared to the tumor models described above – i.e. the angiogenic or lymphangiogenic response can be induced by a defined factor or couple of factors. Considering all these benefits – this assay is probably the strongest *in vivo* assay to study the molecular biology of angiogenesis and lymphangiogenesis (Jensen et al, 2009).

However, the assay has been criticized for being poor at describing physiologic or pathological angiogenesis as the amount of factor used in the pellet usually give rise to supra-physiological concentrations in the cornea. However, the amount of factor used per pellet can be lowered accordingly, or instead of using purified recombinant growth factors, researchers may implant tumor tissue, or a small piece of suture, which lead to corneal inflammation, which then subsequently drives the angiogenic response although via a less defined and controllable pathway. The major limitation of the assay is the technical difficulty of implanting pellets into the mouse cornea. In the beginning, researchers circumvented this problem by using larger animals such as rats or rabbits, but in order to take advantage of the growing number of transgenic or knock-out mice available today it is a major benefit to use this animal if possible.

8. Zebrafish models of angiogenesis

Zebrafish have recently gained much attention as an angiogenesis model system. Zebrafish embryos develop outside of the uterus, which greatly facilitates imaging during development. Furthermore, pigmentation can be inhibited chemically, or by using non-pigmented strains, such that the embryos maintain transparency throughout development. Many eggs are produced each breeding cycle, zebrafish are relatively easy and cheap to maintain, compared to rodents, and their embryonic development is much faster. Additional benefits of zebrafish-based models include passive uptake of chemicals added to the water, which eliminate the invasive administration procedures which are commonly needed in rodents and their small size which means that all tissues can be oxygenated by passive uptake from the water during development, thus allowing researchers to study phenotypes of embryos with severely disrupted vasculature, which are embryonically lethal and therefore difficult to study in mice (Weinstein et al, 1995).

Because of these unique characteristics of the zebrafish embryo, several genes with essential functions in the vasculature have been discovered using this system. One approach which has yielded the identification and characterization of novel genes important for developmental angiogenesis and vascular maturation is to perform unbiased screens of mutant embryos generated by random mutagenesis induced either by radiation, chemicals or insertion of small genetic fragments (Gaiano et al., 1996; Haffter et al., 1996; Knapik et al., 1996). Such screens have in the past yielded information on human congenital disorders of which the genetic background was previously unknown (Alders et al, 2009), and have – combined with the ability of the embryos to passively take up chemicals added to the water – been used to screen for, and identify novel drugs which efficiently correct the pathological vascular phenotypes (Peterson et al, 2004). Such an approach is called chemical genetics, due to the screening of novel chemicals for therapeutic effects in genetically modified zebrafish models, and is today being frequently used as a discovery tool in the pharmaceutical industry.

Recently researchers have further expanded the benefit of zebrafish-based model systems by generating many transgenic zebrafish strains which express fluorescent markers in particular cell types, organs or tissues, including endothelial cells of the vasculature (Jensen et al, 2009). By continuous observation of such transgenic embryos under the microscope, it is possible to follow the dynamics of growing vessels during zebrafish development in real time.

Such studies, which are difficult to perform in mice due to their *in utero* development, have yielded valuable insights into the process of vasculogenesis, which is the formation of the first embryonic vessels – the aorta, cardinal vein and thoracic duct – and on the origin of blood cells as well as the mechanism by which blood flow is initiated (Herbert et al, 2009; Lida et al, 2010; Yaniv et al, 2006).

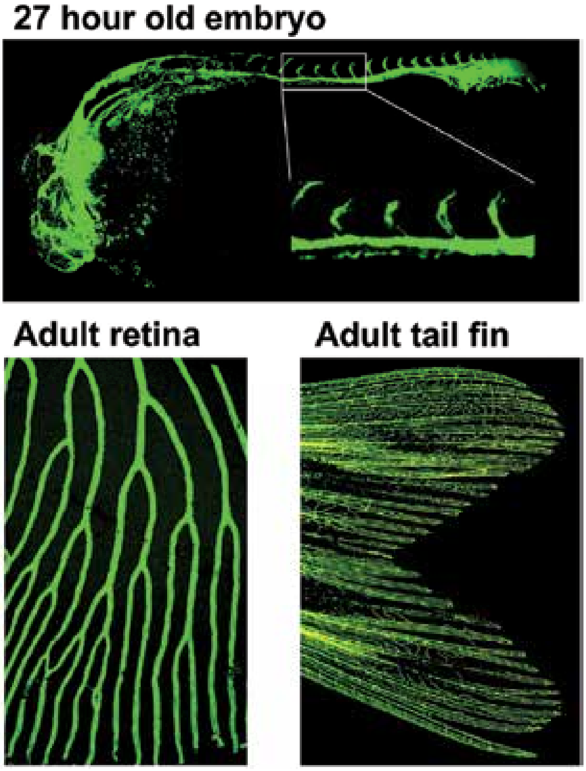


Fig. 14. Blood vessels in the *Fli1:EGFP* transgenic zebrafish line. In the *fli1:EGFP* transgenic zebrafish, the promoter drives EGFP expression specifically in the vasculature. The top image show EGFP positive blood vessels (green) in a 27 hours old embryo, on the bottom left is shown blood vessels in the adult retina and on the bottom right in the adult tail fin.

The zebrafish genome has been fully sequenced and annotated, which makes it easy to interfere with genes involved in vascular development by injecting specially designed synthetic RNA-like molecules called morpholinos. These morpholinos are designed to have a complementary sequence to a particular target mRNA to which they will anneal and thus specifically target that transcript for destruction. The addition of morpholic acid groups on the nucleotide backbone of morpholinos makes them stable inside the cell and they are therefore not degraded along with their mRNA target (Nasevicius & Ekker, 2000). Morpholinos, as well as mature mRNA - if over-expression rather than inhibition of a particular gene is being investigated - can be injected in the yolk of newly fertilized 1-cell stage embryos, and will then be present in all cells during development. As the cell number increases, the concentration of morpholino per cell will gradually decrease, and most morpholinos will therefore only be effective for up to 4 days after fertilization. This period is – due to the fast development of the zebrafish embryo – usually more than sufficient for angiogenesis studies, and thus does usually not pose a restriction on the study.

Morpholino-mediated disruption of a particular gene product is called knock-down, due to its transient nature compared to permanent knock-outs, and is today a widely used method to study the involvement of particular genes in angiogenesis and vascular functions during zebrafish development.

By using a zebrafish strain with enhanced green fluorescent protein expressed in endothelial cells under the *fli1* promoter (Lawson & Weinstein, 2002) (Fig. 14), knock-down of VEGF leads to a concentration-dependent inhibition of angiogenesis which can be quantified already 24 hours after fertilization (Nasevicius et al, 2000). In mice, even heterozygous deletion of VEGF leads to early embryonic lethality (Carmeliet et al, 1996) and the developmental consequence of subtle concentration differences in this critical angiogenic factor can therefore only be studied in zebrafish. Thus, zebrafish have been used recently to show that VEGF is only important for arterial growth, whereas growth of veins are practically unaffected by reduced VEGF levels.

9. Embryonic zebrafish metastasis and tumor model

Zebrafish embryos in the first few weeks of development lack an adaptive immune system, which makes them unable to reject non-zebrafish grafts and they are therefore perfect recipients for implantation of mammalian tumor cells. Due to the transparent nature of zebrafish embryos, fluorescently labeled tumor cells may furthermore be traced in a completely non-invasive fashion, leading to continuous monitoring of the behavior of the tumor cells, including their dissemination from the primary tumor mass *in vivo* (Rouhi et al, 2010).

The induction of tumor angiogenesis is a hallmark of advanced tumors, and therefore also of established tumor cell lines. Thus, mammalian tumor cells, implanted into a space around the yolk sac called the perivitelline space of transgenic *fli1:EGFP* zebrafish embryos, is a powerful way of studying early events of tumor-induced angiogenesis and especially of tumor cell dissemination and metastasis via the vasculature (Lee et al, 2009; Rouhi et al, 2010). Depending on their invasive capabilities, implanted tumor cells may start disseminating from the primary site very shortly after implantation (Fig. 15). Angiogenic factors produced by the tumor cells can affect the host's existing and developing vasculature which furthermore is a good way of studying the function and morphology of vessels induced by the tumor-derived factors in question. Different kinds of manipulation e.g. knock-down or over-expression of a specific gene in the tumor cells can affect the dissemination and invasion pattern of a specific cell line implanted in the embryo. Genetic manipulation of tumor cells can be performed with help of different tools e.g. small interfering RNA (siRNA), small hairpin RNA (shRNA) or by retroviral integration of expression vectors containing angiogenic factors.

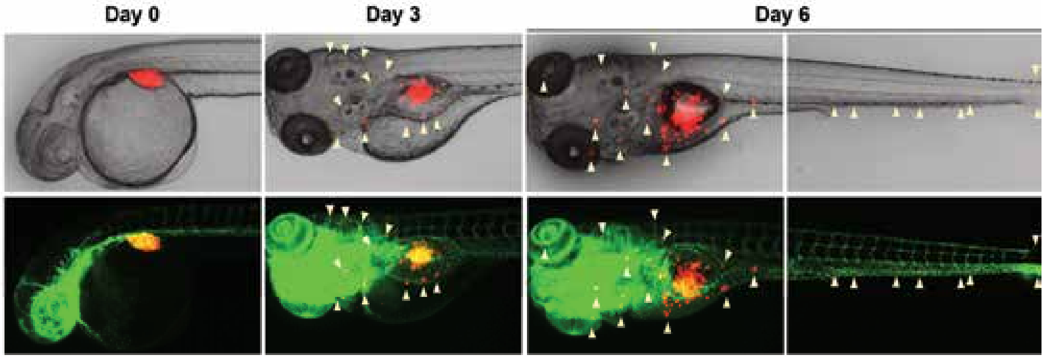


Fig. 15. Metastasis model in zebrafish embryos. DiI labeled tumor cells (red) are injected in the perivitelline space of approximately 2 day old zebrafish embryos (left panels). 3 days after injection many disseminated cells are observed (arrowheads) proximal to the tumor (middle panels). After 6 days tumor cells have spread throughout the embryo, via the vasculature (right panels). The top row shows brightfield images combined with the red-fluorescent signal from the tumor cells. The bottom row shows the EGFP-positive vasculature (green) in transgenic F*li1:EGFP* zebrafish embryos combined with the red-fluorescent signal from the tumor cells.

It is also possible to study the effects of host gene expression on tumor cell dissemination and metastasis by modifying expression of host genes by injecting either morpholinos or mature mRNA immediately after fertilization, and thus prior to tumor implantation, as described above (Lee et al, 2009). This model is thus a simple system to parse out the contributions of factors derived from the tumor relative to the host on tumor angiogenesis, dissemination and metastasis.

Since disseminated tumor cells in this model will be detected by fluorescent microscopy, they should either permanently express a fluorescent protein or they can – prior to injection – be labeled with fluorescent dyes such as DiI, which can be easily distinguished from the green fluorescent color emitted from the embryo vasculature. Tumor cells are preferably implanted in the non-vascularized area in the perivitelline cavity as it can otherwise be difficult to distinguish between existing vessels, those formed by normal developmental processes and those induced by the tumor mass.

Prior to implantation, embryos will be dechorionated, anesthetized and placed on a modified agarose gel. Transplantation of tumor cells is easily carried out using a micromanipulator which is connected to a microinjector. Approximately 100 cells will be implanted into each embryo.

After microinjection, based on the purpose of the study, embryos will be immediately transferred into appropriate embryo water such as E3 water or Danieu's buffer (Fig. 15). As mentioned above, drugs may be added to the water and their effects on tumor angiogenesis, dissemination and metastasis can thus easily be evaluated.

Each tumor bearing embryo may be placed in a separate well of a multi-well plate and examined individually to monitor tumor angiogenesis, dissemination and metastasis in as high temporal resolution as required in the experiment. This model furthermore allows generation of time-lapse video sequences of the tumor cells invading into and out from the vasculature at metastatic sites, and thus examine the tumor-endothelium communications involved in tumor dissemination and metastasis in detail.

As described in more detail below, zebrafish are – compared to mice – highly amenable to studies on the physiological and pathological effects of hypoxia. In order to investigate the effects of hypoxia on tumor angiogenesis, dissemination and metastasis, tumor cell-transplanted embryos can be placed in hypoxic water in a special aquarium (Lee et al, 2009). As tumor hypoxia in mammalian models is very difficult to control and monitor, this system allows studies of the effects of highly defined oxygen concentrations and periods of hypoxia exposure. In all conditions above fish embryos will be kept at 28.5°C, which is the standard temperature for rearing zebrafish embryos and larvae. Death rate is relatively high in hypoxia experiments and in order to have enough embryos at the end to make statistically correct conclusions, a relatively high number of embryos should be implanted with tumor cells and placed inside the hypoxia chamber.

10. Spontaneous tumor models in zebrafish

As it is the case for mice, there are also several genetically engineered spontaneous tumor models available in the zebrafish. While these models may be regarded to more closely model the clinical pathogenesis of cancer, the same limitations are relevant for the zebrafish models as those described for mice, most notably the very heterogeneous tumor onset and growth both for different tumors within each fish but also between different fish.

10.1 Peripheral nerve sheath tumors

Homozygous *tp53M214K* mutant zebrafish which carry a mis-sense mutation in exon 7 of the *P53* gene, giving rise to spontaneous formation of abdominal and periocular tumors at an age of 8-16.5 months. HE staining of ocular and abdominal tumors showed that these tumors are mainly composed of spindle cell and epitheloid cell neoplasms which recapitulate characteristics of human malignant peripheral nerve sheath tumor (MPNST).

Intraperitoneal implantation of extracted zMPNST cells into irradiated wild type adult zebrafish led to formation of typical zMPNST tumors after less than one month post transplantation (Berghmans et al, 2005).

10.2 Melanoma

By crossing the *tp53* mutant zebrafish with a transgenic line expressing a mutant, constitutively active form of BRAF (V600E, which is the same mutation frequently observed in human melanomas) under control of the melanocyte microphthalmia-associated transcription factor (*mitfa*) promoter, it was found that these fish develop melanomas spontaneously in preference to MPNSTs. In these fish formation of small nevi can be observed already at a young age, and these nevi continue to progress into large malignant, metastatic melanomas in multiple locations from approximately 4-6 months of age.

Malignant melanoma tumor cells from the tumor bearing fish can be transplanted into sub-lethally gamma irradiated wild type zebrafish, where they will grow and spread much like tumors in the mouse xenograft models described previously (Patton et al, 2005).

10.3 T-ALL

Aberrant expression of Myc in humans induces lymphoma and leukemia. Similarly in zebrafish a T cell acute lymphoblastic leukemia (T-ALL) zebrafish model was generated by microinjection of the mouse *c-myc* (*mMyc*) gene under control of the zebrafish *Rag2* promoter (*zRag2*). Onset of tumors varied between 30 to 131 days post injection of the *Rag2-mMyc* expression plasmid. Fish developing T-ALL were characterized by inflated abdominal cavities and infiltration of malignant cells (transformed lymphoblasts) throughout the body, under the skin, into base of the pectoral fin, olfactory region, and the retrorbital soft tissue that led to splayed eyes.

These malignant lymphoblasts were transplantable into irradiated wild type adult zebrafish giving rise to small tumors appearing already one week post transplantation (Langenau et al, 2003)

10.4 Rhabdomyocarcoma

Rhabdomyosarcoma (RMS) is a very aggressive soft tissue sarcoma with high incidence seen among children compared to other types of cancer. The zebrafish RMS model was generated by injecting a Rag2-kRASG12D construct, expressing a constitutively active isoform of RAS under the Rag2 promoter, into the zebrafish embryo at one cell stage. Visible highly invasive tumors in liver, intestine, kidney, and testes appeared already 10 days post injection, in accordance with the early onset of these tumors in human patients, distinguishing this model as one of the fastest spontaneous, vertebrate tumor models available (Langenau et al, 2007).

11. Adult zebrafish models

While embryonic zebrafish models have yielded many valuable insights into human vascular or vessel-related pathologies, most diseases strike adult patients, and accordingly, their clinical symptoms are expected to be more closely recapitulated by adult disease models (Dahl Ejby Jensen et al, 2009). To this end there are today also a growing number of adult zebrafish models available which are highly valuable in the study of angiogenesis-related disorders including retinopathy, regeneration/wound healing and cancer.

11.1 Zebrafish retinal angiogenesis models

Diabetic retinopathy and age-related macular degeneration are severely debilitating disorders in which angiogenesis is a major driving force of the pathology. During progression of these diseases, retinal hypoxia induces pathological growth of immature and fragile blood vessels in the retina, which is associated with the progression to severe states of the diseases. Mice are not convenient to use for studies on retinal hypoxia, as the traditional methods of generating tissue hypoxia in mice (vessel occlusion/ligation) are not applicable in the retina, as there are no easily reachable arteries which can be ligated without causing excessive damage to the animal. Furthermore, mice cannot be exposed to severely hypoxic environments, as their respiration system is not adapted to withstand low atmospheric oxygen levels.

Zebrafish are much more robust to environmental hypoxia, and can withstand even very low oxygen levels in the water for a long time. In contrast to mammals, by incubating zebrafish in an aquarium, where the oxygen levels in the water can be tightly controlled by regulating the perfusion of nitrogen gas, researchers are given full control on the precise degree of hypoxia, the amount of time the tissue experiences hypoxia and the possibilities of distinguishing hypoxia-effects from other effects of restricted circulation such as acidosis and accumulation of waste products.

The zebrafish retinal vasculature is furthermore particularly amenable to studies on angiogenesis due to its remarkably simple structure (Fig. 16). As in mice and humans, the retinal vasculature is supplied with blood from a major optic artery, which branches out at the center of the optic disc to form approximately 4-7 so-called grade I arteries which cover the inner surface of the retina. These arteries branch a few more times before anastomosing with protrusions from the circumferential vein. Thus, the zebrafish retina has a simple monolayer vasculature that is organized from the center to the periphery as arteries-capillaries-veins, which is in contrast to that in mice where the vasculature is multi-layered and arteries, capillaries and veins are co-localized throughout the retina (R Cao et al, 2008).

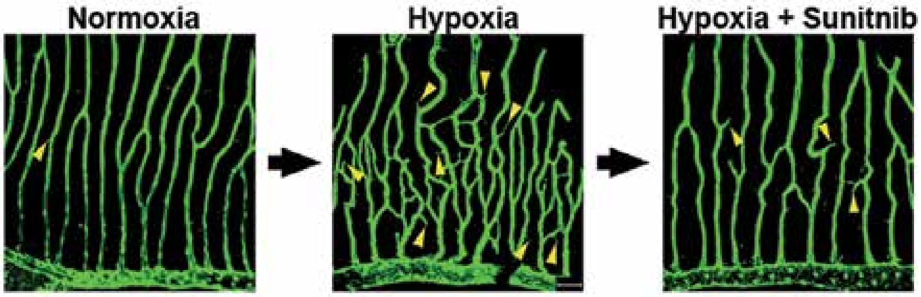


Fig. 16. Hypoxia-induced retinal angiogenesis model. Adult *Fli1:EGFP* transgenic zebrafish were incubated in 10% air-saturated water for 12 days, which induce marked angiogenic expansion of the capillary area of the retinal vasculature (middle figure) compared to fish in normoxic water (left figure). Hypoxia-induced retinal angiogenesis is dependent on VEGF signaling as the VEGF receptor inhibitor sunitnib is able to block the formation of new vessels under hypoxia in the zebrafish retina (right figure). Yellow arrowheads indicate angiogenic sprouts. Figure adapted from R Cao et al, 2008.

This retinal vasculature is highly sensitive to hypoxia. After only a few days of exposure to environmental hypoxia, a marked angiogenic expansion of the capillary area is clearly observed, which can be modulated either by the degree, or the amount of time, the fish is kept in hypoxia. An example of hypoxia-induced retinal angiogenesis in the adult zebrafish is given in Fig. 16.

In this model, adult zebrafish are put into a specifically designed aquarium where the water oxygen levels can be automatically controlled and monitored using an oxygen sensor coupled to a valve which opens for perfusion of nitrogen gas when the oxygen levels are above a set level defined by the researcher. Zebrafish can withstand very low oxygen levels better when being gradually adapted, so by slowly reducing the oxygen concentration in the hypoxia aquarium over the course of 1-2 days, the fish can be maintained in water with an oxygen concentration of only 8-10% of that in fully air-saturated water (Z Cao et al, 2010). After exposure to this level of hypoxia for a few days, retinal neovascularization is readily detected post mortem, under a fluorescent microscope. Also in this assay, it is convenient to add orally active drugs to the water during exposure to hypoxia, to study how such drugs may interfere with hypoxia-induced retinal neovascularization. Hypoxia-induced newly formed vessels can easily be distinguished due to the normally very simple and organized vasculature of the retina in adult zebrafish. Quantifying the number of sprouts, new branches and vascularization area are convenient ways of comparing the effects of the investigated drugs compared to non-treated controls (Z Cao et al, 2010). While capillary sprouting begins on day 3 after exposure to hypoxia, formation of a neovascular network becomes obvious on day 6 after exposure to hypoxia (R Cao et al, 2008).

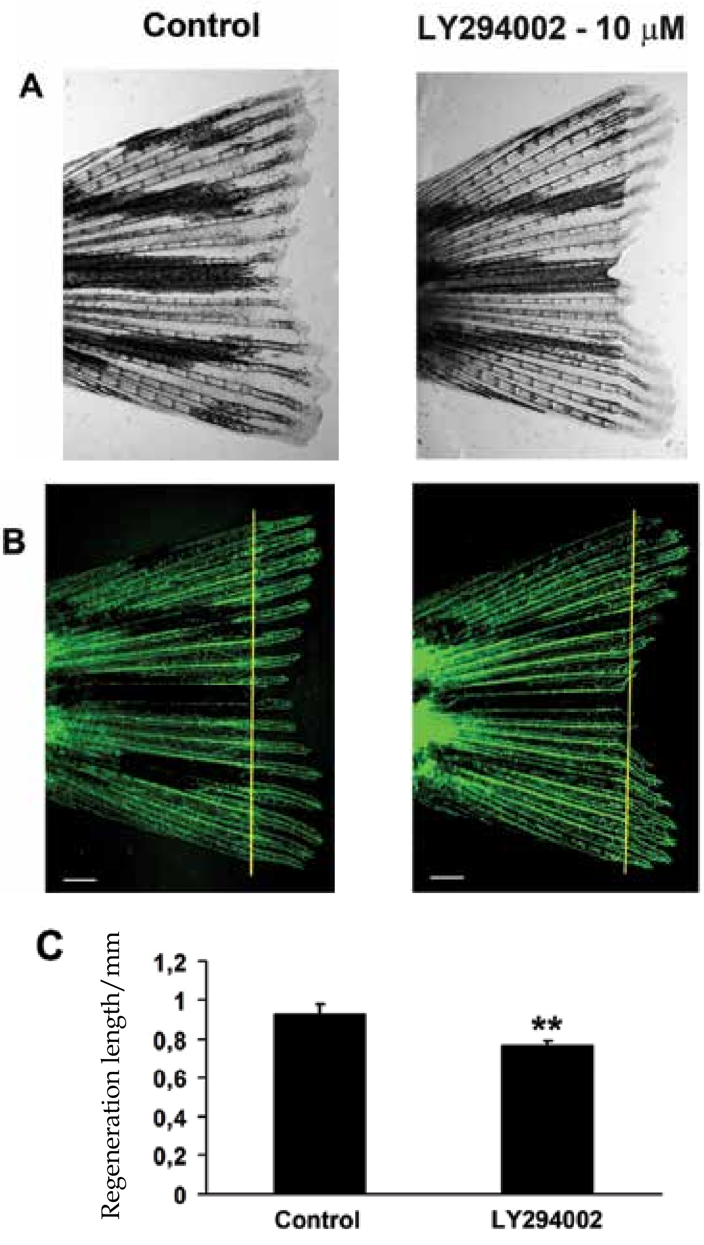


Fig. 17. Regenerating tail fin assay. The distal 1/3 of adult *Fli1:EGFP* transgenic zebrafish tail fins was amputated under anaestisia, and the fish were placed in either clean water (left panels) or water containing the phosphatidyl-inositol-3-kinase (PI3K) inhibitor LY294002 (right panels). After 6 days the fish were sacrificed, the tail fins were fixed in 4% PFA and the regeneration could be determined by bright-field microscopy (A). The vasculature in the regenerating tail fin tissue was observed by fluorescence microscopy (B). Quantification of the regeneration length is an easy and convenient way to demonstrate the effects of the tested drugs (C). Bars indicate 0.5 mm. Figure reproduced from Alvarez et al, 2009.

This experiment allows researchers to study the mechanisms of hypoxia-induced angiogenesis *in vivo*, to study neovascularization, vascular remodeling, and leakiness under pathological settings, to correlate vascular changes with disease development, to screen potential orally active therapeutic agents, and to assess the beneficial effects of known antiangiogenic agents for the treatment of retinopathy. However, there are also limitations. Unlike the clinical situation, this model does not induce local tissue hypoxia but rather rely on systemic exposure to hypoxia. Also, owing to the limitation of available antibodies against zebrafish proteins, staining of particular cell subsets in the retina may be difficult, reducing the amount of information that can be obtained from histological examinations compared to mouse models.

11.2 Angiogenesis in the regenerating zebrafish fin

Adult zebrafish have a remarkable regenerative capability. Many tissues which may not be regenerated in mammals are quickly regenerated in zebrafish. Among these are the heart, retina, maxillary barbell and fins (Poss et al, 2002; Vihtelic & Hyde, 2000; Alvarez et al, 2009). Importantly, as they regenerate, new blood and lymph vessels grow into the regenerating tissue – which enables studies on regenerative angiogenesis. One commonly used assay in the adult zebrafish, based on this principle is the regenerating tail fin. After amputation, the tail fin will re-grow and after approximately 1 month, the fin is back to its original size (Huang et al, 2003). This process of fin regeneration encompasses many of the same mechanisms as in human wound healing and regeneration, and is therefore a good model of regenerative angiogenesis. As the fin is largely transparent, and as morpholinos can be introduced by microinjection and electroporation, this assay is almost as versatile as the zebrafish developmental angiogenesis assays, the major difference being that it is performed in an adult animal. Also in the fins, the vasculature is remarkably simple (Fig. 17), and as the fin grows back various levels of vascular remodeling can be observed and therefore studied in detail (Huang et al, 2003). This regeneration model is probably the most commonly used, adult zebrafish angiogenesis model, and is considered to be complementary to the developmental angiogenesis models.

12. References

Akhtar N., Dickerson EB. & Auerbach R. (2002). The Spongue/Matrigel Angiogenesis Assay. *Angiogenesis*, 5, 1-2, pp. 75-80

Alders M., Hogan BM., Gjini E., et al. (2009). Mutations in CCBE1 cause generalized lymph vessel dysplasia in humans. *Nat Genet.*, 41, 12, pp. 1272-4

Alitalo K., Tammela T. & Petrova TV. (2005). Lymphangiogenesis in development and human disease. *Nature*, 438, 7070, pp. 946-53

Alvarez Y., Astudillo O., Jensen L., et al. (2009). Selective inhibition of retinal angiogenesis by targteting PI3 kinase. *PLoS ONE*, 4, 11, pp. e7867

Arbiser JL., Paningrathy D., Klauber N., et al., (1999). The antiangiogenic agents TNP-470 and 2-methoxyestradiol inhibit the growth of angiosarcoma in mice. *J Am Acad Dermatol*. 40, 6 Pt 1, pp. 925-9.

Berghmans S., Murphey RD., Weinholds E., et al. (2005). Tp53 mutant zebrafish develop malignant peripheral nerve sheath tumors. *Proc Natl Acad Sci U S A*, 102, 2, pp. 407-12

Brakenhielm E., Cao R., Gao B., et al. (2004). Angiogenesis inhibitor, TNP-470, prevents diet-induced and genetic obesity in mice. *Circ Res*, 94, 12, pp. 1579-88.

Brakenhielm E. & Cao Y. (2008). Angiogenesis in adipose tissue. *Methods Mol Biol*, 456, pp. 65-81

Cannon B. & Nedergaard J. (2004). Brown adipose tissue: function and physiological significance. *Physiol Rev*, 84, 1, pp. 277-359.

Cao R., Brakenhielm E., Pawliuk R., et al. (2003). Angiogenic synergism, vascular stability and improvement of hind-limb ischemia by a combination of PDGF-BB and FGF-2. *Nat. Med*, 9, 5, pp. 604-13

Cao R., Bjorndahl M., Religa, P., et al. (2004). PDGF-BB induces intratumoral lymphangiogenesis and promotes lymphatic metastasis. *Cancer Cell*. 6, 4, pp. 333-45

Cao R., Eriksson E., Kubo H., et al. (2004). Comparative evaluation of FGF-2-, VEGF-A-, and VEGF-C-induced angiogenesis, lymphangiogenesis, vascular fenestrations, and permeability. *Circ. Res*, 94, 5, pp. 664-70

Cao R., Jensen LD, Soll I., et al. (2008). Hypoxia-induced retinal angiogenesis in zebrafish as a model to study retinopathy. *PLoS ONE*, 3, 7, pp. e2748

Cao Y. (2005). Emerging mechanisms of tumour lymphangiogenesis and lymphatic metastasis. *Nat. Rev. Cancer*, 5, 9, pp. 735-43

Cao Y., Hong A., Schulten H., et al. (2005). Update on therapeutic neovascularization. *Cardiovasc Res*, 65, 3, pp. 639-48

Cao Y. (2007). Angiogenesis modulates adipogenesis and obesity. *J Clin Invest*, 117, 9, pp. 2362-8.

Cao Y. (2008). Molecular mechanisms and therapeutic development of angiogenesis inhibitors. *Adv. Cancer Res*, 100, pp. 113-31

Cao Y. (2008). Why and how do tumors stimulate lymphangiogenesis?. *Lymphat Res Biol*, 6 ,3-4, pp. 145-8

Cao Y. (2010). Therapeutic angiogenesis for ischemic disorders: what is missing for clinical benefits?. *Discov. Med*, 9, 46, pp. 179-84

Cao, Y. (2010). Adipose tissue angiogenesis as a therapeutic target for obesity and metabolic diseases. *Nat Rev Drug Discov*, 9, 2, pp. 107-15.

Cao Z., Jensen LD., Rouhi P., et al. (2010). Hypoxia-induced retinopathy model in adult zebrafish. *Nat. Protoc*, 5, 12, pp. 1903-10

Carmeliet P. (2003). Angiogenesis in health and disease. *Nat. Med*, 9, 6, pp. 653-60

Carmeliet P., Ferreira V., Breier G., et al. (1996). Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. *Nature*, 380, 6573, pp. 435-9

Cypess AM., Lehman S., Williams G., et al. (2009). Identification and importance of brown adipose tissue in adult humans. *New Engl. J Med*, 360, 15, pp. 1509-17

Dahl Ejby Jensen L., Cao R., Hedlund EM., et al. (2009). Nitric oxide permits hypoxia-induced lyphatic perfusion by controlling arterial-lymphatic conduits in zebrafish and glass catfish. *Proc Natl. Acad. Sci. U S A*, 106, 43, pp. 18408-13

Davis GE. & Senger DR. (2005). Endothelial extracellular matrix biosynthesis, remodeling, and functions during vascular morphogenesis and neovessel stabilization. *Circ. Res*, 97, 11, pp. 1093-107

Eriksson A., Cao R., Pawliuk R., et al. (2002). Placenta growth factor-1 antagonizes VEGF-induced angiogenesis and tumor growth by the formation of functionally inactive PlGF-1/VEGF heterodimers. *Cancer Cell*, 1, 1, pp- 99-108

Folkman J. (1971). Tumor angiogenesis: therapeutic implications. *N. Engl. J. Med,* 285, 21, pp. 1182-6

Folkman J. (1995). Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nat Med,* 1, 1, pp. 27-31.

Frederich RC., Löllmann B., Hamann A., et al. (1995). Expression of ob mRNA and its encoded protein in rodents. Impact of nutrition and obesity. *J Clin Invest*, 96, 3. pp. 1658-63.

Friedman JM. & Halaas JL. (1998). Leptin and the regulation of body weight in mammals. *Nature*, 395, 6704, pp. 763-70.

Hanahan D. (1985). Heritable formation of pancreatic beta-cell tumors in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. *Nature*, 15, 6015, pp. 115-22

Hedlund EM., Hosaka K., Zhong Z., et al. (2009). Malignant cell-derived PlGF promotes normalization and remodeling of the tumor vasculature. *Proc. Natl. Acad. Sci. U S A*, 106, 41, pp. 17505-10

Herbert SP., Huisken J., Kim TN., et al. (2009). Arterial-venous segregation by selective cell sprouting: an alternative mode of blood vessel formation. *Science*, 326, 5950, pp. 294-8

Huang CC., Lawson ND., Weinstein BM., et al. (2003). Reg6 is required for branching morphogenesis during blood vessel regeneration in zebrafish caudal fins. *Dev. Biol*, 264, 1, pp. 263-74

Hummel KP., Dickie MM & Coleman DL. (1966). Diabetes, a new mutation in the mouse. *Science*, 153, 740, pp. 1127-8.

Ingalls JM., Dickie MM. & Snell GD. (1950). Obese, a new mutation in the house mouse. *J. Hered.*, 41, 12 pp. 317-8

Jensen LD., Hansen AJ. & Lundbaek JA. (2007). Regulation of endotheilal cell migration by amphiphiles – are changes in cell membrane physical properties involved?, *Angiogenesis*, 10, 1, pp. 13-22

Jensen LD., Cao R., & Cao Y. (2009). In vivo angiogenesis and lymphangiogenesis models. *Curr Mol Med*, 9, 8, pp. 982-91

Koh YN., Kim HZ., Hwang SI., et al. (2010). Double antiangiogenic protein, DAAP, targeting VEGF-A and angiopoietins in tumor angiogenesis, metastasis, and vascular leakage. *Cancer Cell,* 18, 2, pp. 171-84

Kopelman PG. (2000). Obesity as a medical problem. *Nature*, 404, 6778, pp. 635-43.

Kukk E., Lymboussaki A., Taira S., et al. (1995). VEGF-C receptor binding and pattern of expression with VEGFR-3 suggests a role in lymphatic vascular development. *Development*, 122, 12, pp. 3829-37

Langenau DM., Traver D., Ferrando AA., et al. (2003). Myc-induced T cell leukemia in transgenic zebrafish. *Science,* 299, 5608, pp. 887-90

Langenau DM., Keefe MD., Storer NY., et al. (2007). Effects of RAS on the genesis of embryonal rhabdomyosarcoma. *Genes Dev*, 21, 11, pp. 1382-95

Lawson ND. & Weinstein BM. (2002). In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev Biol*, 248, 2, pp. 307-18

Lee SL., Rouhi P., Dahl Jensen L., et al. (2009). Hypoxia-induced pathological angiogenesis mediates tumor cell dissemination, invasion, and metastasis in a zebrafish tumor model. *Proc. Natl. Acad. Sci. U S A,* 106, 46, pp. 19485-90

Iida A., Sakaguchi K., Sato K., et al. (2010). Metalloprotease-dependent onset of blood circulation in zebrafish. *Curr Biol*, 20, 12, pp 1110-6

Lundberg G., Luo F., Blegen H., et al. (2002). A rat model for severe limb ischemia at rest. *Eur. Surg. Res.*, 35, 5, pp. 430-8

Nasevicius A & Ekker SC. (2000). Effective targeted gene ”knockdown” in zebrafish. *Nat. Genet*, 26, 2, pp. 216-20

Nasevicius A., Larson J. & Ekker SC. (2000). Distinct requirements for zebrafish angiogenesis revealsed by a VEGF-A morphant. *Yeast*, 17, 4, pp. 294-301

Neels JG., Thinnes T. & Loskutoff DJ. (2004). Angiogenesis in an in vivo model of adipose tissue development. *FASEB J*, 18, 9, pp. 983-5.

Patton EE., Widlund HR., Kutok JL., et al. (2005). BRAF mutations are sufficient to promote nevi formation and cooperate with p53 in the genesis of melanoma. *Curr. Biol*, 15, 3 pp. 249-54

Peterson RT., Shaw SY., Peterson TA., et al. (2004). Chemical suppression of a genetic mutation in a zebrafish model of aortic coarctation. *Nat Biotechnol*, 22, 5, pp. 595-9

Poss KD., Wilson LG. & Keating MT. (2002). Heart regeneration in zebrafish. *Science*, 298, 5601, pp. 2188-90

Risau W. (1997). Mechanisms of angiogenesis. *Nature*, 386, 6626, pp. 671-4

Rouhi P., Lee SL., Cao Z., et al. (2010). Pathological angiogenesis facilitates tumor cell dissemination and metastasis. *Cell Cycle*, 9, 5, pp. 913-7

Rouhi P., Jensen LD., Cao Z., et al. (2010). Hypoxia-induced metastasis model in embryonic zebrafish. *Nat Protoc*, 5, 12, pp. 1911-8

Sclafani A. & Springer D. (1976). Dietary obesity in adult rats: similarities to hypothalamic and human obesity syndromes. *Physiol Behav*, 17, 3. pp. 461-71.

van Marken Lichtenbelt WD., Vanhommerig JW., Smulders NM., et al. (2009). Cold-activated brown adipose tissue in healthy men. *N Engl J Med*, 360, 15, pp. 1500-8.

Vihtelic TS & Hyde DR. (2000). Light-induced rod and cone cell death and regeneration in the adult albino zebrafish (Danio rerio) retina. *J. Neurobiol*, 44, 3, pp. 289-307

Virtanen KA., Lidell ME., Orava J., et al. (2009). Functional brown adipose tissue in healthy adults. *N Engl J Med*, 360, 15, pp. 1518-25.

Weinstein BM., Stemple DL., Driever W., et al. (1995). Gridlock, a localized heritable vascular patterning defect in the zebrafish. *Nat. Med*, 1, 11, pp. 1143-7

Winzell MS. & Ahren B. (2004). The high-fat diet-fed mouse: a model for studying mechanisms and treatment of impaired glucose tolerance and type 2 diabetes. *Diabetes*, 53, 3, pp. S215-9

Xue Y., Religa P., Cao R., et al. (2008). Anti-VEGF agent confer survival advantages to tumor-bearing mice by improving cancer-associated systemic syndrome. *Proc. Natl. Acad. Sci. U S A*, 105, 47, pp. 18513-8

Xue, Y. Petrovic N., Cao R., et al. (2009). Hypoxia-independent angiogenesis in adipose tissues during cold acclimation. *Cell Metab*, 9, 1, pp. 99-109.

Yaniv K., Isogai S., Castranova D., et al. (2006). Live imaging of lymphatic development in the zebrafish. *Nat. Med*, 12, 6, pp. 711-6

31

Ethical and Legal Considerations in Human Biobanking: Experience of the Infectious Diseases BioBank at King's College London, UK

Zisis Kozlakidis[1,3], Robert J. S. Cason[2], Christine Mant[1] and John Cason[1,3]
[1]Department of Infectious Diseases, King's College London,
2nd Floor Borough Wing, Guy's Hospital,
[2]School of Law, Birkbeck College London,
[3]National Institute of Health Research's (NIHR) comprehensive Biomedical Research Centre (cBRC) at Guy's and St Thomas' NHS Foundation Trust,
UK

1. Introduction

Since the dawn of time *Homo sapiens* have collected human body-parts for a variety of reasons (Lassila &Branch, 2006; Aquaron *et al.*, 2009; Daily Telegraph, 2011). Similarly, representations of pathological lesions have been collected for educational purposes for at least three hundred years (*e.g.* the Hunterian Museum in Glasgow has preserved plaster casts of diseased tissues). A biobank is a generic term to describe any collection of biological materials and may take many forms, ranging from the preservation of plant seeds (*e.g* the Svalbard Global Seed Vault, Norway) or, the storage of human materials for transplants (*e.g.* corneal biobanks). Others collect human materials for artificial insemination (sperm, eggs and embryos), for forensic investigations and animal materials to assist in the preservation of endangered species such as the Iberian lynx (Leon-Quinto *et al.*, 2009). Some biobanks only collect a single type of sample such as DNA (genebanks), whilst others archive a wide variety of clinical materials. Until recently the *modus operandi* of most medical researchers was to use fresh clinical materials to test a specific hypothesis. The premise was either proven, or not, and then the process repeated to answer subsequent questions. This approach is incredibly wasteful since materials not directly needed to test each argument were discarded. In contrast, clinical biobanks can archive and distribute complete sets of materials from patients with diseases to multiple researchers thereby maximising the benefit of every donation. They can also revolutionise the understanding of very rare conditions by gradually accumulating sufficient numbers of samples –or, by the exchange of samples between multiple biobanks (networking) - to permit statistically-significant conclusions to be derived.

These advantages of biobanks were recognized by *Time* magazine as '*one of the ten ideas that are changing the world right now*' (Park, 2009). This Chapter will be confined to those issues confronting biobanks which collect human materials for medical research. Such archives can be subdivided into those which have the aim of answering one specific research question (*e.g.* the Multiple Sclerosis Brain bank) as opposed to systematic biobanks such as the Infectious Diseases Biobank (IDB) at King's College London (KCL) (Williams *et al.*, 2009), which collects clinical materials with no specific research question in mind. The growing popularity of biobanks in medical research in recent years has inevitably raised new and important ethic and legal questions regarding how they should be managed and regulated. For example, recently, the German Ethics Council has proposed that biobanks should be regulated on the basis of five 'pillars' including the concepts of: confidentiality; open informed consent; careful ethics review; sample quality-assurance; and, a transparency of the biobank's goals (Deutscher Ethikrat, 2010). Here some of the most contentious ethico-legal issues facing clinical archives are considered, including: (i) the nature of the contract (*i.e.* informed consent) between the subject and the researcher; (ii) the concept of property or ownership rights in respect to body tissues and fluids; (iii) the duty of care of a biobank to the donor, the sample, the researcher and, society. This is contextualised against historic turning points which have led to the regulatory structures currently in force in the UK. Finally, the organization of the UK's IDB at KCL is described and proposed as a model system for facilitating research into infectious diseases.

2. Ethical considerations in human biobanking

2.1 Novel challenges associated with biobanking

The idea of a biobank to facilitate medical research would appear to be a worthwhile and commendable activity to most people. However, the establishment of such archives raise not only many of the same ethical problems that face the medical community (particularly those involved in recruiting organ donations), but also some unique questions of their own, for example:

- How can a volunteer provide fully 'informed consent' when neither they nor the recruiter have any idea about the nature of future research which will be performed on the donated sample?
- Donating a sample for no pressing medical reason could be questioned since this relatively benign procedure carries an appreciable risk of adverse events (in one study the rate was 0.59% of 89,000 blood donations: of these ~15% were haematomas and 77% vaso-vagal reflex: Garozzo *et al.*, 2010).
- Who actually owns the donated sample?
- Does a donated sample have a commercial value and can it be sold?
- What happens if research discovers that a volunteer has a potentially deadly disease?
- How can biobanks insure that they are representative of the local community?
- Should biobank samples be used for 'trivial' (*e.g.* cosmetics development) or 'controversial' (*e.g.* stem cell, cloning *etc.*) research?

2.2 An ethical framework

To answer such questions biobanks (and their regulators) must draw on contemporary ethical codes, attitudes and opinions to provide guidance to best practice. Whilst this approach can provide discussion points to it does not always produce definitive answers (Gillon, 1985). The earliest consideration of medical ethics was probably the Hippocratic Oath, which introduced the concepts of respecting patients as individuals and doing no harm (Farnell, 2004). Similar sentiments are expressed in the payer of Maimonides, originally believed to have been written by the 12th-century physician-philosopher (Friedenwald, 1917). More probably, this prayer was written by M. Herz, a physician and pupil of the Konigsberg philosopher Immanuel Kant, as print versions can only be traced back until 1793. More contemporary views on medical ethics were crystallized in a 1902 book by Dr Albert Moll on 'Ethics of the Physician:' Hahn, 1984). Two major ethical issues raised by biobanks revolve around consent and the ownership in human tissues.

The justification for consent stems from the notion of personal sovereignty; the exclusive right an individual holds over their own person. This concept is historically rooted in liberal and political thought, as noted by JS Mills: '*over himself, over his own body and mind, the individual is sovereign*' (Mills 1972). Equally though, Kant believed that the humans '*exists as an end in itself, not merely as a means to be used by this or that will at its discretion*' (Gregor, 1998). Indeed, personal sovereignty now serves as the justification for the majority of articles enshrined in the Universal Declaration of Human Rights. Whilst consent is a necessary component in everyday life and medical research it is debatable how 'informed' consent need be. On the one extreme consent procured through misleading information (or under duress) cannot be considered valid. At the other end of the scale an individual may be informed of the risks and side effects of a medical procedure, but is not expected to comprehend the full complexities of the issues. Consent is often reduced to a subtle paternalism in regards to the unequal position of patient-subject to the researcher, as well as addressing how 'informed' a *research* project can be.

Although respecting personal sovereignty is necessary, there are also ethical principles in favour of a communal duty to society. As the aim of a biobank is to facilitate medical research (which in turn will aid the development of future treatments for the general good of society), the question arises as to whether there is an obligation to assist such endeavours. As Mills states '*there are also many positive acts for the benefit of others that he may rightfully be compelled to perform: such as to give evidence in courts of justice; to bear his fair share in a common defence; or in any other joint work necessary to the interests of the society of which he enjoys the protection*' (Mills, 1972). Thus, respecting personal sovereignty does not negate the argument in favour of a public duty to assist such endeavours. Indeed, utilitarian arguments for the 'greatest good for the greatest number' (Bentham's *'felicific calculus'*: Mitchell, 1918) and Kant's transcendental deduction of a moral duty (Paton, 1948) may to some degree also imply an obligation to donate samples to a biobank.

Locke's concept of property is based on the premise that an individual owns the labour of their body, which when mixed with something in nature, confers a property right in the produced object. Indeed a Lockean justification of property rights was accepted in the US case *John Moore v The Regents of the University of California* (1990) as a foundation for a claim on a human cell-line. Thus, Lockean justification for ownership of samples in a biobank could be constructed in a similar fashion; the labour expended in collecting, preparing and storing A biobank's samples confers a right of ownership. Although a degree of ownership exists in relation to human samples it is better to conceive this as conditional ownership (or 'custodianship') rather than an absolute ownership.

2.3 Some historical precedents leading to research ethics regulation

Self-regulation of biobanks based upon general ethical principles may seem a reasonable approach to managing a few samples of blood or urine which have been willingly donated for research. However, a series of notorious cases from the 19^{th} century up to the present day have so shocked the public that legislation of medical ethics and the storage of human body parts became inevitable: some of the most infamous cases are outlined below.

Body-snatching: The UK Murder Act of 1752 meant that the only legal source of corpses for anatomy was those of executed prisoners: however, this was insufficient to supply the demand from medical schools. Stealing a corpse was only regarded as a minor crime and thus evolved into a lucrative business. In 1827/8, the Edinburgh grave-robbers Burke and Hare realized that institutions paid more for fresh corpses and thus graduated from body-snatching to murder in order to meet this demand (Lancet, 1829; Howard & Smith, 2004). The subsequent conviction of Burke in 1829 led to the UK Anatomy Act of 1832 which stipulated that anyone practising anatomy must hold a licence and be responsible for the correct treatment of corpses. This act was repealed by the Anatomy act of 1984 which, in turn, was replaced by the Human Tissue Act of 2004 (below).

Genocide: In the 1930s/40s the National Socialist German Worker's party (NSDAP) became obsessed with the ideas of social Darwinism, eugenics and the Nietzsche concept of 'superman' (Taha, 2005). On this basis the regime initially justified killing those with congenital defects in the T4 (Tiergarten-4) euthenasia programme (Freidlander, 1995). This was criminal programme was subsequently extended to include anyone that the NSDAP deemed 'sub-human' (political opponents, Russian prisoners of war, and, notably the near genocides of European Jews and Roma: Bachrach, 2004). As part of this holocaust some victims were also subjected to non-consensual medical experiments (*e.g.* LD_{50} type testing of humans exposed to hypobaric or hypothermic conditions). Additionally, the NSDAP also assembled a collection of skeletons from euthanized prisoners for the Institute of Racial Hygiene to act as a historic record (and the basis for scientific study of) extinct human 'races'. At the end of the war the Nuremberg 'Doctors Trial' sentenced some of those responsible and resulted in the development of the Nuremberg code of practice for research involving humans (Table 1: US Government Printing Office, 1949). This is an important document since it has served as the basis of almost all subsequent refinements in medical ethics such as the most recent version of the Declaration of Helsinki (World Medical Association [WMA], 2009).

Unit 731: A less-well publicised series of medical crimes of the Second World War included those perpetrated by the Imperial Japanese Army's Unit 731. This was euphemistically named the 'Epidemic Prevention and Water Purification Department' of the Kwantung Army Group in Harbin, occupied China (Alibewk & Handelman, 1999). This unit experimented on over 10,000 humans in studies involving conscious, non-anaesthetised, *vivisections*, weapons testing (*e.g.* the effects of flamethrowers, hand grenades *etc.* upon live humans), as well as bio-weapons research (Harris, 1994; Kristof, 1995; Barrenblat, 2004).

Tuskagee syphilis study: A study of 400 poor African-American men with syphilis was initiated in 1932. To induce participation, recruits were given free medical care, meals and burials and in return provided samples of blood and cerebro-spinal fluids to researchers (Roy, 1995; Crenner, 2011). At no point were the recruits informed that they had syphilis, nor were they treated for it. The 40-year study was particularly controversial because the researchers failed to treat patients even after the discovery that penicillin was an effective cure. In 2010 it was subsequently revealed that in Guatemala the same study had been extended, between 1946-1948, to include actually infecting prisoners, soldiers, and patients in a mental hospital. A total of 696 men and women were exposed to syphilis without their informed consent. As a direct result of these revelations the US Congress passed the National Research Act in 1974 and created a commission to study construct regulations governing studies which involve human participants (Prograis, 2010).

1. The voluntary consent of the human subject is absolutely essential. This means that the person involved should have legal capacity to give consent; should be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, over-reaching, or other ulterior form of coercion; and should have sufficient knowledge and comprehension of the elements of the subject matter involved as to enable him/her to make an understanding and enlightened decision. This latter element requires that before the acceptance of an affirmative decision by the experimental subject there should be made known to him the nature, duration, and purpose of the experiment; the method and means by which it is to be conducted; all inconveniences and hazards reasonable expected; and the effects upon his health which may possibly arise from participation. The duty and responsibility for ascertaining the quality of the consent rests upon each individual who initiates, directs or engages in the experiment. It is a personal duty and responsibility which may not be delegated to another with impunity.

2. The experiment should be such as to yield fruitful results for the good of society, unprocurable by other methods or means of study, and not random and unnecessary in nature.

3. The experiment should be so designed and based on the results of animal experimentation and a knowledge of the natural history of the disease or other problem under study that the anticipated results will justify the performance of the experiment.

4. The experiment should be so conducted as to avoid all unnecessary physical and mental suffering and injury.

5. No experiment should be conducted where there is a prior reason to believe that death or disabling injury will occur; except, perhaps, in those experiments where the experimental physicians also serve as subjects.

6. The degree of risk to be taken should never exceed that determined by the humanitarian importance of the problem to be solved by the experiment.

7. Proper preparations should be made and adequate facilities provided to protect the experimental subject against even remote possibilities of injury, disability, or death.

8. The experiment should be conducted only by scientifically qualified persons. The highest degree of skill and care should be required through all stages of the experiment of those who conduct or engage in the experiment.

9. During the course of the experiment the human subject should be at liberty to bring the experiment to an end if he has reached the physical or mental state where continuation of the experiment seems to him to be impossible.

10. During the course of the experiment the scientist in charge must be prepared to terminate the experiment at any stage, if he has probable cause to believe, in the exercise of the good faith, superior skill and careful judgment required of him that a continuation of the experiment is likely to result in injury, disability, or death to the experimental subject.

Table 1. The Nuremberg code for medical research involving humans.

Alder Hey hospital scandal: An investigation into the retention of hearts at hospitals in Bristol UK in the early 1990s led to a public inquiry. This subsequently found that a large number of hearts were also being held by the Alder Hey Children's Hospital and the Walton Hospital. In 2001 the Redfern Report (Royal Liverpool Hospital Children's enquiry, 2001) was published and this led to public outcry when it was revealed that Prof van Velzen had archived organs from every child subjected to a *post mortem.* Around 500,000 tissue samples were being held without any realistic likelihood of them ever being used for research. These revelations led to the creation of the Human Tissue Authority and the 2004 Human Tissue Act and in the UK.

Desecration of Alaister Cooke's remains: In 2005 it was discovered that the bones of Alaister Cooke (a distinguished BBC correspondent) and those of others had been surgically excised without permission prior to cremation by Biomedical Tissue Services Ltd. (Smit, 2008). The company then sold the treated bones for use as surgical grafts. Cooke was suffering from bone cancer when he died which would have made his tissues unsuitable for such a purpose. Reports revealed that the people involved in selling the bones altered his death certificate to hide this fact: subsequently M. Mastromarino, a former New Jersey dentist, was sentenced to between 18 and 54 years imprisonment.

The more contemporary of these cases illustrate that body-snatching is a practice which is not restricted to the dark days of the 19th century and will no doubt continue in illicit markets for the foreseeable future. Common themes linking all of these examples include: a dereliction of basic medical responsibilities by physicians; lack of compassion; complete disregard of the dignity and autonomy of the participants (and/or that of the relatives of the deceased); the *storage of body parts*; and, *the absence of consent* by the participants.

2.4 Informed consent and tissue banks

A fundamental requirement of contemporary medical ethics is that of 'informed consent' be provided by a participant before any study, or procedure, can be performed as discussed above (2.2). However, the phrase is fundamentally misleading (Kaye, 2004), since it implies a comprehension of the relevant facts and all possible outcomes of the research. However, how can non-medically qualified members of the public truly be considered to be fully 'informed'? Indeed, by definition the researchers themselves can only best-guess the possibilities (*'if we knew what we were doing, it wouldn't be called research, would it?'* Albert Einstein). This situation is exacerbated in the case of biobanks where samples may be used in future research projects that have not yet been envisioned using techniques and technologies which have yet to be developed. Indeed, one study of biobank donors found that they did not consider themselves well informed about what their samples would be used for (Hoeyer *et al.*, 2005).

This issue was been addressed by the German Ethics Council which takes the view that: *'if donors have been informed of the indefinite nature of the actual future applications, they will be aware that they are agreeing to an uncertainty. This uncertainty is not acceptable if it involves more than minimum health risk which is not the case with Biobanks'* (Deutscher Ethikrat, 2010). Thus 'open' or 'broad' consent to future usage of donated samples has been proposed as best practice for biobanks (Hansson *et al.*, 2006). Similarly, the council of Europe's biobanking recommendation acknowledges the conflict between the traditional informed consent and the needs of population genetic databases, and as a result stated that consent need not be specific, but it must be as specific as possible with regard to unforeseen uses (Council of Europe, 2006).

Such proposals are not without their critics who see equivalence between the broadening of consent and the dilution of ethics which may result in increased public distrust (Hofmann, 2009). Though others have noted that actually the reverse may be true (Lipworth *et al.*, 2009). Some commentators have gone further, suggesting that for genebanks and for population databases, informed consent should be abandoned altogether (Kaye, 2004). Furthermore, in countries such as the UK where free healthcare is provided to all by the state, there is debate as to whether there is an automatic moral obligation upon patients to automatically donate any excess clinical material taken for diagnostic purposes to medical research: *i.e.* the introduction of an 'opt-out' as opposed to the current 'opt-in' system. Such 'opt-out' genebanks are already in operation in Europe (*e.g.* the Vanderbilt DNA databank) and have driven the development of new approaches to the governance as well as innovative public education and communications strategies (Pulley *et al.*, 2010).

2.5 Inclusivity of biobanks

There are many problems facing researchers in gaining the public's confidence in donating samples to biobanks. In the case of donations to genebanks public refusal to consent (revealed by a questionnaire) was explained by a lack of personal relevance of the contribution and feelings of discomfort related to the possibility that the DNA would be used for purposes other than the original study (Melas *et al.*, 2010). The underlying concerns revolved around issues of integrity, privacy, suspiciousness, and insecurity. Interestingly though despite concerns about privacy another study of 4,569 US participants revealed that 60% would be willing to participate (Kafam *et al.*, 2009). However, the same study noted that ethnic minorities, women and those without a College degree, were concerned that the government could gain access to their personal information. Such concerns may be translated into an unwillingness not to engage with biobanks. Indeed, there are well acknowledged problems in recruiting sufficient organ donations from ethnic minorities resulting in a higher mortality amongst these communities from diseases necessitating transplants (Bratton *et al.*, 2011; Salim *et al.*, 2010). In the USA educational schemes have been introduced and appear to have partially resolved this problem (Callender *et al.*, 2010).

In industrialized nations the ethical points of reference for regulating medical research have inevitably been drawn from classical western moral philosophies and Judeo-Christian religious traditions. Relatively recently, the UK population has transformed from a predominantly Caucasian European Christian admix into a diverse multi-cultural/ethnic society as a result of immigration. Whilst more recent migrants from Eastern Europe share many of the cultural and religious traditions of the former, others from Africa, the Indian sub-continent and Asia often do not. Research biobanks (like organ donation schemes) need to be representative of their communities, consequently they must: (i) appreciate the cultural and religious sensitivities of ethnic minorities; (ii) understand historic negative perceptions of Western medical research, and, (iii) use this information to insure that ethnic minorities become fully engaged in such research projects.

An example of such cultural differences include Chinese tradition where self-determination is not a recognized phenomenon (Bowman & Hui, 2010), meaning that the family –rather than the patient- receive clinical information and make decisions to coordinate treatment. In Judaism, bodies are buried undisturbed and quickly after death as a matter of respect. Discussion within the Jewish faith about whether it is permissible to harvest organs for transplant from brain-dead persons is ongoing (Bresnahan & Mahler, 2000). Similarly, amongst Hindus and Sikhs the individual is caste-bound in its decision, and there is the concept of purification by death/rebirth axis: organ donation or contribution to a biobank might be seen to interfere with. In a systematic review of the opinions of twenty-eight major religions only one, Shintoism, was noted to be completely antagonistic to the idea of organ or tissue donation after death (United Network for Organ Sharing [UNOS], 2000). This is based on the concept that the cadaver is impure and dangerous, and injuring it is a serious crime as it damages the '*itai*' (the relationship between the dead and the bereaved).

Many ethnic minorities resident in industrialized nations also have perceptions of 'scientific imperialism' or 'bio-colonialism' (Emerson *et al.*, 2011) or scientific racism (*e.g.* the Tuskegee study 2.3 above). Not surprisingly there can be considerable distrust of western medical research. A recent example of this was the reticence of the Indonesian government to share samples of the H5N1 influenza virus with the international scientific community (Gelling, 2007). One suggestion to restore public faith is the proposal to establish a tissue trust to serve the interests of the common good (Emerson *et al.*, 2011) and would act by involving tissue donors and community members in research governance. These issues are of considerable importance to the current 'opt-in' model for biobanking in the UK. Indeed, for biobanks to be effective they must collect tissues from all of the community. Failure to do this may mean that downstream medical research using a biobanks samples may effectively result in further examples of scientific racism in that some research may be race-specific. These concerns are also a persuasive argument for locating biobanks in ethnically-diverse regions of a country.

Extending this concept of inclusivity further is the idea of harmonizing legal and ethical permissions internationally. This is important as a major ambition of biobanks world-wide is to establish networks for the international exchange of important clinical samples (Pearson, 2004). Whilst differences in national laws may complicate this process it has been suggested that if all countries simply abide by the Helsinki Declaration (WMA, 2009) any additional regulation would be counterproductive (Hansson, 2011). In contrast, others have argued in favour of a greater harmonization of ethics legislation between nations (Chalmers, 2011). Harmonization of biobank regulation is an important future goal since a survey of 126 European biobanks noted that most had currently only a very limited networking activity, and just a half having policies for cross-border sharing of samples (Zika *et al.*, 2011).

3. UK regulations and statutory bodies

In the UK there are three major governmental bodies which regulate medical research. Regulation of research ethics is by one of two types of review bodies: universities (*e.g.* the King's College London's College Research Ethics Committee) and the government's National Research Ethics Service (NRES, 2007). Whilst there is some overlap between these two bodies (*e.g.* human studies not involving NHS patients can be considered by both, investigations using NHS patient samples can only be considered by the latter), most medical researchers use the NRES's local research ethics committees (LREC) scheme. LRECs have evolved over recent years so that now specialized committees exist which are trained in issues arising from biobanking.

Storage of human tissues is regulated by a different body, the Human Tissue Authority (HTA) and premises keeping human tissues for research are required to hold a specific type of HTA licence and are subject to periodic inspections. The definition of tissues by the HTA differs considerably from the biological meaning (a collection of the same kind of cells with a common structure and function: *e.g.* muscle, skin and bone). For the HTA, a tissue is considered to be a mixture of different cells acting with common purpose (*e.g.* such as cells of the immune system). Thus, blood is considered a tissue under the HTA act, though so too are faeces and urine since they also contain a mix of immunological cells. Conversely, a cell-line derived within a week of isolation from the body is not considered a tissue on the basis of its homogeneity. Similarly, hair – not containing cellular architecture - is not regarded by the HTA as a tissue. Confusingly, if tissue architecture architecture is immediately disrupted the resultant biochemical mix is not considered a tissue by the authority (*e.g.* DNA extracted from a human tissue).

Anyone who collects, stores, uses, discloses or destroys identifiable personal information about living individuals, must also comply with the UK's 1998 Data Protection Act (DPA) and the Common Law duty of confidence. For the deceased, researchers must comply with the latter only. Anonymised personal information (as most frequently collected by biobanks) whether concerning the living or the deceased, falls outside the scope of these legal requirements. The DPA applies to `personal data´, which are data that relate to a living individual who can be identified either from those data alone or from those data taken in conjunction with other information that is available to the person who controls the data. When gathering identifiable personal information researchers should aim at all times to ensure that its processing is defensible as both `fair and lawful´. This requires as much transparency as possible about the uses to which data will be put and any risks that might be involved. The net effect of the DPA act is that centres which recruit and annoymise human data and clinical materials prior to submitting them to a biobank are subject to the provisions of the DPA, thus such personal information must be kept secure at all times. Dependent on the nature of the human samples being archived the Health and Safety Executive (HSE) will often need to be consulted by biobanks, particularly if this includes the use of clinical materials from patients infected with dangerous human pathogens (*e.g.* the IDB below).

4. The KCL infectious diseases biobank as a model system

4.1 Vision

Biological resource collections such as the Multicenter AIDS Cohort Study (Kingsley *et al.*, 1987) and the Sidney blood bank cohort (Oelrichs *et al.*, 1998) have helped drive important advances in the understanding of the pathogenesis of human immunodeficiency virus (HIV). In 2005 there were many biobanks in the UK dedicated to the collection of brains or cancer biopsies, but no equivalent facility for the collection of materials for HIV, or indeed any other infectious diseases, research. Even within Europe only three other biobanks held stocks of publicly-accessible material for HIV research, the Spanish HIV biobank (Garcia-Merino *et al.*, 2009,2010), the Sapienza University HIV biobank in Italy, and the Picardie biobank which holds only sera from infected subjects (Chaigneau *et al.*, 2007). A consultation exercise with researchers at KCL indicated that an infectious diseases tissue bank facility would be welcomed by many. This led to the establishment of a group of clinicians and scientists to develop what has now become the KCL IDB. The central issue of the IDB was to collect materials which are of significant value to researchers. For example, around this time many pathology departments were (and still are) rebranding themselves as biobanks to attract research funding. The problem with this approach is that such pathological collections are plentiful and the types of samples preserved do not always coincide with the requirements of researchers (*e.g.* materials suitable for molecular biology studies).

It was therefore established early on that the IDB would not be a genebank, but rather an archive of a broad range of clinical materials which would enable a spectrum of proteomic and genomic studies to be performed (*e.g.* containing live lymphocytes, RNA, DNA, plasma, sera, cerebro-spinal fluids *etc.*). These would be prepared and stored to a high-standard and fully documented in terms of sample tracking and processing details. The initial patient cohorts selected for study were those infected with pathogens that were of significance to the local community and also to local researchers. These were patients infected with HIV, hepatitis B virus or, with bacteraemia (especially methicillin-resistant *Staphyloccocus aureus*). In the case of the major sample collection from HIV-infected subjects it was further decided to selectively recruit those with particularly interesting clinical histories. For example, those initially recruited were HIV-1 clade B infected individuals who either progressed to disease unusually quickly (rapid progressors) or, very slowly (long-term non-progressors) as these extremes are most likely to yield important answers to the determinants of pathogenesis. Importantly, none of these patients were to be receiving medication so that the natural history of the infection and disease processes could be studied.

4.2 Location and setting

The IDB is uniquely located for purpose as the local population in Lambeth and Southwark is large (~4 million) and extremely diverse. Indeed, after English the second most common spoken language is Yaruba (African) and then Portuguese (Lambeth census, 2001). This community also suffers from some of the highest rates for HIV infections, the viral hepatitides and sexually-transmitted infections in Europe. The prevalence of UK HIV infections is highest within this area and over 10% of all UK HIV cases are treated by local clinics. This is exemplified by the facts that amongst pregnant women attending St Thomas' Hospital to deliver their babies around 1% are infected with HIV and 2% with hepatitis B (Health Protection Agency, 2008). The IDB is embedded within the KCL Department of Infectious Diseases which is affiliated to King's Health Partners and Guy's And St Thomas' NHS Foundation Trust. The latter hosts an Academic Health Science Centre (AHSC) and also an NHS National Institute of Health Research (NIHR) comprehensive BioMedical Research Centre (cBRC). The latter offer considerable advantages since it has established two clinical research facilities (CRFs) that effectively comprise of two wards and resources in which to conduct clinical trials.

4.3 Ethical permissions for the IDB

The prerequisites for the IDB's ethics included the concept of the dignity and autonomy of the volunteers yet also acknowledged the uncertain future research uses of biobank samples. Thus, the patient's information sheets and consent forms were designed to make it absolutely clear about the uncertainty of future usage. They also make it transparent that their samples would probably be used for genetic research and, the possibility that they would be used both for academic or commercial research purposes anywhere in the world. It is also made clear to participants that their healthcare would be unaffected by their decision to donate a sample, that they could withdraw from the biobank project at any time (and also demand that previously donated samples be destroyed) in line with most recommendations (Gertz, 2008) and that all samples would be annonymised. An additional safeguard was that should a downstream third-party researcher make a finding that was pertinent to the health care of the volunteer, this information would be passed back up through the management chain *via* the biobank to the clinicians at the tissue collection centre (TCC) who could then break the code and advise patients accordingly (since codes linking the patient's NHS number and the biobank code are only held at TCCs). These core principles were consequently remarkably similar to those proposed by the German Research Ethics Council some four years later (Deutscher Ethikrat, 2010).

The timing of the establishment of the IDB was far from ideal since the HTA act was just being implemented. Like any such legislation it is the interpretation which sets the precedents, a process which can take some time. Currently the IDB has ethical permission from the Southampton and South West Hampshire Research Ethics Committee (B) which extends until 2014 (reference # 09/H0504/39) to collect research samples (blood, urine and faeces) and (any) residual diagnostic samples from patients (adults, children and infants) with any infectious or inflammatory disease who are attending a routine clinical appointment.

The IDB cannot recruit from patients who are prisoners or those who are incapable of providing informed consent (other than children where the parent/guardian can consent for them). There are also restrictions for researchers and IDB samples cannot be used for 'trivial' or 'controversial' research projects such as those involving: fertility, reproduction, stem cells, cosmetics or animals. The IDB can establish TCCs at any NHS location in England, Wales or Northern Ireland with the co-operation with a local medical Consultant. Local NHS R&D

offices have to be informed that a TCC is being established but can play no other role in the process. The IDB governance committee is also enabled (through devolved ethics powers) to act as an LREC and provide ethical opinions upon studies wishing to access IDB samples. The IDB stores samples under the authority of an HTA research license held by the Guy's Hospital campus (reference # 12521). This not only covers the storage of materials by the IDB, but also those researchers who remove IDB samples to other sites for the duration of their ethical permission. A diagram of the IDBs operations is provided (Figure 1).

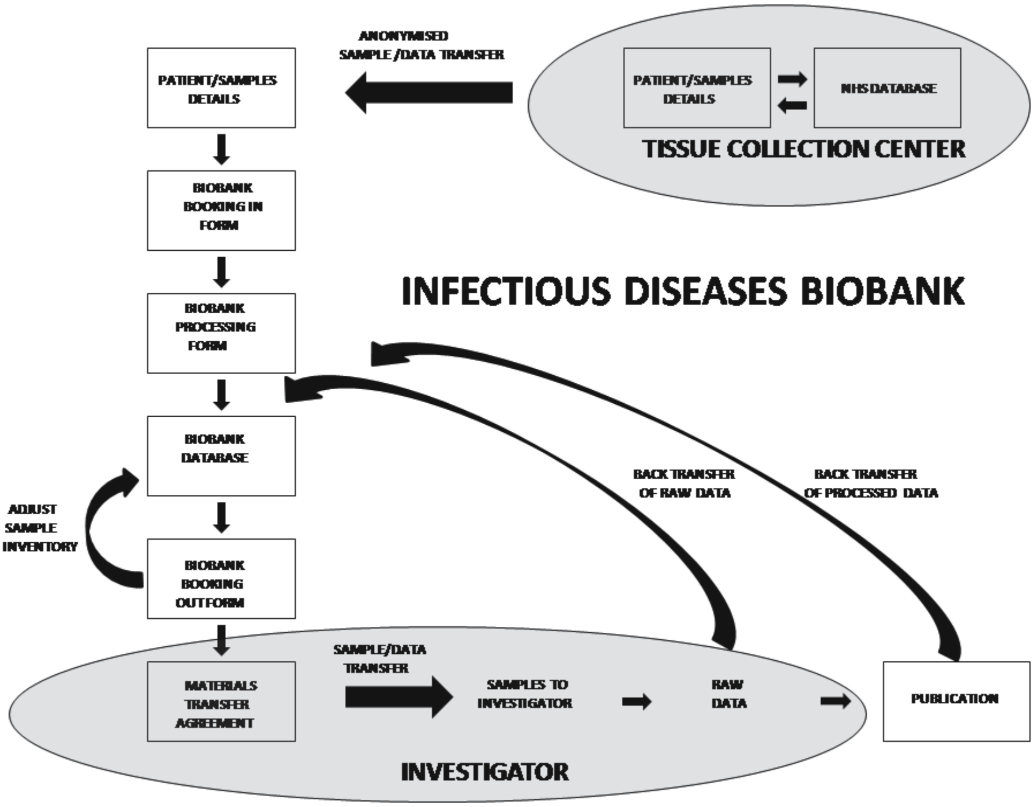

ANONYMISED
SAMPLE /DATA TRANSFER
PATIENT/SAMPLES DETAILS
PATIENT/SAMPLES DETAILS
NHS DATABASE
TISSUE COLLECTION CENTER
BIOBANK BOOKING IN FORM
INFECTIOUS DISEASES BIOBANK
BIOBANK PROCESSING FORM
BIOBANK DATABASE
BACK TRANSFER OF RAW DATA
BACK TRANSFER OF PROCESSED DATA
ADJUST SAMPLE INVENTORY
BIOBANK BOOKING OUTFORM
SAMPLE/DATA TRANSFER
MATERIALS TRANSFER AGREEMENT
SAMPLES TO INVESTIGATOR
RAW DATA
PUBLICATION
INVESTIGATOR

Fig. 1. Management of the IDB.

4.4 Governance

The IDB has a Governance Committee comprising of scientists, doctors, representatives of funding bodies and of the patients. This Committee is responsible for managing the IDB, strategy decisions, insuring that it conforms to current legislative requirements and also to local Medical school regulations. The IDB is regularly audited by the internal KCL Medical School representatives and by the HTA (~four times in 2010) and the results of these checks are passed back to the Governance Committee. Researchers wishing to access the IDB's samples submit a simple two-page application detailing what they propose to do and, the type and numbers of samples required. These details are scrutinized by the IDB's Governance Committee for scientific validity and also for any ethical dimensions. If successful, researchers sign a materials transfer agreement (MTA) and the samples are released.

Staff involved in recruiting volunteers must all have completed 'in house' courses on 'consent taking', 'good clinical practice' and phlebotomy. Copies of these certificates are held by the IDB. Members of the IDB staff are also encouraged to undertake an academic module in ethics, philosophy and religion (a three year 'Associate of KCL' course). The Governance Committee has also established a clear policy on charging researchers to access the IDB's material; they may either agree to pay a fixed rate for the individual samples or contribute funds towards the salaries of IDB staff, to offset the processing and storage costs incurred by the IDB. Researchers are encouraged to approach the IDB early on during the preparation of grants so that projected costs can be included in their applications.

4.5 Quality control

The IDB utilizes standardized operating procedures (SOPs) based upon EEC standards (ISO guideline 34,# 17025:2005) and works within the UNE-EN-ISO 9001:2000 guidelines to facilitate future inter-biobank networking capabilities. Samples are continuously tracked from the time of venepuncture, the time of courier collection from the TCCs, through to freezing at the IDB, with a target of processing >75% of samples within four hours of the bleed. All materials from patients with infections are processed in negative-pressure category III laboratory and stored in locked -80 °C freezers. All of the freezers are: on a protected hospital electricity supply; alarmed to the IDB's staff mobile telephones; and, checked daily for temperature fluctuations. None of the released samples from the collections have undergone a freeze-thaw cycle. Purified DNA samples are tested for the concentration of DNA and, by polymerase chain reaction (PCR) amplification of the housekeeping gene β-globin, for the absence of PCR inhibitors prior to release. In-house assessment of viral RNA viability has also revealed that viral RNA and sequences can be recovered from all plasmas of HIV patients so far tested (for those with viral loads of >350 copies *per* ml). Similarly, an independent analysis of human genomic RNA integrity has demonstrated that all of 104 samples were of high-quality (mean RNA integrity values of 9.3, on a scale where 1=degraded and 10=completely intact RNA) and were successfully used to generate DNA for transcriptome analysis (Kozlakidis *et al.*, 2011). Ultimately, the IDB aims to have all of its procedures, SOPs and operations validated by the International Standards Organisation.

In addition to merely maintaining the samples under the IDB's custiodianship we have also sought to enhance their research value. For example, for the core cohort of HIV infected patients approximately 33.3% of samples have been genotyped for their HLA class I and II alleles and their plasma viruses have had their Gag genes sequenced. The Gag region is important as protein products of this reading frame are believed to be important determinants in viral escape from the innate and adaptive arms of the immune system (Deml *et al.*, 2005). To date, several hundreds of full-length Gag genes have been cloned and sequenced (and the latter data deposited in Genebank: accession numbers FN597659-FN600533). These cloned Gag genes are available to researchers by arrangement. The intention is to obtain equivalent data for the complete HIV cohort. The other type of quality control that the IDB is actively involved in is monitoring that the samples being collected are those which are of (a) most clinical significance and (b) representative of the local community. Initial analyses of the first 200 HIV patient volunteers indicated that it was collecting a population greatly enriched for those with unusual rates of disease progression and that the ethnicity of the volunteers matches well with that of the local community (Kozlakidis *et al.*, In Press).

4.6 Transparancy

Given the multiplicity of studies that any individual sample may be used in, the Governance committee decided that logistically it would be impossible to provide individual volunteers with research feedback on their individual samples, despite reports that this is the preference of potential volunteers (Meulkamp *et al.*, 2010). However, the IDB does attempt to provide feedback in the types of studies performed and these data are displayed on the IDB's website (http://www.kcl.ac.uk/schools/medicine/research/diiid/centres /pii/ biobank/index.html). The IDB has also publicised its mission and was reviewed in *Nature Medicine* (Towie, 2006), has published in *Retrovirology* (Williams *et al.*, 2009), *Biopreservation and Biobanking* (Kozlakidis *et al.*, 2011) and has made presentations to national and international meetings (*e.g.* the 2010 Biobanking Conference; the European Virology Congress in 2008 & 2010, Nuremberg and Como). The IDB director has served on the feasibility study of Biobanking in Northern Ireland (NI) for the NI NHS R&D committee (2008), the KCL College Research ethics committee and associated sub-committees (2010-) and, is chair of the IDB Governance and ethics committees. The IDB has also consulted with patient representatives and has, as a consequence, increased its electronic footprint by establishing KCL IDB sites on the social networking sites 'Facebook' and 'Linkedin'.

4.7 Growth of the IDB

Since sample collection was initiated in January 2007 there has been a logarithmic growth in the number of patient visits to TCCs as well as the numbers of studies approved to access IDB samples (Figure 2). Indeed, currently the IDB is processing over three litres of peripheral venous blood *per* week. These examples of growth have also been matched by the expansion of the IDB into new categories of diseases. These now include those with hepatitis C virus or papillomavirus virus infections and patients with inflammatory diseases (including: as diabetes, Crohn's disease, ulcerative colitis, systemic lupus erythematosus, and, pre-multiple sclerosis syndrome). Some of the currently approved studies are listed (Table 2), and a steady stream of research publications is starting to ensue (Nath *et al.*, 2006, 2007; Alvarez *et al.*, 2008; Thorborn *et al.*, 2010). In addition, publications arising from the contract research work of the IDB are expected to increase significantly (4.8 below).

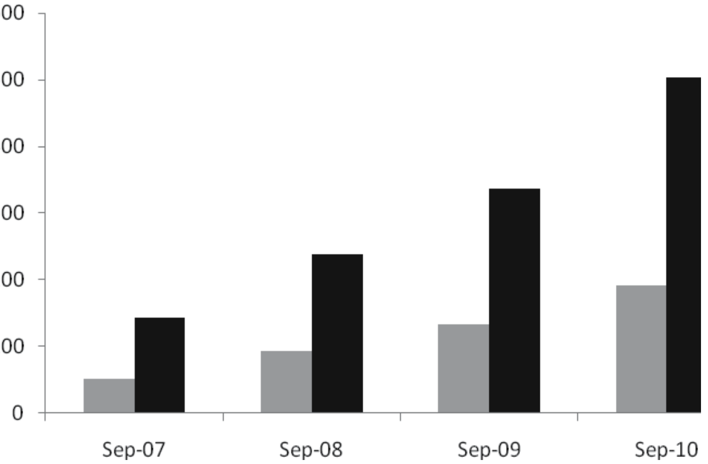


Fig. 2. Year-on-year growth of the IDB's HIV sample collection. Gray bars indicate the recruitment of new patients in total, whereas black bars indicate the number of donations per year.

T-Regulatory cells in HIV infection
Ps20 studies in HIV infections
The roles of vpu and tetherinin HIV/AIDS pathogenesis
The control of inflammation in immunity and autoimmunity
Naive B cell responses in older people
Gene expression signatures of HIV-1 infection *in vivo* and *in vitro*
ccess and study residual clinical samples made available during routine joint surgery.
Genetic variations in IL28B on the natural history of hepatitis B and C and their treatment response.
Non-infectious co-morbidities in HIV infection
Renal function and bone homeostasis in patients starting HAART
Defining the function of CD161+ CD8+ T cell subsets in HIV infection and their response to therapy
Investigating the effect of Maraviroc on Microbial Translocation in HIV infected individuals who are receiving antiretroviral therapy
The metabolic impact of Darunavir/ritonavir maintenance monotherapy after successful viral suppression with standard Atripla in HIV-1-infected patients

Table 2. Some of the types of studies currently accessing the IDB.

4.8 The IDB as a contract service

The IDB's skill and expertise is also being currently utilized to provide research support for clinical studies performed locally, these are often intervention studies and hence require independent ethical permissions from LRECs other than that of the IDB. The largest of these is currently the KCL Human Immune Response Dynamics (HIRD) study. This ground-breaking longitudinal study is investigating the response in humans to vaccination with the H1N1 ('swine flu') influenza vaccine using a protocol was approved by the Brent LREC (09/H0717/88). Briefly, the study involves the IDB collecting and archiving of peripheral venous blood samples from overnight fasted volunteers at the CRFs. Two pre-vaccination samples are harvested and, after vaccination (with PandremixTM H1N1 vaccine: GlaxoSmithKline Biologicals Ltd), a further four samples are collected until six weeks post vaccination. To date over 170 volunteers have completed the course of bleeds and vaccination.

5. Conclusions

This Chapter has summarized the development of some key ethical concepts which specifically impinge upon the newly emerging discipline of human tissue biobanking in terms of classical notions as well as historical turning points which have resulted in current UK legislation. In particular, this paper highlights the fundamental dilemma between the rights of the individual and their duty to the society and describes in practical terms how the KCL Infectious Diseases Biobank is regulated and managed. Whilst, the essential ethical keystone to any biobank is the principle of informed consent, the effective functioning of a biobank necessitate that there should be as few restrictions as possible. A major challenge not mentioned above is that of the education of the general public about biobanks as one study reported strong evidence that more people are likely to embrace the idea of biobank research if they are informed (Gaskell & Gottweis, 2011). Interestingly, they proposed the use of social networking sites to promote the concept.

Now that the IDB has been established for several years most of the governance and ethics issues have been established, nevertheless new challenges are arising. Most pressing for the immediate future is improving the IDB's IT and data processing capabilities. This arises from the facts that the IDB's MTA which requires researchers to feedback raw experimental data and with some of the initial studies drawing to a close the quantity of information is becoming overwhelming. Ultimately though this will eventually permit the IDB to amass a detailed database on the patient volunteers and subsequently permit the multivariate analyses on these cohorts to identify important pathogenic markers.

6. Acknowledgements

The IDB is grateful for generous funding from Guy's and St Thomas' Charity and from the NIHR cBRC. In addition the participation of the volunteers and enthusiasm of the IDB staff is greatly appreciated.

7. References

Alibk K, Handelman S. Biohazzard: The chilling story of the largest covert biological weapons programme in the world – told from inside by the man who ran it. 1999. Delta (2000) ISBN 0-385-33496-6

Alvarez R, Reading J, King DF, Hayes M, Easterbrook P, Farzaneh F, Ressler S, Yang F, Rowley D, Vyakarnam A. WFDC1/ps20 is a novel innate immunomodulatory signature protein of human immunodeficiency virus (HIV)-permissive CD4+ CD45RO+ memory T cells that promotes infection by upregulating CD54 integrin expression and is elevated in HIV type 1 infection. *J Virol*. 2008; 82: 471-486.

Aquaron R, Djatou M, Kamdem L. Sociocultural aspects of albinism in Sub-Saharan Africa: mutilations and ritual murders committed in east Africa (Burundi and Tanzania) *Med Trop* 2009; 69: 449-453.

Bachrach S. In the name of public health-Nazi racial hygiene. *N.Eng J Med* 2004; 351: 417-420.

Barenblat D. A plague upon humanity: the secret genocide of Axis Japan's germ warfare operation. Harper-Collins, 2004. ISBN 0-06-018625-9

Bowman KW, Hui EC. Bioethics for clinicians: 20. Chinese bioethics. *CMAJ*. 2000; 163:1481-1485.

Bratton C, Chavin K, Baliga P.Racial disparities in organ donation and why. *Curr Opin Organ Transplant*. 2011; 16: 243-249.

Bresnahan MJ, Mahler K. Ethical debate over organ donation in the context of brain death. *Bioethics* 2010; 24: 54-60.

Callender CO, Miles PV. Minority organ donation: the power of an educated community. *J Am Coll Surg*. 2010; 210: 708-715.

Chaigneau C, Cabioch T, Beaumont K, Betsou F. Serum biobank certification and the establishment of quality controls for biological fluids: examples of serum biomarker stability after temperature variation. *Clin Chem Lab Med*. 2007; 45: 1390-1395.

Chalmers D. Genetic research and biobanks. *Methods Mol Biol*. 2011; 675: 1-37.

Council of Europe, 2006. https://wcd.coe.int/wcd/ViewDoc.jsp?id=977859

Crenner C. The Tuskegee Syphilis Study and the Scientific Concept of Racial Nervous Resistance. *J Hist Med Allied Sci.* 2011 Feb 12. [Epub ahead of print]

Daily Telegraph, 2011. http://www.telegraph.co.uk/culture/culturenews/5319434Museum-returns-old-Aboriginal-skull-to-Australia.html

Deml L, Speth C, Dierich MP, Wolf H, Wagner R.Recombinant HIV-1 Pr55gag virus-like particles: potent stimulators of innate and acquired immune responses. *Mol Immunol*. 2005; 42: 259-277.

Deutscher Ethikrat, 'Human biobanks for research'. Press release: 05/2010 http://www.thikrat.org/press/press-releases/2010/press-release-05-2010

Emerson CI, Singer PA, Upshur RE. Access and use of human tissues from the developing world: ethical challenges and a way forward using a tissue trust. *BMC Med Ethics*. 2011; 12:2.

Farnell LR. 2004. Chapter 10. Greek hero cults and ideas of immortality. Kessinger Publishing. Pp234-279. ISBN 978-1417921348. P269. The famous Hippocratean oath may not be an authentic deliverance of the great master, but is an ancient formula current in his school.

Freidenwald H. 1917. Prayer of Mimondes. *Johns Hopkins Hospital Bulletin*, August 1917.

Freidlander H. The origins of the Nazi genocide: from euthanasia to the final solution. Chapel Hill: University of North Carolina Press, 1995.

García-Merino I, de Las Cuevas N, Jiménez JL, Gallego J, Gómez C, Prieto C, Serramía MJ, Lorente R, Muñoz-Fernández MA; Spanish HIV BioBank. The Spanish HIV BioBank: a model of cooperative HIV research. *Retrovirology* 2009; 6:27.

García-Merino I, de Las Cuevas N, Jiménez JL, García A, Gallego J, Gómez C, García D, Muñoz-Fernández MA. Pediatric HIV BioBank: a new role of the Spanish HIV BioBank in pediatric HIV research. *AIDS Res Hum Retroviruses* 2010; 26: 241-244.

Garozzo G, Crocco I, Giussani B, Martinucci A, Monacelli S, Randi V. Adverse reactions to blood donations: the READ project. *Blood Transfus.* 2010; 8: 49-62.

Gaskell G, Gottweis H. Biobanks need publicity. *Nature* 2011; 471: 159-160.

Gelling P. 2007. http://www.nytimes.com/2007/03/26/world/asia/26cnd-flu.html?_r=1&scp=7&sq=indonesia %20h5n1&st=cse

Gertz R.Withdrawing from participating in a biobank--a comparative study. *Eur J Health Law* 2008; 15: 381-389.

Gillon R."It's all too subjective": scepticism about the possibility or use of philosophical medical ethics. *Br Med J (Clin Res Ed)*. 1985; 290: 1574-1575.

Gregor MJ. 1998. *Groundwork of the metaphysics of morals*. Cambridge University Press. Cambridge, U.K., New York. ISBN 0521622352.

Hahn SZ. Medical ethics in the life of a physician of the soul- reflections on the medical ethics concept of Albert Moll (1862-1936). *Gesamte Inn Med* 1984; 39: 558-561.

Hansson MG. The need to downregulate: a minimal ethical framework for biobank research. *Methods Mol Biol.* 2011; 675: 39-59.

Hansson MG, Dillner J, Bartram CR, Carlson JA, Helgesson G.Should donors be allowed to give broad consent to future biobank research? *Lancet Oncol*. 2006; 7: 266-269.

Harris SH. Factories of death: Japanese biological warfare 1932-1945 and the American cover-up. Routledge, 1984. ISBN 0-415-09105-5

Health Protection Agency, 2008. http://www.hpa.org.uk/web/HPAwebfile/HPAweb_C/1200471685793

Hoeyer K. The role of ethics in commercial genetic research: notes on the notion of commodification. *Med Anthropol.* 2005; 24: 45-70.

Hoeyer K, Olofsson BO, Mjörndal T, Lynöe N.The ethics of research using biobanks: reason to question the importance attributed to informed consent. *Arch Intern Med.* 2005; 165: 97-100.

Hofmann B. Broadening consent--and diluting ethics? *J Med Ethics* 2009; 35: 125-129.

Howard A, Smith M. 2004. William Burke and William Hare. River of blood: serial killers and their victims. Universal. p 50. ISBN 1581125186

Kaufman DJ, Murphy-Bollinger J, Scott J, Hudson KL.Public opinion about the importance of privacy in biobank research. *Am J Hum Genet*. 2009 Nov;85(5):643-54. Epub 2009 Oct 29.

Kaye J. Abandoning informed consent: the case of genetic research in population collections –genetic databases: socio-ethical issues in the collection and use of DNA. Tutton R, Corrigan O (ed.) London, New York: Routledge, 2004.

Kingsley LA, Detels R, Kaslow R, Polk BF, Rinaldo CR Jr, Chmiel J, Detre K, Kelsey SF, Odaka N, Ostrow D, et al.Risk factors for seroconversion to human immunodeficiency virus among male homosexuals. Results from the Multicenter AIDS Cohort Study. *Lancet* 1987; 1(8529): 345-349.

Kozlakidis Z, Mant C, Abdinur F, Cope A, Steiner S, Peakman M, Hayday A, Cason J. Variation of peripheral blood mononuclear cell RNA quality in archived samples. *Biopreservation and Biobanking* 2011: In the press.

Kozlakidis Z, Mant C, Peters B, Post F, Fox J, Philpott-Howard J, Tong W C-Y, Edgeworth J, Peakman M, Malim M, Cason J. How representative are research tissue banks of the local population? Experience of the Infectious Diseases BioBank at King's College London. *Biopreservation and Biobanking:* In press.

Kristof ND. Unmasking horror. (March 17, 1995). New York Times. A special report. Japan confronting gruesome war atrocity.

Lambeth census, 2001. www.lambeth.gov.uk/services/councildemocracy/statisticscensus information *Lancet* editorial, 1828-9 (1) pp 818-821. 28th March 1829.

Lassila KD, Branch MA. Whosse skull and bones? Yale Alumni Magazine (May/June 2006). http://www.yalealumnimagazine.com/issues/2006_05/notebook.html

Leon-Quinto T, Simon MA, Cadenas R, Jones J, Martinez-Hernandez FJ, Moreno JM, Vargas A, Martinez F, Soria B.Developing biological resource banks as a supporting tool for wildlife reproduction and conservation The Iberian lynx bank as a model for other endangered species. *Anim Reprod Sci.* 2009; 112: 347-361.

Lipworth W, Morrell B, Irvine R, Kerridge I.An empirical reappraisal of public trust in biobanking research: rethinking restrictive consent requirements. *J Law Med.* 2009; 17: 119-132.

Melas PA, Sjöholm LK, Forsner T, Edhborg M, Juth N, Forsell Y, Lavebratt C.Examining the public refusal to consent to DNA biobanking: empirical data from a Swedish population-based study. *J Med Ethics*. 2010; 36: 93-98.

Meulenkamp TM, Gevers SK, Bovenberg JA, Koppelman GH, van Hylckama Vlieg A, Smets EM.Communication of biobanks' research results: what do (potential) participants want? *Am J Med Genet A.* 2010; 152A: 2482-2492.

Mills JS. 1972. On liberty in utilitarianism, liberty and representative government, 73, Everyman library. London : J.M. Dent (ed.)

Mitchell WC. Bentham's *Felicific Calculus*. *Political Science Quarterly* 1918; 33: 161-183.

Moore v. Regents of the University of California (1990) 51 Cal. 3d 120; 271 Cal. Rptr. 146; 793 P.2d 479.

Nath R, Mant CA, Kell B, Cason J, Bible JM.Analyses of variant human papillomavirus type-16 E5 proteins for their ability to induce mitogenesis of murine fibroblasts. *Cancer Cell Int*. 2006; 6: 19.

Nath R, Mant C, Luxton J, Hughes G, Raju KS, Shepherd P, Cason J.High risk of human papillomavirus type 16 infections and of development of cervical squamous intraepithelial lesions in systemic lupus erythematosus patients. *Arthritis Rheum.* 2007; 57: 619-625.

NRES, 2007. http://www.nres.npsa.nhs.uk

Oelrichs R, Tsykin A, Rhodes D, Solomon A, Ellett A, McPhee D, Deacon N.Genomic sequence of HIV type 1 from four members of the Sydney Blood Bank Cohort of long-term nonprogressors. *AIDS Res Hum Retroviruses* 1998; 14: 811-814.

Park A. 2009. http://www.time.com/time/specials/packages/article/0,28804,1884779_1884 72_1884766,00.html

Paton HJ. The categorical imperative: A study of Kant's moral philosophy. The University of Chicago Press. Chicago, Illinois, 1948, p3.

Pearson H. Summit calls for clear view of deposits in all biobanks. *Nature* 2004; 432: 426.

Prograis LJ Jr.Tuskegee Bioethics Center 10th anniversary presentation: "Commemorating 10 years: ethical perspectives on origin and destiny". *J Health Care Poor Underserved*. 2010; 21(3 Suppl): 21-25.

Pulley J, Clayton E, Bernard GR, Roden DM, Masys DR. Principles of human subjects protections applied in an opt-out, de-identified biobank. *Clin Transl Sci.* 2010; 3: 42-48.

Roy B. The Tuskegee Syphilis Experiment: biotechnology and the administrative state. *J Natl Med Assoc*. 1995; 87: 56-67.

Royal Liverpool Children's enquiry, 2001. Report ordered by the House of Commons. The Stationary Office. http://www.rlcinquiry.org.uk/

Salim A, Berry C, Ley EJ, Schulman D, Desai C, Navarro S, Malinoski D. The impact of race on organ donation rates in Southern California. *J Am Coll Surg*. 2010; 211: 596-600.

Smit M. 2008. http://www.telegraph.co.uk/news/worldnews/1582092/Alaistair-Cooke-body-snatch-leader-pleads-guilty.html

Taha A. Nietzsche, Prophet of Nazism: the cult of the superman-unveiling the Nazi secret doctorine. Author House, Bloomingto Indianna, 2005.

Thorborn G, Pomeroy L, Isohanni H, Perry M, Peters B, Vyakarnam A. Increased sensitivity of CD4+ T-effector cells to CD4+CD25+ Treg suppression compensates for reduced Treg number in asymptomatic HIV-1 infection. *PLoS One* 2010; 5: e9254.

Towie N. London hospital launches infectious disease 'biobank'. *Nat Med*. 2007; 13: 653.

UNOS. Organ and tissue donation. A reference for clergy, 4th ed., 2000. Cooper ML, Taylor GJ (eds.) Richmond, Virginia.

US Government Printing Office. Trials of war criminals before the Nuremberg Military Tribunals under Control Council Law No. 10, Vol. 2, pp. 181-182. Washington D.C.: 1949.

Williams R, Mant C, Cason J. The Infectious Diseases BioBank at King's College London: archiving samples from patients infected with HIV to facilitate translational research. *Retrovirology* 2009; 6: 98.

World Medical association Inc. Declaration of Helsinki. Ethical principles for medical research involving human subjects. *J Indian Med Assoc*. 2009; 107: 403-405.

Zika E, Paci D, Braun A, Rijkers-Defrasne S, Deschênes M, Fortier I, Laage-Hellman J, Scerri CA, Ibarreta D. A European survey on biobanks: trends and issues. *Public Health Genomics* 2011; 14: 96-103.

Part 5

Physiological Systems Engineering in Medical Assessment

32

Cardiac Myocardial Disease States Cause Left Ventricular Remodeling with Decreased Contractility and Lead to Heart Failure; Interventions by Coronary Arterial Bypass Grafting and Surgical Ventricular Restoration Can Reverse LV Remodeling with Improved Contractility

Dhanjoo N. Ghista[1], Liang Zhong[2], Leok Poh Chua[3],
Ghassan S. Kassab[4], Yi Su[5] and Ru San Tan[2]
[1]*Department of Graduate and Continuing Education, Framingham State University, Framingham, Massachusetts,*
[2]*Department of Cardiology, National Heart Centre,*
[3]*School of Mechanical and Aerospace Engineering, Nanyang Technological University,*
[4]*Departments of Biomedical Engineering, Surgery, Cellular and Integrative Physiology, Indiana University-Purdue University Indianapolis, Indianapolis, Indiana,*
[5]*Institute of High Performance Computing, Agency for Science, Technology and Research,*
[1,4]*USA*
[2,3,5]*Singapore*

1. Introduction

1.1 Theme and scope

In this chapter, we are studying the course (i) of cardiomyopathy diseased LVs (with myocardial infarcts) progressing to heart failure (HF) through LV remodelling and decreased LV contractility, and (ii) their recovery through surgical therapeutic interventions of CABG and Surgical Ventricular Restoration (SVR), by restoration of myocardial ischemic segments, reversal of LV remodeling and improvement in LV contractility.

For this purpose, we first provide the methodology for detecting myocardial infarcts. Then, we characterize LV remodeling of cardiomyopathy diseased LVs (with myocardial infarcts) in terms of reduced change in curvedness from end-diastole to end-systole. In these LVs, there is also reduced contractility; so we provide an index for cardiac contractility, in terms of maximal rate-of-change of normalized wall stress, $d\sigma^*/dt_{max}$, and its decrease in an infarcted LV progressing to heart failure. We provide clinical studies of remodeled cardiomyopathy diseased LVs, in terms of reduced values of their curvedness index and contractility index.

By way of therapeutic interventions, we have presented the hemodynamic flow simulation of the CABG (carried out to enhance myocardial perfusion in the region around the blocked coronary artery), and pointed out certain factors and sites of wall shear stresses that cause intimal damage of vessels and hyperplasia, as potential causes for decreased patency. We have shown that surgical ventricular restoration (SVR), in conjunction with CABG, is seen to benefit the ischemic-infarcted heart, by (i) restoration of cardiac remodeling index of 'end-diastolic to end-systolic curvedness change', (ii) reduction of regional wall stresses, and (iii) augmentation of the cardiac contractility index value.

2. Myocardial infarction: What it entails

In Cardiology, the etiology of congestive heart failure (CHF) is coronary artery disease in approximately two-thirds of cases. The majority of these patients have hearts with myocardial infarcted segments. This infarcted myocardial wall mitigates adequate contraction of the wall. So, the end-result of an infarcted left ventricle (LV) is poor intra-LV velocity distribution and pressure-gradient distribution, causing impaired outflow from the LV into the aorta.

In the infarcted myocardial segments, the myocardial infrastructure of actin and myosin filaments (and their cross – bridges) is disrupted, and hence there is no contraction within these infarcted myocardial segments. The below figure 1 illustrates a myocardial sarcomere segment's bioengineering model, composed of two symmetrical myocardial structural units (MSUs). In these MSU(s), the contractile elements represent the actin-myosin contractile components of the sarcomere segment. The disruptions of these contractile elements impair the contractile capability of that sarcomere segment. Hence, a LV with infarcted myocardial segments will have diminished contractility, inadequate and improper intra-LV flow, and poor ejection.

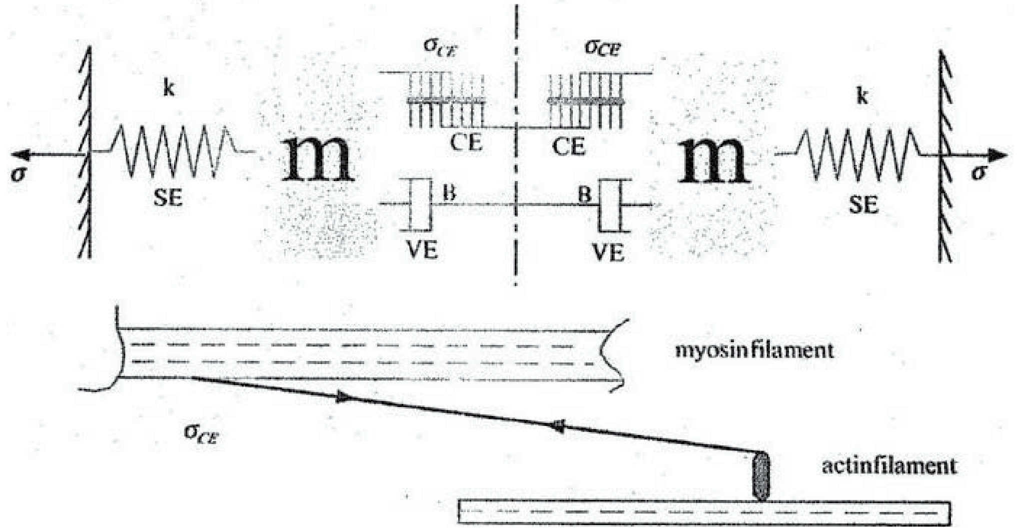
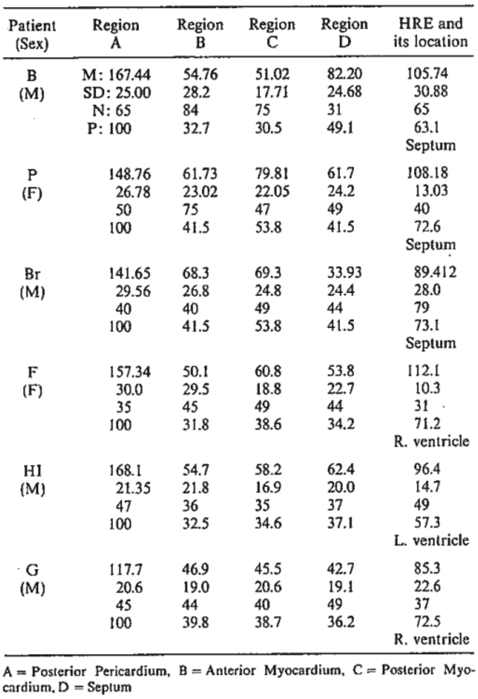


Fig. 1. Based on the conventional Hill three-element model and Huxley cross bridge theory, we have developed a myocardial model involving the LV myocardial mass, series-elastic element (CE),. In this figure we have linked the anatomical associations of these myocardial model elements with microscopic structure of the heart muscle. This figure illustrates the sarcomere element contractile model, involving: the effective mass (*m*) of the muscle tissue that is accelerated: elastic parameter k of the series element stress σ_{SE} (k = elastic modulus of the sarcomere), viscous damping parameter B of the stress σ_{VE} in the parallel viscous element VE, the generated contractile stress σ_{CE} between myosin (thick) and actin (thin) filaments.

3. Detection of myocardial infarcted segments

Now, infarcted myocardial segments can be detected as highly reflectile echo zones (HREZs) in 2-dimensional B-scan echocardiograms. In this context, we have shown how infarcted myocardial segments can be detected (in shape and size), by echo-texture analysis, as highly reflectile echo zones or HREZs. Each myocardial tissue component of the heart generates a grey scale pattern or texture related to the tissue density and fibrous content. In diseased states (such as myocardial ischemia, myocardial fibrosis, and infiltrative diseases), changes in myocardial tissue density have been recognised by employing echo intensity and mean grey level of pixel as the basis for recognition of such myocardial disorders. It is found that hyper-reflectile echoes (HREs) correlated well with diseased cardiac muscle, and that myocardial tissue containing HREs corresponded with foci of sub-endorcardial necrosis and even calcification.

In our study [1], in order to determine highly reflectile echo zones (HREZs), echocardiograms were recorded, and each image was made up of 256 x 256 pixels, with each pixel having a resolution of 0-256 grey scales. The echocardiographic images were digitised into 256 grey scales. Then, echo intensity levels from normal infants were used to delineate the range of echo intensities for normal tissues. The upper bound of the echo intensity was set to 100 per cent in each normal infant, and the intensities from the rest of the image were referenced to this level. Normally, pericardium had the highest intensity level. It was found that the upper-bound of the echo intensity value for healthy tissue (expressed as a percentage of the pericardial echo intensity value) was 54.2.



Patient (Sex)	Region A	Region B	Region C	Region D	HRE and its location
B	M: 167.44	54.76	51.02	82.20	105.74
(M)	SD: 25.00	28.2	17.71	24.68	30.88
	N: 65	84	75	31	65
	P: 100	32.7	30.5	49.1	63.1
					Septum
P	148.76	61.73	79.81	61.7	108.18
(F)	26.78	23.02	22.05	24.2	13.03
	50	75	47	49	40
	100	41.5	53.8	41.5	72.6
					Septum
Br	141.65	68.3	69.3	33.93	89.412
(M)	29.56	26.8	24.8	24.4	28.0
	40	40	49	44	79
	100	41.5	53.8	41.5	73.1
					Septum
F	157.34	50.1	60.8	53.8	112.1
(F)	30.0	29.5	18.8	22.7	10.3
	35	45	49	44	31
	100	31.8	38.6	34.2	71.2
					R. ventricle
HI	168.1	54.7	58.2	62.4	96.4
(M)	21.35	21.8	16.9	20.0	14.7
	47	36	35	37	49
	100	32.5	34.6	37.1	57.3
					L. ventricle
G	117.7	46.9	45.5	42.7	85.3
(M)	20.6	19.0	20.6	19.1	22.6
	45	44	40	49	37
	100	39.8	38.7	36.2	72.5
					R. ventricle

A = Posterior Pericardium, B = Anterior Myocardium, C = Posterior Myocardium, D = Septum

Table 1. Echo intensity values for various anatomic regions of diseased pediatric hearts (based on long axis view). The numbers in the four rows represent Mean (M), Standard Deviation (SD), Number of Pixels (N), Percentage of Posterior Pericardial Intensity (P). This figure is adopted from Ref [1].

For patients whose echo-texture analysis showed presence of HREs, it was found that the echocardiographic intensities of the HREs from these patients intensities were distinctly higher than the echo intensity range of normal tissue (as depicted in Table 1).

Myocardial tissue pixels having echo-intensity values greater than 200 are generally noted to be infarcted. This region's echo-intensity values can remain unaffected by administration of a myocardial perfusing agent. This infarcted sub-region is seen to be surrounded by an ischemic sub-region whose pixels are noted to have echo intensity values between 100 and 200. This region's echo-intensity can be reduced by the administration of a myocardial perfusing agent. The surrounding healthy tissue has echo intensity less than 100. Figure 2(a) depicts an echo image of an infant with visible scars regions 1 and 2, while figure 2(b) depicts printouts of the echo intensities from these two regions, wherein the infarcted segments are depicted in dark colour and the surrounding ischemic segments are depicted in a lighter shade.

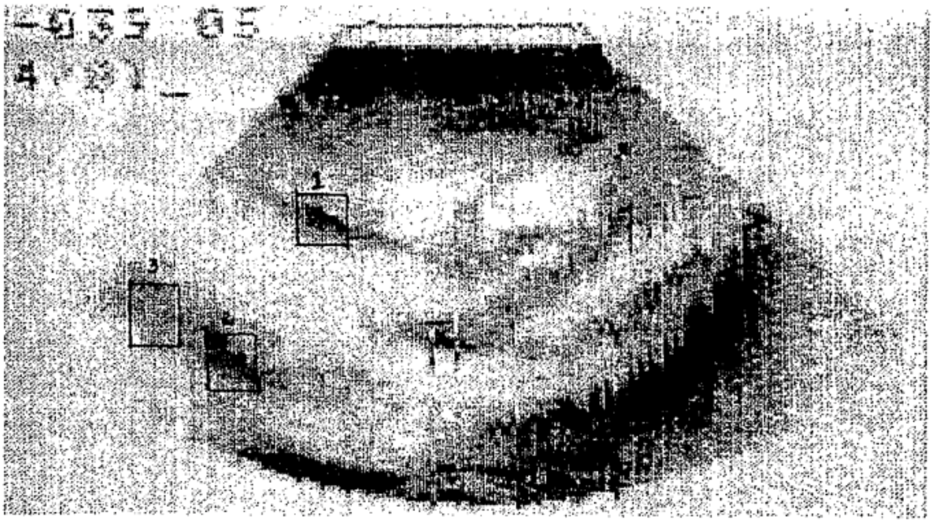
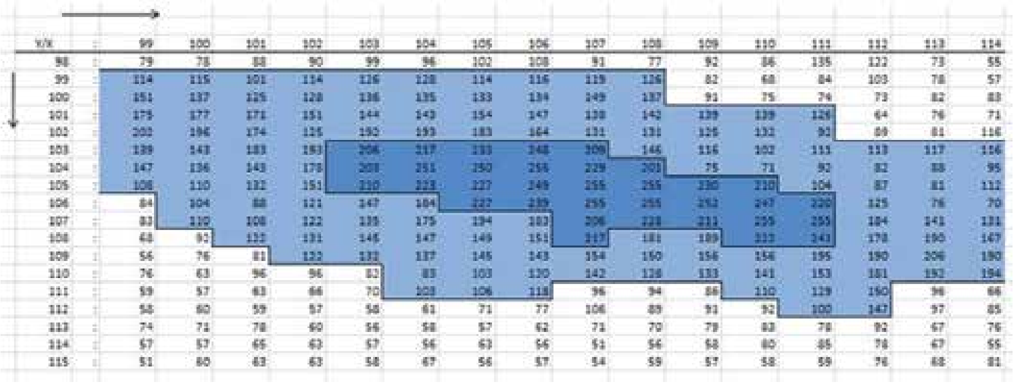


Fig. 2. (a) Long axis view of a pediatric patient's heart showing HRE regions 1 and 2 and a healthy region 3 [1].



Y/X	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114
98	79	78	88	90	99	96	102	108	91	77	92	86	135	122	73	55
99	114	115	101	114	126	128	114	116	119	126	82	68	84	103	78	57
100	151	137	125	128	136	135	133	134	149	137	91	75	74	73	82	83
101	175	177	171	151	144	143	154	147	138	142	139	139	126	64	76	71
102	202	196	174	125	190	193	183	164	131	131	125	132	92	89	81	116
103	139	143	183	193	206	217	233	248	209	146	116	102	111	113	117	156
104	147	136	143	178	208	251	250	255	229	203	75	71	92	82	88	95
105	108	110	132	151	210	223	227	249	255	255	230	210	104	87	81	112
106	84	104	88	121	147	184	227	239	255	255	252	247	220	125	76	70
107	83	110	108	122	135	175	194	183	208	228	211	255	255	184	141	131
108	68	93	132	131	146	147	149	153	217	181	189	233	243	178	190	167
109	56	76	81	122	132	137	145	143	154	150	156	156	195	190	206	190
110	76	63	96	96	83	83	103	130	142	128	133	141	153	181	192	194
111	59	57	63	66	70	103	106	118	96	94	86	110	129	150	98	66
112	58	60	59	57	58	61	71	77	106	89	91	92	100	147	97	85
113	74	71	78	60	56	58	57	62	71	70	79	82	78	92	67	76
114	67	57	65	63	57	56	63	56	51	56	58	80	85	78	67	55
115	51	60	63	63	58	67	56	57	54	59	57	58	59	76	68	81

Fig. 2. (b) Pixel values correcponding to highly reflectile echo region 1. The central region having echo-intensity values greater than 200 is infarcted, while its immediately surrounding region shown in lighter shade is ischemic.

In this way, in each highly reflectile echo zone (HREZ) made up of , say, N number of pixels, we can determine the number (I) of infarcted pixels. The ratio I/N represents the infarcted potion of that HREZ myocardial segment. The total number of all the infarcted pixels in all the HREZs provides an indication of the amount of infarcted myocardium of the heart or of the LV.

4. Cardiac remodeling following myocardial infarction and progression to heart failure

Myocardial infarction (MI) reduces the amount of ventricular contractile myocardium, which in turn reduces the left ventricular (LV) contractile capacity for pumping out adequate cardiac output (CO). As MI extends, it progresses into heart failure (HF). A manifestation of MI and its decreased contractile capacity is the inability of the LV to retain its compact systolic curved shape. In other words, in HF resulting from MI, the left ventricle undergoes remodeling and its curvedness index decreases. As MI progresses into HF, LV size increases, LV function deteriorates, and symptoms of HF become evident. Thus, cardiac remodeling (expressed in terms of regional curvedness of the LV) constitutes a measure of the progression of HF after MI, and interventions causing some reversal of this LV remodeling process have been shown to improve mortality in patients with HF. Therefore, reversal of LV remodeling (through medical or surgical treatment) has been emerging as a therapeutic target in HF of all etiologies. The challenge is to develop specific measures of LV remodeling that can be incorporated into the clinical management pathway. For patients with HF after MI, the LV shape is more spherical in terms of the global sphericity index. However, focus on global sphericity index may be misleading, since the simple plane ratio reflects a linear alteration in the two axes of the LV chamber. Hence, regional curvedness index, determined from 3D magnetic resonance imaging (MRI)-based LV model, is proposed as a measure of LV shape. The curvedness value describes the magnitude of the curvature at a surface point, i.e., a measure of degree of curvature a point; the regional curvedness describes the curvedness of the segment of the LV. In MI patients, the regional curvedness index value does not increase significantly from end-diastole to end-systole (due to decreased LV contractile capacity), as in the case of normal patients.

In the case of patients with HF following MI, the LV cannot generate adequate contractile force, as explained earlier in section 1. As a result, the overall LV contractility index of maximal rate-of-change of normalized systolic wall stress $d\sigma^*/dt_{max}$ is decreased, and the LV ejection force is also diminished. Now, LV regional wall stress is proportional to the wall surface radius-of-curvature and inversely proportional to the wall thickness. Hence, for patients with HF after MI, the inability of the remaining viable myocardium to compensate for the increased wall stress associated with LV dilatation and thinning is a trigger for LV enlargement.

4.1 Remodeling quantification in terms of local curvedness indices and regional curvedness index

4.1.1 Overall approach

The remodeling is characterized in terms of the *Regional Curvedness* index of the 16 segments of the LV endocardial surface. The method entails generation of the LV endocardial shape and its compartmentalization into 16 segments or regions (as explained later on). Each segment is discretized into triangular meshes, so that a point on the endocardial surface is a vertex of a triangular mesh. At each vertex point, we determine the *Local Curvedness* index

[2], by employing its surrounding neighboring vertices in the form of *n*-rings around the local point. To ensure accurate curvedness computation without over smoothing, the optimal number of *n*-rings is determined to be 5. Then, for each segment, we determine the *Regional Curvedness* index, as the mean of the *Local Curvedness* indices in the segment.

The values of *Regional Curvedness* are determined for normal and MI patients, at end-diastolic and end-systolic instants. For normal patients, the *Regional Curvedness* index changes significantly from diastole to systole, as given by the *Diastole-to-Systole Change in Curvedness* (%ΔC) [2]

$$\%\Delta C = \frac{C_{ED} - C_{ES}}{C_{ED}} \times 100 \tag{1}$$

wherein C_{ED} and C_{ES} are end-diastolic and end-systolic curvedness, defined later on. The mean alteration in regional curvedness index value of %ΔC is determined for all the segments, for normal subjects and MI patients (and logged in Table 3). It can be seen that for MI patients, the regional curvedness index or (%ΔC) does not change significantly from end-diastole to end-systole, in comparison with normal subjects.

4.1.2 Clinical application methodology

Our study involved 10 normal subjects and 11 patients after myocardial infarction (MI). The hemodynamic and volumetric parameters of the subjects are summarized in Table 2.

Human Subjects and MRI Scans: The study to characterize regional curvedness index or (%ΔC) involved ten normal subjects and 11 patients after myocardial infarction [2]. All subjects underwent diagnostic MRI scans. For each subject, short-axis MRI images were taken along the plane which passes through the mitral and aortic valves of the heart at an interval of 8mm thickness. Each image has a spatial resolution of 1.5mm, acquired in a single breath hold, with 25 temporal phases per heart cycle. Of these images, the set of images corresponding to the cardiac cycle at end-diastole and end-systole are then used for the study.

LV Endocardial Surface Reconstruction and Segmentation: The MRI images were processed, by using a semi-automatic technique that is included in the CMRtools suite (Cardiovascular Solution, UK). The contours demarcating the myocardium and the LV chamber were defined by means of B-spline curves. The endocardial surface of the LV was reconstructed by joining the series of contours to form a triangle mesh. In order to facilitate quantification of the LV segmental regional curvedness, the endocardial surface was partitioned into 16 segments; the method of segmentation of the LV endocardial surface is provided in our paper [2].

Left Ventricular Shape Analysis: To quantify LV remodeling, we first define a measure known as the *Local Curvedness* index [2]. This is essentially a shape descriptor used to quantify how curved the surface is in the vicinity of a vertex on the LV endocardial surface. This is done by using the 3-d mesh of the LV endocardial surface as an input. Each vertex of the mesh is processed by fitting a quadric surface over a local region around the vertex as described in our paper [2]. The extent of this local region is determined by the *n*-ring parameter. Next, the *Local Curvedness* index of each vertex can be calculated from the coefficients of the fitted quadric surface by [2]:

$$C = \sqrt{\frac{\kappa_1{}^2 + \kappa_2{}^2}{2}} = \frac{1}{A^2}\sqrt{\frac{2B^2 + A^2\left(4ac - b^2\right)}{A}} \tag{2}$$

such that

$$A = \sqrt{d^2 + e^2 + 1}$$

$$B = a + ae^2 + c + cd^2 + bde$$

where a, b, c, d and e are the coefficients of the fitted quadric surface at the vertex.

In order to derive the *Regional Curvedness*, the endocardial surface is partitioned into 16 segments (Fig 3). The *Regional Curvedness* for the each segment is the mean of the *Local Curvedness* indices in the segment. The flowchart of the overall workflow for the regional LV shape analysis is shown in Fig. 4.

4.1.3 Clinical studies results

In our clinical studies, it was found that (1) MI patients exhibit decreased curvedness and $\%\Delta C$, (2) MI patients exhibit increased variation of curvedness and variation of $\%\Delta C$, and (3) LV ejection fraction is positively correlated with curvedness and $\%\Delta C$, and inversely correlated with variation of $\%\Delta C$.

The *Diastole-to-Systole Change in Curvedness* ($\%\Delta C$), as defined by equation (1), is a measure of regional deformity due to contraction. Positive values of $\%\Delta C$ indicate regions of increasing inward concavity of the LV wall during systole, while negative values of $\%\Delta C$ indicate wall regions of decreasing inward concavity. The $\%\Delta C$ measure can be employed to relate the regional differences in hypokinesis due to myocardial infarction.

Variation of curvedness: The extent of LV surface inhomogeneity is characterized by a coefficient of variation of curvedness at end-diastole (CV_C_{ED}) and end-systole (CV_C_{ES}):

$$CV\_C = \frac{\sigma(C)}{\mu(C)} \tag{3}$$

where $\sigma(C)$ is the standard deviation of the regional curvedness and $\mu(C)$ is the mean of the regional curvedness of the segments of the LV mesh.

To evaluate the extent of functional non-uniformity of LV regions, the index CV_DC was determined as [2]:

$$CV\_\Delta C = \frac{CV\_C_{ED} - CV\_C_{ES}}{CV\_C_{ED}} \times 100 \tag{4}$$

In general, the larger the values of CV_C_{ED} and CV_C_{ES}, the more inhomogeneous the LV endocardial surface appears. Hence, the larger the value of index $CV\_\Delta C$, the more functionally non-uniform are the LV shape changes due to LV contraction.

Curvedness, Variation of curvedness and Diastole-to-systole change $\%\Delta C$ in MI patients compared to Normal subjects: The hemodynamic and volumetric parameters of the subjects are summarized in Table 2. For patients after MI, the LV ejection fraction was significantly lower than that in the control subjects. In addition, their LV end-diastolic and end-systolic indexed volumes were greater than those in the control subjects.

The values of regional curvedness from apex to base in MI patients and normal subjects are given in Table 3 and Fig. 5, to highlight the regional variations of the LV curvature. In the normal group, there was a significant increase in the curvedness at the apex from diastole to systole. However, in MI patients, there was no significant difference in curvedness in all segments. Significant differences in end-diastolic curvedness C_{ED} and end-systolic curvedness C_{ES} were noted between MI and normal groups. Among the 16 segments of the LV, the variation coefficient of C_{ES} (CV_C_{ES}) was significantly lower in MI patients than in the normal group (18±4% in MI vs 31±8% in normal, $p<0.0001$), indicating fair homogeneity of LV shape in MI at end-systole. Correspondingly, the diastole-to-systole change in curvedness (%ΔC) was significantly lower, and the variation of %ΔC was higher in MI patients compared to normal group, indicating ventricular functional non-uniformity due to the pathologic state.

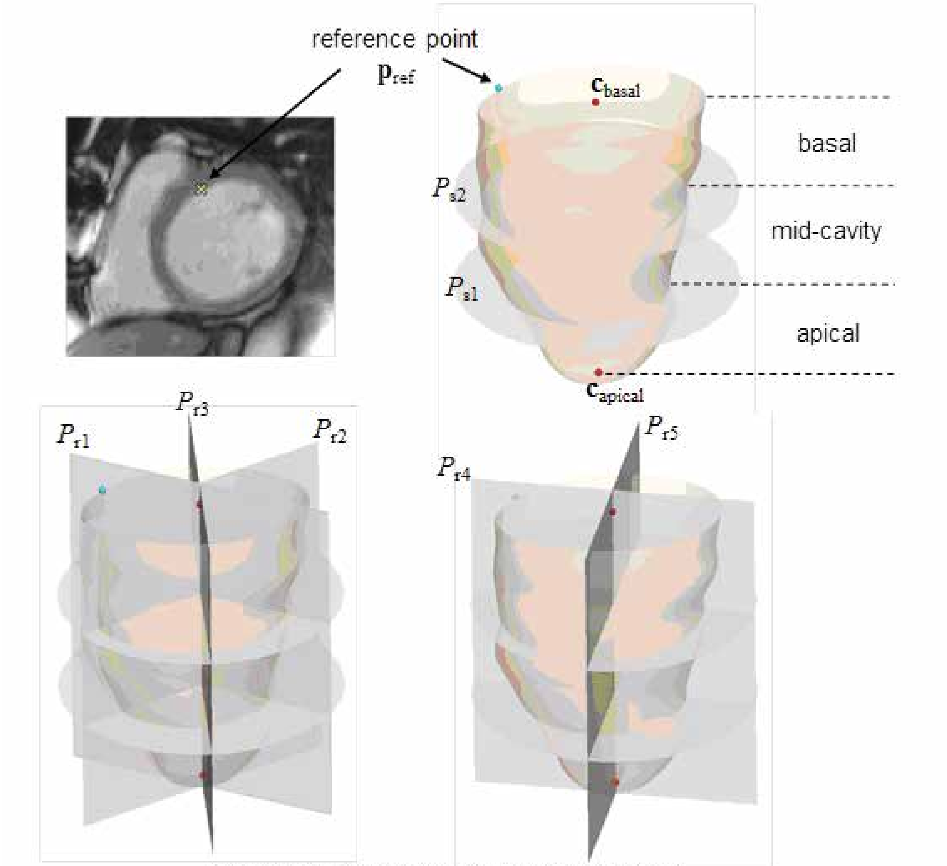
	Control (*n*=10)	MI (*n*=11)	p value
Age (years)	41 ± 16	60 ± 6	0.003
Weight (kg)	67 ± 15	65 ± 14	0.30
Height (cm)	169 ± 8	165 ± 10	0.86
Diastolic pressure (mmHg)	73 ± 12	74 ± 18	0.79
Systolic pressure (mmHg)	122 ± 17	116 ± 20	0.50
HR (beats/min)	70 ± 9	84 ± 13	0.012
CI (L/min/m^2)	3.3 ± 0.4	2.2 ± 0.5	<0.001
EDVI (ml/m^2)	73 ± 10	148 ± 40	<0.001
ESVI (ml/m^2)	26 ± 6	122 ± 38	<0.001
EF (%)	65 ± 5	18 ± 5	<0.001
Sphericity index	0.52 ± 0.06	0.62 ± 0.08	0.01
LV mass index	56 ± 12	83 ± 13	0.004

Table 2. Characteristics of Normal Control and Patients after MI, involved in the study.

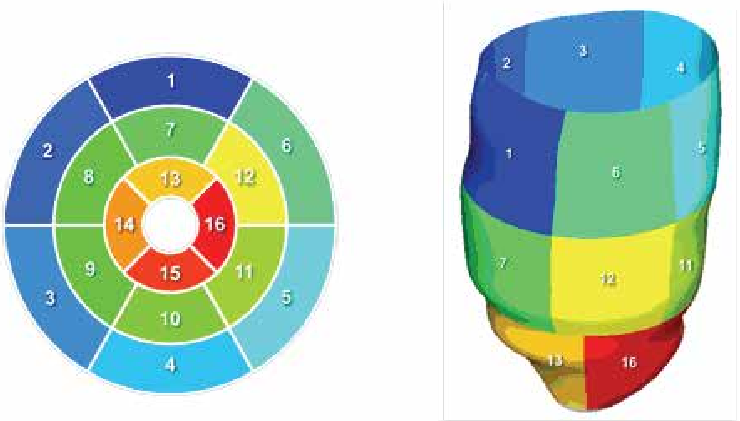
	Controls (*n*=10)			MI (*n*=11)		
Segment	C_{ED}(x10^{-2} mm^{-1})	C_{ES}(x10^{-2} mm^{-1})	ΔC (%)	C_{ED}(x10^{-2} mm^{-1})	C_{ES}(x10^{-2} mm^{-1})	ΔC (%)
1. Basal anterior	4.1 ± 0.8	5.6 ± 0.7	38 ± 22	3.4 ± 0.5*	3.7 ± 0.5ξ	7 ± 17#
2. Basal anterior septal	3.4 ± 0.6	5.3 ± 1.1	57 ± 30	3.7 ± 1.0	3.7 ± 0.9*	4 ± 26ξ
3. Basal inferior septal	3.1 ± 0.5	4.8 ± 0.6	61 ± 25	3.6 ± 0.9	4.0 ± 1.0*	12 ± 23ξ
4. Basal inferior	4.0 ± 0.5	5.8 ± 1.0	45 ± 27	3.7 ± 1.0	4.1 ± 1.0*	13 ± 21*
5. Basal inferior lateral	3.5 ± 0.5	5.3 ± 0.9	50 ± 22	3.0 ± 0.6*	3.3 ± 0.9ξ	13 ± 25*
6. Basal anterior lateral	3.6 ± 0.8	5.0 ± 0.7	46 ± 26	3.1 ± 0.8	3.2 ± 0.8ξ	8 ± 24*
7. Middle anterior	3.9 ± 0.5	6.0 ± 0.1	56 ± 20	3.4 ± 0.2*	3.6 ± 0.3ξ	6 ± 12ξ
8. Middle anterior septal	3.9 ± 0.4	0.6 ± 1.4	55 ± 26	3.3 ± 0.6*	3.4 ± 0.4ξ	5 ± 17ξ
9. Middle inferior septal	3.6 ± 0.6	5.2 ± 1.1	44 ± 18	3.5 ± 0.7	3.4 ± 0.5ξ	1 ± 16ξ
10. Middle inferior	4.1 ± 0.5	6.0 ± 1.2	51 ± 31	3.8 ± 0.7	4.0 ± 0.7ξ	7 ± 17*
11. Middle inferior lateral	3.6 ± 0.3	5.5 ± 0.8	52 ± 25	3.1 ± 0.6*	3.4 ± 0.6ξ	10 ± 14ξ
12. Middle anterior lateral	3.2 ± 0.3	4.7 ± 0.9	45 ± 19	2.9 ± 0.5	3.0 ± 0.3ξ	4 ± 14ξ
13. Apical anterior	4.8 ± 1.0	9.3 ± 2.0	96 ± 33	3.8 ± 0.7*	4.1 ± 1.0ξ	6 ± 15ξ
14. Apical septal	4.9 ± 0.6	9.0 ± 1.9	83 ± 34	4.4 ± 0.7	4.6 ± 0.9ξ	4 ± 14ξ
15. Apical inferior	5.7 ± 0.9	11 ± 2.8	90 ± 44	4.7 ± 1.0*	4.9 ± 1.0ξ	5 ± 14ξ
16. Apical lateral	4.4 ± 0.7	8.8 ± 2.2	103 ± 65	3.8 ± 0.4	4.0 ± 0.7ξ	2 ± 12ξ
Mean	4.0 ± 0.4	6.5 ± 1.0	61 ± 18	3.6 ± 0.5*	3.8 ± 0.5ξ	7 ± 9ξ
Coefficient of variation (%)	21 ± 5	31 ± 8	51 ± 14	19 ± 5	18 ± 4ξ	392 ± 501*

*, $p<0.05$; #, $p<0.01$; ξ, $p<0.001$

Table 3. Left ventricular Regional curvedness, Diastole-to-systole change in curvedness (%ΔC), Variation of curvedness and %ΔC in MI compared to normal state, as defined by equation (1). This table is related to our work in Ref [2].


reference point
$\mathbf{p}_{\mathrm{ref}}$
$\mathbf{c}_{\mathrm{basal}}$
basal
P_{s2}
mid-cavity
P_{s1}
apical
$\mathbf{c}_{\mathrm{apical}}$
P_{r3}
P_{r1}
P_{r2}
P_{r5}
P_{r4}

Reference planes for 3D mesh partitioning


1
2
6
7
8
12
13
14
16
15
9
11
3
5
10
4
2
3
4
1
6
5
7
12
11
13
16

2D nomenclature viewed from the apex

3D mesh partitioning into 16 segments

Fig. 3. Partitioning of LV mesh into 16 segments; this figure is based on our work in Ref. [2].

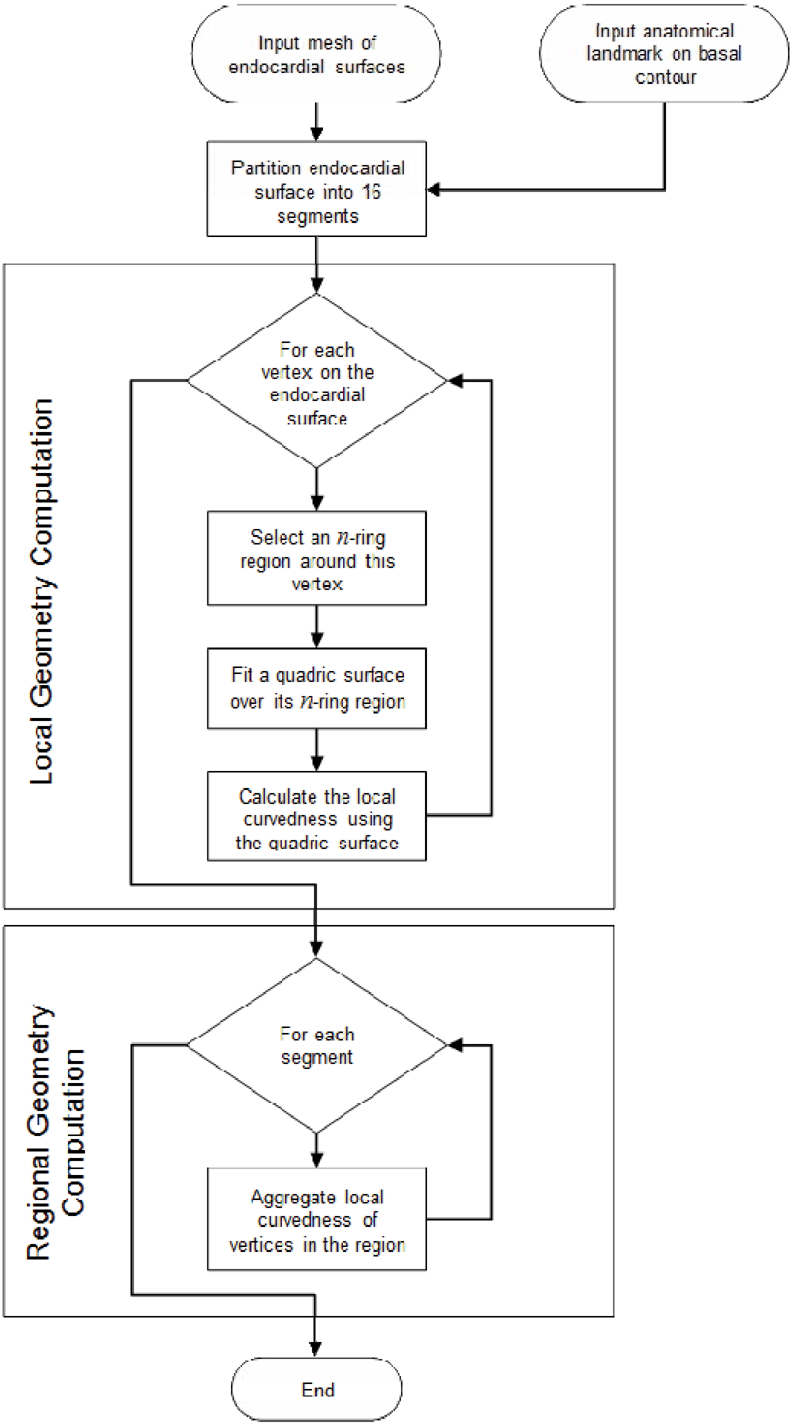

Input mesh of endocardial surfaces
Input anatomical landmark on basal contour
Partition endocardial surface into 16 segments
Local Geometry Computation
For each vertex on the endocardial surface
Select an n-ring region around this vertex
Fit a quadric surface over its n-ring region
Calculate the local curvedness using the quadric surface
Regional Geometry Computation
For each segment
Aggregate local curvedness of vertices in the region
End

Fig. 4. Flowchart of the overall workflow for the regional LV shape analysis; this figure is based on our work in Ref. [2].

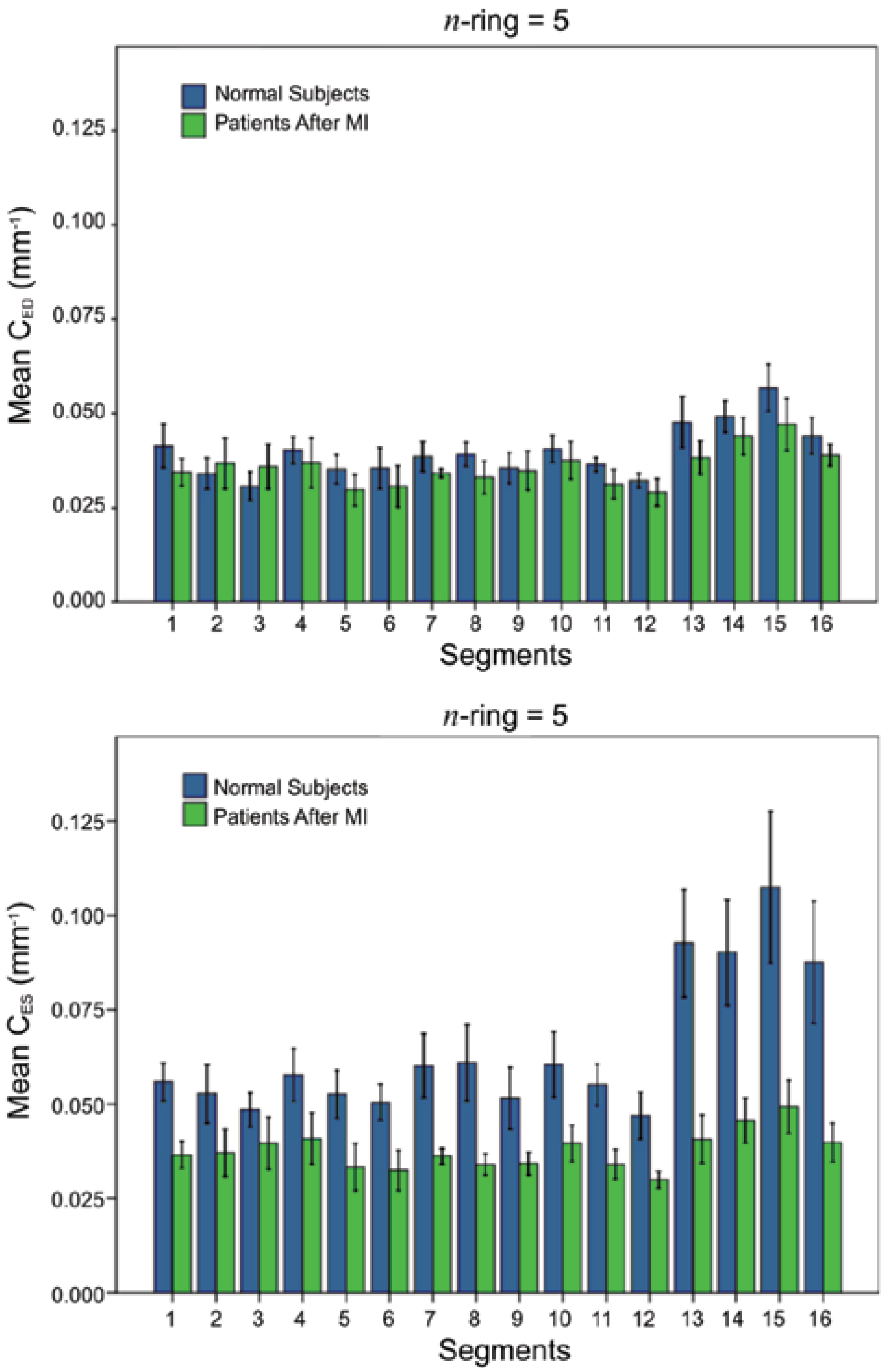


Fig. 5. Comparison of regional curvedness (5-ring selection) in normal subjects and patients after myocardial infarction; this figure is based on our work in Ref. [2].

5. Employment of cardiac contractility Index dσ*/dt$_{max}$ to demonstrate (i) its excellent correlation with traditional contractility index dP/dt$_{max}$, in patients with varying ejection fractions, and (ii) its capacity to diagnose Heart failure with normal ejection fraction (HFNEF, or diastolic heart failure) and with reduced ejection fraction (HFREF, or systolic heart failure)

5.1 Cardiac contractility expressed in terms of maximum rate of change of pressure-normalized LV wall stress, dσ*/dt$_{max}$

Left Ventricular (LV) contractility can be termed as the capacity of the LV to develop intra-myocardial stress, and thereby intra-cavitary pressure, to eject blood volume as rapidly as possible. It is hence rational and appropriate to formulate a LV contractility index on the basis of LV wall stress. The traditional dP/dt_{max} is based on left ventricular intra-cavitary pressure which is generated by an active myocardial stress. Hence, analogous to dP/dt_{max}, which is based on LV intra-cavitary pressure, we have formulated a novel LV contractility index based on LV wall stress, namely the maximum rate of change during systole of LV wall stress normalized to LV intra-cavitary pressure, $d(\sigma/P)/dt_{max}$ or $d\sigma^*/dt_{max}$, where $\sigma^* = \sigma/P$ [3].

Model Analysis and Index Formulation: For mathematically simplicity, we have approximated the LV as a thick-wall spherical shell consisting of incompressible, homogeneous, isotropic, elastic material. The maximum circumferential wall stress (σ_θ) can be expressed at the endocardium, as:

$$\sigma_\theta(r_i) = P\left[\frac{r_i^3/r_e^3 + 1/2}{1 - r_i^3/r_e^3}\right] \tag{5}$$

where r_i and r_e are the inner and outer radii, P is LV intracavitary pressure.
By normalizing wall stress to LV intra-cavitary pressure (P), we obtain:

$$\sigma^*(r_i) = \frac{\sigma_\theta}{P} = \frac{r_i^3}{r_e^3 - r_i^3}\left(1 + \frac{r_e^3}{2r_i^3}\right) \tag{6}$$

Since the maximum wall stress occurs at the inner endocardial wall, we have:

$$\sigma^*(r = r_i) = \left(\frac{V/(V_m + V) + 1/2}{1 - V/(V_m + V)}\right) = \left(\frac{3V + V_m}{2V_m}\right) = \left(\frac{3V}{2V_m} + \frac{1}{2}\right) \tag{7}$$

where P is LV intra-cavitary pressure; σ_θ is the wall stress; V $(= 4\pi r_i^3/3)$ denotes LV volume; V_m $(= 4\pi\left(r_e^3 - r_i^3\right)/3)$ denotes LV myocardial volume; r_i and r_e are the inner and outer radii of the LV, respectively. Differentiating equation (7) with respect to time, we get:

$$d\sigma^*/dt_{\max} = \left|\frac{d(\sigma_\theta/P)}{dt}\right|_{max} = \frac{3}{2V_m}\left|\left(\frac{dV}{dt}\right)\right|_{max} \tag{8}$$

It can be thus noted that in contrast to the indices of dP/dt_{max}, E_{es} and $E_{a,max}$, our $d\sigma^*/dt_{max}$ index can be determined solely from non-invasive assessment of LV geometry and flow. Normalizing LV wall stress to LV pressure obviates the need for invasive LV pressure measurement.

Our LV contractility formulation has been based on the premise that LV wall stress (due to LV myocardial sarcomere contraction) is responsible for the development of LV pressure. Hence, it is more rational to base LV contractile function on LV wall stress per pressure. Hence, analogus to dP/dt_{max}, this LV contractility index is formulated as the maximal rate of pressure-normalized wall stress (as given by the above equation 8), to represent the maximal flow rate out of the ventricle (dV/dt) normalized to myocardial volume (V_m). This is somewhat in keeping with cardiac output or maximal volume change having been used as a measure of myocardial contractility in rats or human, provided that the influence of afterload is taken into account.

5.2 Medical application to subjects with varying ejection fractions, to demonstrate excellent correlation of $d\sigma^*/dt_{max}$ with dP/dt_{max}

We have validated $d\sigma^*/dt_{max}$ against dP/dt_{max} and $E_{a,max}$ in 30 subjects with disparate ventricular function in Figure 6, and demonstrated the index's load independence, albeit under conditions of limited preload and after load manipulations [3].

For this study, thirty volunteers [mean 58.1 (range 48–77) yr of age, 13:2 male-to-female ratio] with diverse cardiac conditions were recruited. From their LV pressure-volume data, LVEF and dP/dt_{max} were computed directly from these traces. Active elastance Ea at various times was also computed from the pressure-volume loops, from the data in Table 4, based on our earlier work on Ea definition and determination [4]. $E_{a,max}$ was extrapolated from the peak of the Ea-time curve [4]. The single-beat estimation of end-systolic elastance Ees (SB) was determined, using bilinearly approximated time-varying elastance [5].

The patients were divided into three groups on the basis of tertiles of LVEF, with 10 individuals in each group, as shown in Table 5. Intergroup comparisons show significant differences between the mean values of dP/dt_{max}, $E_{a,max}$, Ees(SB), and $d\sigma^*/dt_{max}$ in those in the highest tertile compared with those in lowest and middle tertiles. There is agreement with regard to the index $d\sigma^*/dt_{max}$ with dP/dt_{max}, $E_{a,max}$, and Ees(SB) across the three tertiles of ascending LVEF values, with statistically significant differences in LV contractility indexes among the three groups. Values of dP/dt_{max}, Ees(SB), and $d\sigma^*/dt_{max}$ were statistically significantly lower in patients in the lowest and middle tertiles had than those in the highest tertile.

Figure 6 summarizes the correlation between $d\sigma^*/dt_{max}$, dP/dt_{max}, and $E_{a,max}$, as well as Ees(SB). Linear regression analysis revealed good correlation between $d\sigma^*/dt_{max}$ and dP/dt_{max}, $E_{a,max}$, and Ees(SB), with significant correlation coefficients in each case: $d\sigma^*/dt_{max}$=0.0075dP/dt_{max}- 4.70 (r = 0.88, P < 0.01), $d\sigma^*/dt_{max}$=1.20$E_{a,max}$ + 1.40 (r = 0.89, P<0.01), and $d\sigma^*/dt_{max}$= 1.60Ees(SB) + 1.20 (r = 0.88, P < 0.01). In contrast, the correlation between $d\sigma^*/dt_{max}$and LVEF is less strong (r = 0.71), as is the correlation between Ees (SB) and LVEF (r = 0.78), underscoring the lack of specificity of LVEF as an index of myocardial contractility.

Frame No. *(i)*	*t*, s	P, mmHg	V, ml	*E*a, mmHg/ml
		Isovolumic contraction		
1	0	18	136.7	0
2	0.02	22	135.7	0.0295
3	0.04	32	134.6	0.1038
4	0.06	52	133.5	0.2536
5	0.08	80	132.5	0.4636
		Isovolumic relaxation		
18	0.34	74	85.0	0.0590
19	0.36	50	85.5	0.1778
20	0.38	30	86.4	0.3127
21	0.40	17	90.6	0.4636

Table 4. Active Elastance *Ea computed at discrete time points during isovolumic contraction and relaxation in a sample subject: i*, time instant in the cardiac cycle (frame number from end-diastole); *t*, time from start of isovolumic contraction; *P*, measured left ventricular intracavitary pressure; *V*, measured left ventricular intracavitary volume; $E_{a,i}$, calculated active elastance at instant *i*. This table is related to our work in Ref [4].

	Lowest tertile	Middle tertile	Highest tertile
Ejection fraction	0.38 ± 0.12*	0.49 ± 0.13*	0.63 ± 0.05
Age,yr	58.30 ± 8.86	56.10 ± 6.15	59.90 ± 6.17
Heart rate,beats/min	71.18 ± 10.72	71.77 ± 10.68	71.46 ± 9.09
dP/dt_{max},mmHg/s	960 ± 115*	1,121 ± 113*	1,360 ± 97
*E*a,max, mmHg/ml	0.95 ± 0.32*	1.85 ± 0.59*	3.61 ± 0.62
*E*es(SB),mmHg/ml	0.72 ± 0.26*	1.51 ± 0.20*	2.81 ± 0.51
$d\sigma^*/dt_{max}$,s^{-1}	2.30 ± 0.58*	3.60 ± 1.06*	5.64 ± 1.13

Table 5. *LV contractility indexes classified into tertiles of LVEF:* Left ventricular contractility indices classified into tertiles of left ventricular ejection fraction. Values are expressed as mean ± standard deviation. Asterisks denote statistically significant difference ($p<0.05$) when compared with corresponding values in the highest tertile of left ventricular ejection fraction. dP/dt_{max}, peak first time-derivative of the ventricular pressure; $E_{a,max}$, maximum left ventricular elastance; *E*es(SB), single-beat LV end-systolic elastance; $d\sigma^*/dt_{max}$,left ventricular contractility index. This table is related to our work in Ref [3].

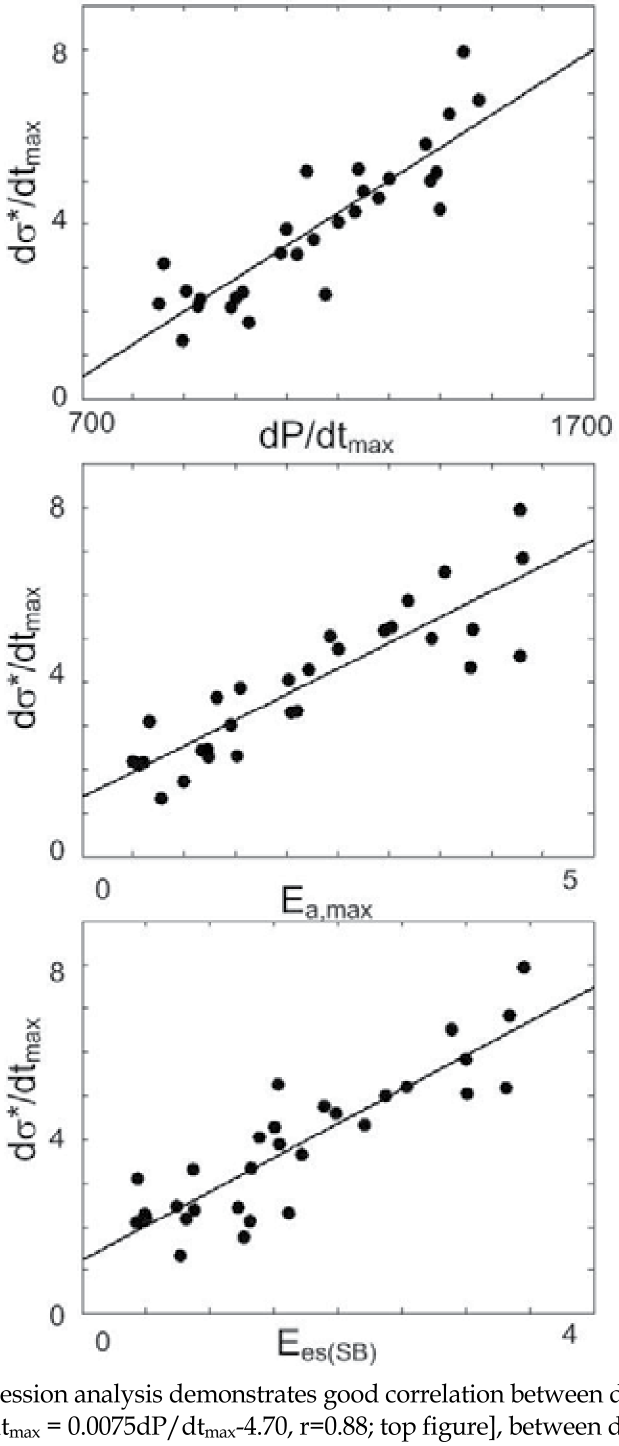


Fig. 6. Linear regression analysis demonstrates good correlation between $d\sigma^*/dt_{max}$ and dP/dt_{max} : [$d\sigma^*/dt_{max} = 0.0075dP/dt_{max}-4.70$, r=0.88; top figure], between $d\sigma^*/dt_{max}$ and $E_{a,max}$ [$d\sigma^*/dt_{max} = 1.20E_{a,max}+1.40$, r=0.89; middle figure], and between $d\sigma^*/dt_{max}$ and $E_{es(SB)}$ [$d\sigma^*/dt_{max} = 1.60E_{es(SB)}+1.20$, r=0.88; bottom figure]. This figure is based on our work in Ref [3].

5.3 Use of cardiac contractility index $d\sigma^*/dt_{max}$ to diagnose heart failure with normal ejection fraction (HFNEF) and with reduced ejection fraction (HFREF)

In another study [6], we assessed the capacity of $d\sigma^*/dt_{max}$ to diagnose heart failure in patients with normal ejection fraction (HFNEF) and with reduced ejection fraction (HFREF).

5.3.1 Introduction

Heart failure (HF) is a major health care burden: it is the leading cause of hospitalization in persons older than 65 years, and confers an annual mortality of 10%. HF can occur with either normal or reduced LV ejection fraction (EF), depending on different degrees of ventricular remodeling. Both heart failure with normal ejection fraction (HFNEF) and heart failure with reduced ejection fraction (HFREF), also commonly known as diastolic and systolic heart failure respectively, have equally poor prognosis [7]. Medical therapy targets to reduce load, by using vasodilators and/or to alter contractile strength using inotropic agents. Alternatively, some therapies target to affect cardiac remodeling, such as passive cardiac support devices, surgical restoration of LV shape (i.e. the Dor procedure), and stem cells therapies.

Assessment of left ventricular (LV) contractility is important for HF management and evaluation of the heart's response to medical and surgical therapies. Although approaches based on pressure-volume analysis, stress-strain analysis, and dP/dt_{max}-EDV relations [8]can provide assessments of contractile function, these relations generally require invasive data measured at several chamber loads and thus are difficult to apply in routine or long-term clinical studies. This is an important limitation, because heart failure often requires longitudinal evaluation. The ideal measure of contractility should have the following characteristics: sensitivity to inotropic changes, independence from loading conditions as well as heart size and mass, ease of application, and proven usefulness in the clinical setting.

LV ejection fraction (EF) is the index overwhelmingly used to assess cardiac function in both clinical and experimental studies, despite the fact that it is highly dependent upon preload and afterload. Based on the National Heart Lung and Blood Institute's Framingham Heart Study, an LVEF50% as cut-off for the presence of normal LVEF has been used in the present study [9].

Usefulness of $d\sigma^*/dt_{max}$ as Contractility Index: During LV systole, LV wall stress is generated intrinsically by sarcomere contraction and results in the development of extrinsic LV pressure. We have shown earlier that our novel LV contractility index, $d\sigma^*/dt_{max}$ (maximal change rate of pressure-normalized wall stress) correlates well with LV dP/dt_{max} [3]. We have proposed and validated a new LV contractility index, $d\sigma^*/dt_{max}$,based on the maximal rate of development of LV wall stress with respect to LV pressure. From the right-hand side of equation (8), this index is also seen to represent the maximal flow rate from the ventricle (cardiac output) normalized to myocardial volume (or mass).

This index is easily measured non-invasively (i.e. from echocardiography or magnetic resonance imaging), is sensitive to LV inotropic changes, and has been demonstrated by us to be preload and afterload independent [3]. Importantly, it is measured at a single steady-state condition, as opposed to the multiple variably loaded cardiac cycles required for many of the other indices. Thus $d\sigma^*/dt_{max}$has several qualities that make it a useful LV contractility index. This study [6] has constituted an important step toward establishing the clinical utility of $d\sigma^*/dt_{max}$ as a tool for diagnosis of HF (both HFNEF and HFREF) as well as for follow-up surveillance of LV function. Hence we see a great potential for application of this novel index to evaluate heart function in diverse heart conditions.

5.3.2 Clinical methodology

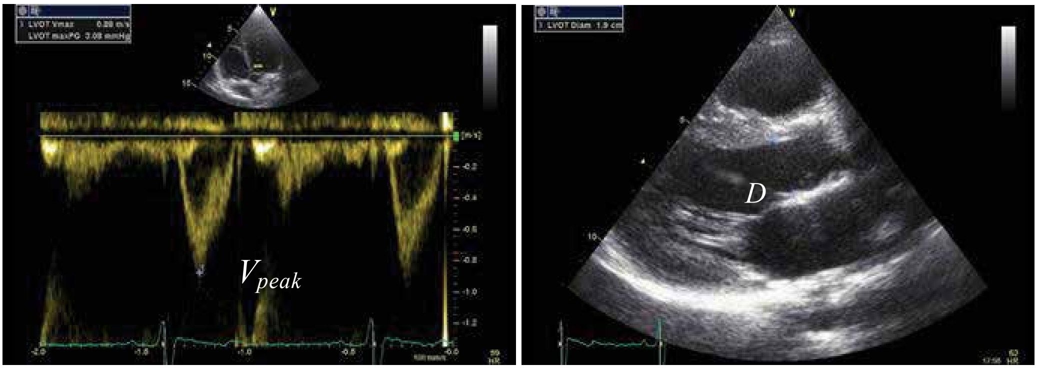
Patients referred to our echocardiography service with symptoms and signs of heart failure underwent echocardiography and electrocardiography (ECG). Patients with atrial fibrillation, more than mild mitral or aortic valvular regurgitation, and unsatisfactory echocardiographic images were excluded. Clinical signs of heart failure were defined as presence of at least one of the following: raised jugular venous pressure, peripheral pedema, hepatomegaly, basal inspiratory crepitation or gallop rhythm. Patients with LVEF ≥ 50% and LVEF < 50% on echocardiography were classified into HFNEF and HFREF, respectively.

Echocardiography Study: With the subject in the left lateral decubitus position, 2D examinations, M-mode measurements and Doppler recordings were performed from the standard left parasternal long- and short-axis as well as the apical four chamber views with simultaneous ECG. The LVEF was assessed by using a 2-dimensional method by an experienced observer; normal LVEF was defined as greater or equal to 50%. Mitral flow velocities were obtained from the apical 4-chamber view using pulsed wave Doppler technique with the sample volume at the tips of the corresponding valve leaflets. LV outflow tract velocity was obtained from apical 5-chamber view, using pulsed wave Doppler technique with the sample volume at the aortic valve level.

The measurements included peak E (peak early trans-mitral filling velocity during early diastole) and A (peak trans-mitral atrial filling velocity during late diastole); wave velocities (cm/s) were measured and E/A ratio was calculated. The E wave deceleration time (DT) was also calculated as the time elapsed between peak E velocity and the point where the extrapolation of the deceleration slope of E velocity crosses the zero baseline measured in milliseconds. LVOT maximal velocity V_{peak} was measured, and LV mass was calculated by using ASE methods [10, 11]. Myocardial tissue Doppler (TDI) velocities were also estimated at the atrioventricular ring, septal positions, in the apical 4 chamber view. All measurements were averaged over two or three cardiac cycle.

Calculation of $d\sigma^*/dt_{max}$ *from Echocardiography:* The contractility index was computed by the above equation (8). M-mode echocardiographic measurements of the LV were obtained, and LV mass calculated using standardized methodology [10, 11]. Myocardial volume was calculated by dividing LV mass with myocardial density (assumed to be 1.05 g/ml). Furthermore, two-dimensional apical four- and two-chamber views of the LV were acquired, and end-diastolic and –systolic endocardial contours were manually outlined. The corresponding LVEDV were then automatically determined using biplane Simpson's method.

From Pulse-wave echo-Doppler interrogation of the LV outflow tract (LVOT), we calculated (in the absence of significant mitral regurgitation or aortic valve dysfunction) the maximal LV volume rate (dV/dt_{max}) during ejection: $dV/dt_{max}=V_{peak}$*AVA, where V_{peak} is the peak velocity sampled at the LVOT and AVA is the aortic valve area (= $\pi D^2/4$, where D is the LVOT diameter measured in the two-dimensional parasternal long-axis image of the heart), as shown in Figure 7. Upon substituting values of myocardial volume and dV/dt_{max} into equation (8), we determined the value of $d\sigma^*/dt_{max}$.



(a) (b)

Fig. 7. Echocardiographic measurement on (a) peak velocity V_{peak} sampled at the LVOT and (b) LVOT diameter D measured in the two-dimensional parasternal long-axis image of the heart. This figure is adopted from Ref. [6].

Clinical Studies: The study involved 26 age- and sex-matched subjects in each of the groups of normal controls, HFNEF and HFREF. The characteristics of 78 subjects are shown in Table 6. It summarizes the subjects' age, BSA, LVEF, peak E, peak A, E/A ratio, DT, heart rate (HR), septal E/E′, lateral E/E′ and our index $d\sigma^*/dt_{max}$. Mean $d\sigma^*/dt_{max}$ was 3.91 s^{-1} (95%CI, 3.56-4.26 s^{-1}) in control subjects; it was reduced in heart failure, HFNEF, to 2.90 s^{-1} (95%CI, 2.56-3.24 s^{-1}); and in HFREF, to 1.84 s^{-1} (95%CI, 1.60-2.07 s^{-1}). There exists no substantially difference between the average values of LVEF, peak E, peak A, E/A ratio, DT, heart rate (HR), septal E/E′, and lateral E/E′ in HENEF compared to normal controls, except for $d\sigma^*/dt_{max}$ (2.90 ± 0.84 vs. 3.91 ± 0.87, $p<0.001$). However, there exists significant difference between the average values of LVEF, peak E, peak A, E/A ratio, DT, septal E/E′, lateral E/E′ and $d\sigma^*/dt_{max}$ in HEREF compared to HFNEF.

Discussion on the Usefulness of $d\sigma^*/dt_{max}$: During LV systole, LV wall stress is generated intrinsically by sarcomere contraction and results in the development of extrinsic LV pressure. LV wall stress is dependent on wall thickness, LV geometry and chamber pressure and sarcomere contraction. Hence, it is rational to quantify the LV wall stress as an intrinsic measure of myocardial contractility. We have proposed and validated a new LV contractility index, $d\sigma^*/dt_{max}$,based on the maximal rate of development of LV wall stress with respect to LV pressure. From the right-hand side of equation (8),this index is also seen to represent the maximal flow rate from the ventricle (cardiac output) normalized to myocardial volume (or mass).

Assessment of heart failure with normal ejection fraction (HFNEF) and reduced ejection fraction (HFREF): Heart failure may be viewed as a progressive disorder that is initiated after an "index event" with a resultant loss of functioning cardiac myocytes, thereby preventing the heart from contracting normally. HF can occur with either normal or reduced LV ejection fraction (LVEF), depending on different degree of ventricular remodeling. Perhaps 50% of patients with heart failure have a normal or minimally impaired LVEF (HFNEF) [12, 13].

Although mechanisms for HFNEF remain incompletely understood, diastolic dysfunction is said to play a dominant role: impaired relaxation, increased passive stiffness, raised end-diastolic pressure (EDP) [14]. The diagnostic standard for HFNEF is cardiac catheterization, which demonstrates increased EDP. However, a more practical noninvasive alternative is echocardiography. Our study has shown that E/A ratio (1.26 ± 0.90 vs. 0.96 ± 0.38, $p<0.05$) and DT (157 ± 41 ms vs. 214 ± 47 ms, $p<0.05$) are significantly different between HFREF and normal controls, and not so between HFNEF and normal controls (Table 6).

Our contractility index, of change rate of normalized wall stress index $d\sigma^*/dt_{max}$, is dependent on lumen and wall volume of LV chamber and represents an integrated assessment of LV systolic performance [3], based on our findings relating $d\sigma^*/dt_{max}$ with HFNEF and HFREF [6]. In this study, as shown in Table 7, we find that there exists significant difference in dV/dt_{max} between HFREF and HFNEF (233±48 ml/s vs. 355±65 ml/s, $p<0.05$), while there exists no difference between HFNEF and normal controls (355±65 ml/s vs. 353±80 ml/s, *NS*). Similarly, there exists significant difference in LV mass between normal controls and HFNEF (147±41 g vs. 202±47 g, $p<0.05$), while there is no difference between HFREF and HFNEF (213±60 g vs. 202±47 g, *NS*).

Our $d\sigma^*/dt_{max}$ index, using dV/dt_{max} normalized with LV mass, can clearly differentiate HFREF, HFNEF and normal controls ($p<0.05$) (Table 7). The average value of $d\sigma^*/dt_{max}$ decreases in HFNEF and HFREF, in relation to normal controls. The mean value of $d\sigma^*/dt_{max}$ was found to be 3.91 s^{-1} (95%CI, 3.56-4.26 s^{-1}) in control subjects; the index was reduced in heart failure patients: in HFNEF, to 2.90 s^{-1} (95%CI, 2.56-3.24 s^{-1}) and in HFREF, to 1.84 s^{-1} (95%CI, 1.60-2.07 s^{-1}). This suggests that poor systolic function of LV is associated with lower $d\sigma^*/dt_{max}$ values. Therefore, it can again be concluded that $d\sigma^*/dt_{max}$ is an appropriate index for representing assessment of LV contractile function in heart failure with/without preserved LV ejection fraction.

	Controls (n=26) (n=21)	HFNEF (n=26)	HFREF (n=26)
Age (years)	72 ± 8	70 ± 8	70 ± 8
Gender (male:female)	16:10	16:10	16:10
BSA (m^2)	1.69 ± 0.20	1.71 ± 0.20	1.61 ± 0.20
LVEF (%)	68.3 ± 5.1	66.5 ± 4.9	33 ± 13.7$^{\S *}$
E/A ratio	0.96 ± 0.38	0.78 ± 0.24	1.26 ± 0.90*
DT (ms)	214 ± 47	214 ± 67	157 ± 41$^{\S *}$
HR (beats/min)	66 ± 10	72 ± 16	80 ± 11§
Septal E/E′	8.48 ± 2.10	9.79 ± 3.29	13.68 ± 4.78$^{\S *}$
Lateral E/E′	6.64 ± 1.55	8.39 ± 2.76	10.28 ± 3.40$^{\S *}$
$d\sigma^*/dt_{max}$ (s^{-1})	3.91 ± 0.87	2.90 ± 0.84$^{\S *}$	1.84 ± 0.59$^{\S *}$

A, mitral atrial flow velocity on echo-Doppler; BSA, body surface area; DT, mitral E deceleration time; E, mitral early velocity; E′, septal mitral annular myocardial velocity on tissue Doppler imaging; HR, heart rate, LVEF, let ventricular ejection fraction.

The values are expressed as mean ± SD. § and * denote statistically significant difference of HF compared to controls, HFREF compared to HFNEF patients, respectively (Bonferroni pairwise test, p value <0.05)

Table 6. Patients characteristics and echocardiographic measurements in Group 1 (Controls), Group 2 (HFNEF) and Group 3 (HFREF). This table is related to our work in Ref [6].

	Controls (95% CI)	HFNEF (95% CI)	HFREF (95% CI)
dV/dt_{max} (ml/s)	353 (320, 385)	355 (329, 381)	233 (213, 252)§*
V_{peak}(cm/s)	106 (98, 115)	112 (104, 119)	73 (68, 78) §*
LV mass (g)	147 (131, 164)	202 (183, 221)§	213 (189, 297)§
$d\sigma^*/dt_{max}$ (s^{-1})	3.91 (3.56, 4.26)	2.90 (2.56, 3.24)§*	1.84 (1.60, 2.07)§*

§ and * denote statistically significant difference of HF compared to controls, HFREF compared to HFNEF patients, respectively

Table 7. Comparison of the maximal flow rate dV/dt_{max}, V_{peak}, LV mass, and $d\sigma^*/dt_{max}$ in Group 1 (Controls), Group 2 (HFNEF) and Group 3 (HFREF). This table is related to our work in Ref [6].

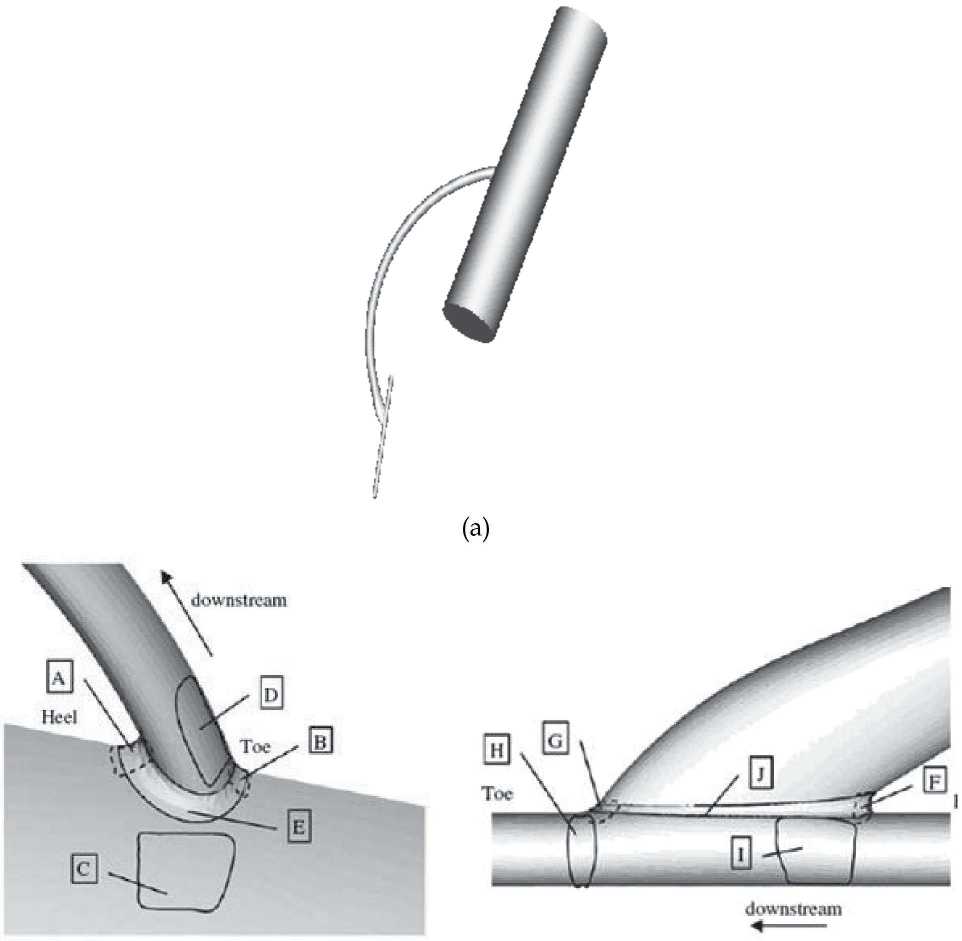
6. Coronary Arterial Bypass Grafting (CABG) to salvage ischemic myocardial segments

As is well known, coronary artery bypass graft (CABG) surgery has been the standard treatment for serious blockages in the coronary arteries and for re-perfusing myocardial ischemic segments to restore them to normal contractile state. During the surgery, one end of the graft is sewn to the aorta (or its subsidiary branches) to create proximal anastomosis, while the other end is attached to coronary artery below the area of blockage to create distal anastomosis. In this way, the oxygen-rich blood is taken directly from the aorta, bypasses the obstruction, and flows through the graft to perfuse and nourish the heart muscle. The most commonly used graft is the saphenous vein. Besides this vein graft, some arterial conduits (such as internal mammary artery, gastroepiploic artery and radial artery etc), and synthetic veins (such as Dacron, Teflon and Polytetrafluoroethylene-PTFE veins) are also suitable for CABG.

Although the number of bypass operations keeps increasing, the CABG has not been without complications. Approximately 15% to 20% vein grafts occlude in the first year, and 22.5% to 30% occlude within the first 2 years. At 10 years, approximately 60% of vein grafts are patent; only 50% of these vein grafts remain free of significant stenosis. In order to intensively investigate the coronary arterial stenosis symptom, numerous research works have been carried out. One direction of these studies is to investigate the pathogenic mechanism of bypass graft failure. In this regard, vascular injury and biomechanical factors (such as wall shear stress related factors, compliance mismatch, etc.) are believed to stimulate cellular responses for pathological changes. In particular, hemodynamic flow patterns of CABG have considerable relevance to the causes and sites of pathogenesis. Hence, we have carried out simulation of hemodynamic flow patterns in CABG models, to look into the hemodynamic causes and mechanisms of lesions in coronary bypass grafts.

The flow characteristics and hemodynamic parameters distributions in a complete CABG model (as shown in Figure 8) have been investigated computationally by us [15]. It is found that disturbed flow (flow separation and reattachment, vortical and secondary flows) patterns occur at both proximal and distal anastomoses, especially at the distal anastomosis. In addition, regions of high-OSI & low-WSS and low-OSI & high-WSS are found in the proximal and distal anastomoses, especially at the toe and heel regions of distal anastomosis. These regions are suspected to initiate the atherosclerotic lesions and are further worsened by the increasing permeability of low-density lipoprotein as indicated by high WSSG. The comparisons of segmental average of HPs (in the Table of Figure 8) further imply that intimal hyperplasia is more prone to form in the distal anastomosis than the proximal anastomosis, especially along the suture line at the toe and heel of distal anastomosis, which was in line with the in-vivo observations.

We then investigated the fluid dynamics of blood flow in two complete models of CABG for the right and left coronary artery separately, as shown in Figure 9 [16]. The results reveal that blood flow through the coronary artery bypass graft primarily occurs only during the diastolic phase of the cardiac cycle, which is in agreement with the physiological observation. However, at the onset of ejection, some backflow from the coronary artery into the bypass graft is found for the CABG to left coronary artery, which is absent for the right coronary artery. This reversal of flow during systole can be explained by the predominant intra-cardiac course of the left coronary artery system. As the same time, this study also found a low WSS region near the heel and a high WSS in the toe region of the anastomosis domain.


(a)

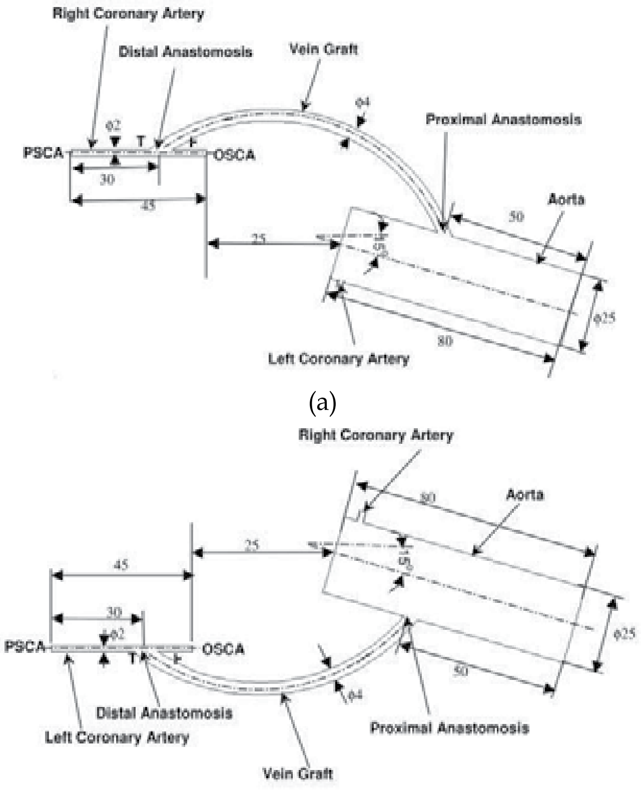
(b)

Fig. 8. (a) The configuration to mimic complete CABG; (b) sketch maps of areas investigated for segmental averages of hemodynamic parameters (HPs) in proximal and distal anastomoses;

Name		Labels in maps	Area (mm^2)	<WSS> (Pa)	<WSSG>	<OSI>
Proximal anastomosis	heel	A	4.23	4.48	11.88	0.41
	toe	B	1.59	8.58	20.24	0.36
	part3	C	9.46	3.17	3.24	0.49
	part4	D	25.50	6.56	14.25	0.07
	suture_line	E	12.00	5.06	9.01	0.29
Distal anastomosis	heel	F	0.55	2.92	12.75	0.24
	toe	G	0.55	**28.04**	**147.17**	**0.02**
	part3	H	9.70	**14.63**	**64.24**	**0.07**
	part4	I	6.81	0.85	3.59	0.22
	suture_line	J	3.18	**9.20**	**41.43**	**0.11**

(c)

Fig. 8. (c) the segmental averages of HPs at these locations. This figure is adopted from our work in Ref 15 (Zhang et al., 2008).

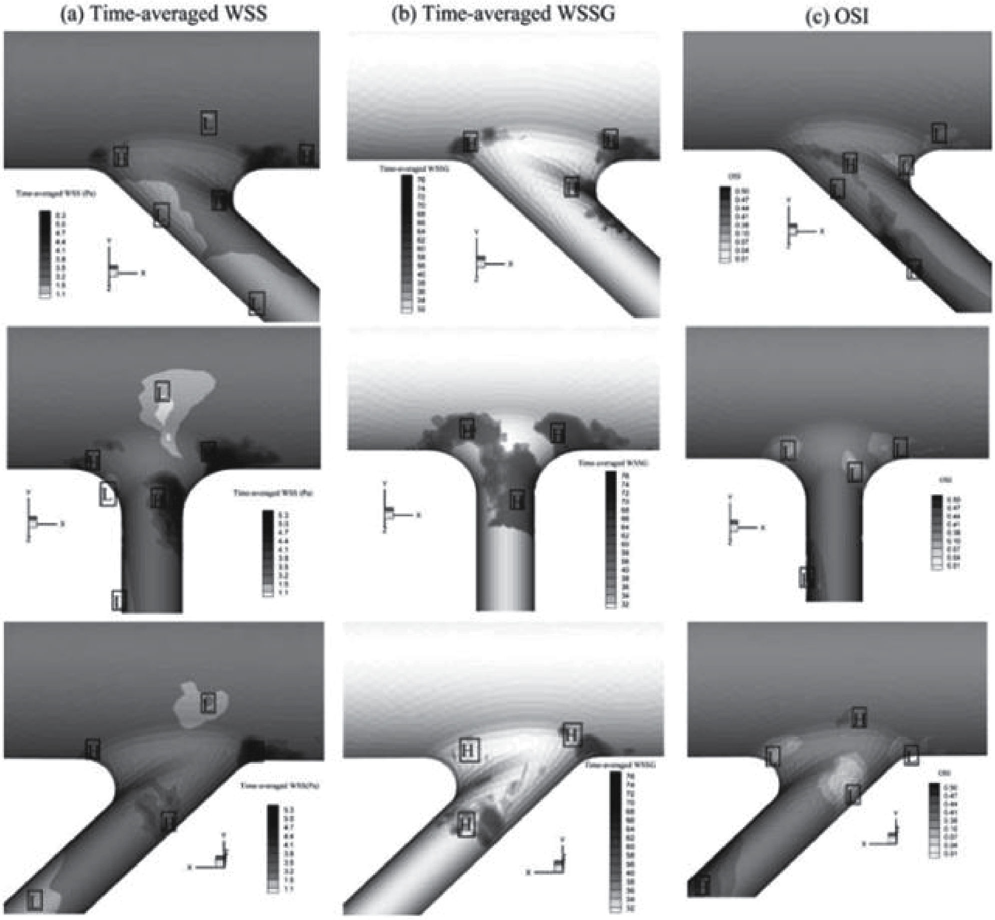

Right Coronary Artery
Distal Anastomosis
Vein Graft
φ4
Proximal Anastomosis
φ2
T
H
PSCA
OSCA
30
45
25
50
Aorta
φ15
φ25
80
Left Coronary Artery

(a)

(b)

Fig. 9. Geometry (plane view) and dimensions (in mm) of the bypass models of: (a) The aorta-right coronary artery bypass model; (b) The aorta-left coronary artery bypass model (PSCA-Perfused Segment of the Coronary Artery; OSCA-Occluded Segment of the Coronary Artery; T-Toe; H-Heel). This figure is adopted from our work in Ref 16: (Sankaranarayanan et al., 2005).

CABG performance is based on flow characteristics at both the proximal and distal anastomoses. So then, let us summarise the optimal geometrical parameters for proximal and distal anastomoses. For proximal anastomosis, a detailed study, on the effect of three grafting angles (viz. 45° forward facing, 45° backward facing, and 90°), has been carried out by Chua et al. [17]. The results show a flow separation region along the graft inner wall immediately after the heel at peak flow phase, which decreases in size with the grafting angle shifting from 45° forward facing to 45° backward facing. The existence of nearly fixed stagnating location, flow separation, vortex, high-WSS-low-OSI, low-WSS-high-OSI, and high WSSG is suspected to lead to graft stenosis. Among these three models, the 45° backward-facing graft is found to have the lowest variation range of time-averaged WSS and the lowest segmental average of WSSG, as shown in Figure 10; these parameters are then recommended for obtaining higher expected patency rates in bypass operations.


(a) Time-averaged WSS

(b) Time-averaged WSSG

(c) OSI

Fig. 10. Contours of HPs: (a) time-averaged WSS; (b) time-averaged WSSG; and (c) OSI, on the surfaces of 45° forward-facing, 90°, and 45° backward-facing models ('H' means the region has high value, 'L' means the region has low values. This figure is adopted from our work in Ref 17 (Chua et al., 2005b).

At the distal anastomosis junction of CABG, the important geometrical parameters for smooth flow are the angle (α) and the diameter ratio (ϕ) between the graft and host artery. The hemodynamics associated with these parametric were investigated by Xiong and Chong [18], over a range of ϕ (1:1, 1.5:1 and 2:1) and α (15°, 30°, 45° and 60°) in physiological coronary flow conditions. It is found that increasing ϕ from 1:1 to 1.5:1 almost eliminates both low WSS and high OSI at the toe. However, further increasing ϕ to 2:1 causes elevated OSI for most of the host artery segment in the anastomotic region and in almost the entire graft. On the other hand, varying α is also found to change certain aspects of hemodynamics, although less than those changes related with different values of ϕ. A smaller value of α is found to be associated with a higher OSI in the anastomotic region, whereas a larger α causes higher WSSG on the artery floor. Therefore, it is suggested that for distal coronary anastomosis (with a 20:80 proximal to distal flow division ratio maintained in the host artery), the geometry associated with $\phi = 1.5$ and $\alpha = 30–45°$ is favorable for enhancing long-term performance.

7. Surgical Ventricular Restoration (SVR), combined with CABG, restores LV shape and improves cardiac contractility

7.1 Introduction

Ischemic heart disease is one of the most widely spread, progressive and prognostically unfavorable diseases of the cardiovascular system. In ischemic dilated cardiomyopathy (IDC) patients, the remodeling process involves a lesser systolic LV curved shape, increase of peak wall stress, and decrease of contractile functional index, compared with normal subjects. Surgical ventricular restoration (SVR) is performed in chronic ischemic heart disease patients with large non-aneurysmal or aneurysmal post-myocardial infarction zones. It involves operative methods, that reduce LV volume and 'restore' ventricular ellipsoidal shape, by exclusion of anteroseptal, apical, and anterolateral LV scarred segments by means of intra-cardiac patch or direct closure.

For patients in heart failure (HF) resulting from serious myocardial diseases of ischemic dilated cardiomyopathy and myocardial infarction (MI), surgical ventricular restoration (SVR), designed to restore the LV to its normal shape (reversal of LV remodeling), is performed usually in conjunction with coronary artery bypass grafting (CABG). In our study [19], in 40 ischemic dilated cardiomyopathy (IDC) patients who underwent SVR and CABG, there was found to be: (i) decrease in end-diastolic volume from 318 ± 63 ml to 206 ± 59 ml ($p<0.01$) and in end-systolic volume from 228 ± 58 ml to 133 ± 61 ml ($p<0.01$), (ii) increase in LV ejection fraction from 26 ± 7% to 31 ± 8% ($p<0.01$), (iii) decrease in LV mass (from 204 ± 49 g to 187 ± 53 g, $p<0.01$), (iv) decrease in peak normalized wall stress (PNWS) (from 4.30 ± 0.95 to 3.31 ± 0.75, $p<0.01$) , (v) increase in end-systolic sphericity index SI (from 0.57 ± 0.094 to 0.67 ± 0.13, $p<0.01$), (vi) increased value of shape (S) index (from 0.44 ± 0.085 to 0.54 ± 0.089, $p<0.01$) during end-systole indicating that LV became more spherical after SVR, and most importantly (vii) improvement in LV contractility index $d\sigma^*/dt_{max}$ (from 2.69 ± 0.74 s^{-1} to 3.23 ± 0.73 s^{-1}, $p<0.01$).

Thus, in IDC patients, surgical ventricular restoration (in combination with CABG) aiming to reverse LV remodeling, has shown to (i) improve ventricular function and decrease wall stress, along with making a more curved apex, and (ii) improve cardiac contractility. It is not the LV shape alone that defines LV contractility. Rather, a complex interaction of the rate of change of shape factor (dS/dt_{max}) along with LV maximal flow rate and LV mass may explain the improvement in LV contractility.

7.2 Clinical study

The study was carried to retrospectively evaluate (with cardiac MRI) the changes on systolic function and LV wall stress, the relationships between LV geometry (shape) and dimensions and systolic function after SVR performed in chronic ischemic heart disease patients with aneurismal postmyocardial infarction zones. The study consisted of 40 patients with ischemic dilated cardiomyopathy who had SVR; the age of the patients averaged 69 years (range, 52-84 years). MRI scans were performed 2 weeks before surgery (pre-surgery) and 1 week after the surgery; the details of the MRI procedure are reported in our earlier work [19].

Cardiac magnetic resonance imaging (MRI) provides the means to study heart structure and function: the ventricular systolic and diastolic volumes (and hence ejection fraction) are easily assessed reproducibly and accurately; the regional wall motion of the asynergy area and the remote myocardium can be measured by several quantitative means, including with myocardial tagging; the presence or absence of nonviable, irreversible scar can be detected with gadolinium-based interstitial contrast agents.

Data Analysis, 3-dimensional modeling of LV: For analysis, the images were displayed on a computer monitor in a cine-loop mode using CMRtools, to reconstruct the 3-dimensional model of the left ventricle (LV). The LV epicardial and endocardial borders were outlined, and all the frames were delineated to produce a volume curve from end-diastolic and end-systolic phases. These measurements were used to determine the end-diastolic volume (EDV), end-systolic volume (ESV), stroke volume (SV), ejection fraction (EF), and LV mass.

Ellipsoidal Shape factor, Eccentricity (E) and sphericity index and normalized wall stress: The LV is modeled as a prolate spheroid, truncated 50% of the distance from equator to base, as shown in figure 11 [20, 21]. Then, the left ventricular cavity wall volume is calculated, from the endocardial anterior-posterior (AP) and base-apex (BA) lengths [20], as:

$$V_m = \frac{9}{8}\left[\left(\frac{BA}{1.5}+h\right)\left(\frac{AP}{2}+h\right)^2 - \left(\frac{BA}{1.5}\right)\left(\frac{AP}{2}\right)^2\right] \tag{9}$$

wherein the BA and AP dimensions are identified in figure 12. The mean wall thickness (h) is calculated at each cavity volume, from the above equation, by assuming that myocardial wall volume (V_m) remains constant throughout the cardiac cycle. The endocardial minor axis dimension (SA) and major axis dimension (LA), shape factor (S), eccentricity (E) and sphericity index (SI) were then calculated as follows (refer figures 11 and 12):

$$\mathrm{SA} = \mathrm{AP}/2;\ \mathrm{LA} = \mathrm{BA}/1.5,\ \mathrm{S} = \mathrm{SA}/\mathrm{LA},\ E = \left(\frac{BA^2 - AP^2}{BA^2}\right)^{0.5},\ \mathrm{SI} = \mathrm{AP}/\mathrm{BP} \tag{10}$$

wherein BA (the LV long axis) is defined as the longest distance from the apex to the base of the LV (defined as the mitral annular plane), as measured on the four-chamber cine MRI view of the heart; AP is defined as the widest LV minor axis (Figure 12). A small value of SI implies an ellipsoid LV, whereas values approaching "1" suggest a more spherical LV. The SI at end-diastole (SI_{ed}) and end-systole (SI_{es}), the % shortening of the long and minor axes, as well as the difference between end-diastolic and end-systolic values of SI, (SI_{ed} - SI_{es}) were calculated and are tabulated in Table 10 below.

The time-varying circumferential normalized wall stress, NWS(t), is calculated from the instantaneous measurements of LV dimensions and wall thickness, by treating the LV as a prolate spheroid model truncated 50% of the distance from equator to base [20, 21]

$$NWS(t) = \frac{AP(t)}{2h(t)}\left[1 - \frac{\frac{9AP(t)}{32h(t)}(SI)^2}{\frac{AP(t)}{h(t)}+1}\right] \tag{11}$$

The LV wall thickness, h(t), is calculated from the following formula (based on the above equation 9), by assuming that the myocardial wall volume (V_m) remains constant throughout the cardiac cycle:

$$\frac{9}{8}\left[\left(\frac{BA(t)}{1.5}+h(t)\right)\left(\frac{AP(t)}{2}+h(t)\right)^2\right] = V_m(t) + \left(\frac{BA(t)}{1.5}\right)\left(\frac{AP(t)}{2}\right)^2 \tag{12}$$

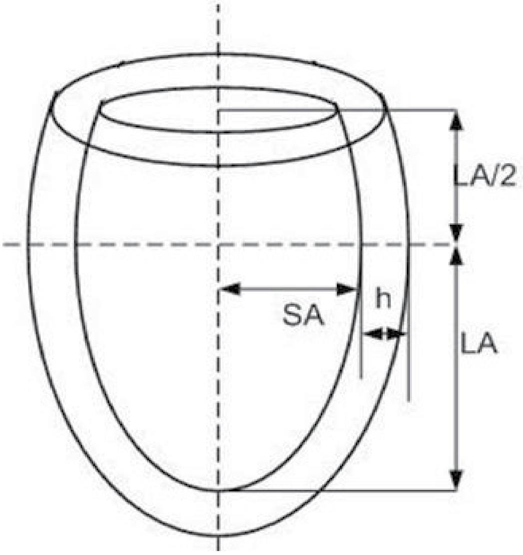
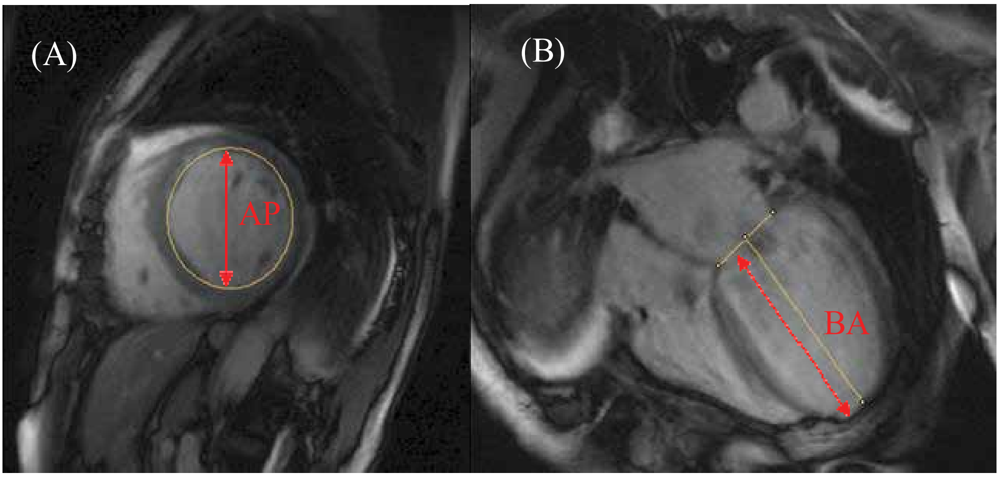


Fig. 11. LV model geometry, showing the major and minor radii of the inner surface of the LV (LA & SA) and the wall-thickness (h).

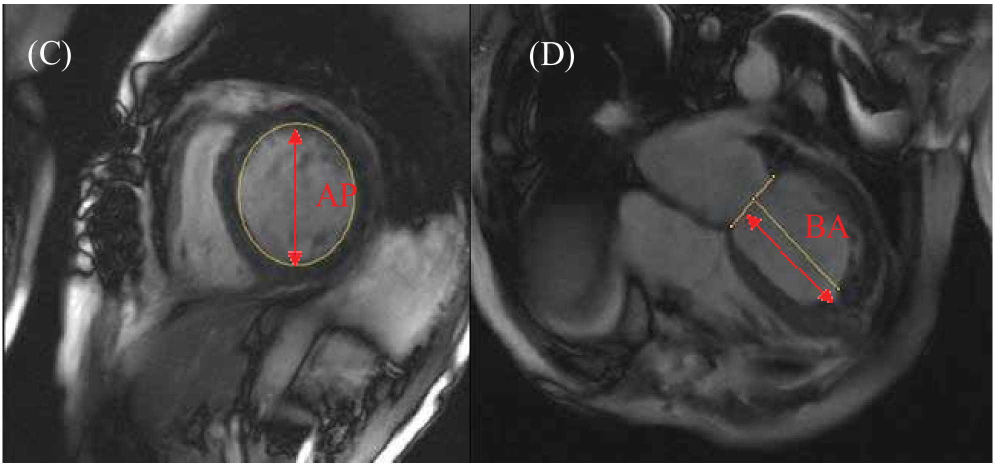
Cardicac Contractility *$d\sigma^*/dt_{max}$*: In order to compute the contractility index by employing the above equation (8), a 6-order polynomial function to curve-fit the volumes-time data to calculate the volume rate (*dV/dt*) by differentiating it. Then the contractility index $d\sigma^*/dt_{max}$ is calculated as:

$$d\sigma^*/dt_{\max} = \left|\frac{d(\sigma_\theta / P)}{dt}\right|_{max} = \frac{3}{2V_m}\left|\left(\frac{dV}{dt}\right)\right|_{max}$$

where V_m is myocardial volume at the end-diastolic phase.



(a) Pre-SVR



(b) Post-SVR

Fig. 12. Short-axis (panels A and C) and long-axis (panels B and D) magnetic resonance images of patients, before (panels A and B) and after (panels C and D) surgical ventricular restoration (SVR). Multiple short-axis cines from the apex to the base of the heart (or orientated axial) and long-axis cines are used to quantify LV function. Anterior-posterior (AP) (panels A and C) and base-apex (BA) (panels B and D) were measured from 2-D CMR imaging before (panels A and B) and after (panels C and D) SVR during cardiac cycle. The shape factor S, eccentricity index and sphericity index (SI) were calculated from Equation (6). It can be noted that the long-axis decreased more dramatically compared with the short-axis dimension, thereby producing a more spherical ventricle. This figure is adopted from our work presented in Ref 19.

7.3 Clinical results

All 40 patients were treated with CABG and SVR (endoventricular circular patch plasty). The age of the patients averaged 69 years (range, 52-84 years). Among them, 19 patients had severe mitral regurgitation and received additional MVS. 28 patients had CHF. The baseline patient characteristics are summarized in Table 8.

Variables	Value
Male : female	36:4
Age (years)	69 ± 9
Body surface area (m^2)	1.98 ± 0.18
Coronary artery disease	39 (98%)
Hypertension	19 (48%)
Diabetes mellitus	13 (33%)
Tobacco	23 (58%)
Congestive heart failure	28 (70%)
Peripheral arterial disease	3 (8%)
Stroke	2 (5%)
Creatinine (mg/dL)	1.17 ± 0.29
Prior cardiac surgery	19 (48%)
New York Heart Association class	
I-II	26 (65%)
III-IV	14 (35%)
Surgery	
Surgical ventricular restoration + coronary artery bypass grafting	21 (52%)
Surgical ventricular restoration + coronary artery bypass grafting + mitral valve surgery	19 (48%)
Values are mean ± SD or numbers of patients (percentages).	

Table 8. Patients' characteristics and clinical data (n=40). This table is related to our work in Ref [19].

7.4 Left ventricular functional indexes changes pre and post-surgery (MRI parameters)

Figure 12 shows typical short-axis and long-axis magnetic resonance images of patient pre- and post-SVR and how to calculate shape, eccentricity and sphericity index. It is noted that the long-axis decreased more dramatically compared with short-axis dimension, thereby producing a more spherical LV. Figure 13 shows typical 3-dimensional modeling of LV from CMR images using LVtools pre and post SVR.

The intraobserver and interobserver data for EDV, ESV and mass for pre- and post-surgery groups are shown in Table 9. Table 10 summarizes the mean LV functional indexes pre and post-SVR. Following SVR, there was a significant decrease in the dimensions of both the long- and short-axes of the LV. However, the long-axis dimension of the LV decreased more than the short-axis dimension, resulting in a more spherical ventricle post-SVR. There was a significantly reduction in end-diastolic volume index (EDVI), end-systolic volume index (ESVI), LV stroke volume index (SVI), LV mass index, and peak normalized wall stress after

SVR (Table 3). The values of LV EDVI, ESVI, LVEF and the contractility index $d\sigma^*/dt_{max}$ pre- and post-SVR are also shown in the scatter plots of Figure 14.

Table 10 provides the sphericity index (SI) values in end-diastole and end-systole and its diastolic-systolic change, as well as the % shortening of the long- and short-axes. During a cardiac cycle, LV shape becomes less spherical in systole (SI smaller) than in diastole (SI closer to '1'). The diastolic-systolic change in SI (SI_{ed}-SI_{es}) is significantly augmented by the operation, despite the LV chamber becoming more spherical. The % shortening of long-axis is not significantly altered, but the % shortening of the short-axis is significantly increased by the operation. Despite the seemingly unfavorable spherical LV shape post-SVR, the LV contractile function is significantly improved, as indicated by the increased value of $d\sigma^*/dt_{max}$.

The scatter plots of figure 14 graphically illustrate pre- and post-SVR values of ventricular end-diastolic volume, end-systolic volume, LVEF and contractility index $d\sigma^*/dt_{max}$. From Tables 9 and 10, we can note a significant reduction in end-diastolic volume (318 ± 63 ml vs. 206 ± 59 ml, $p<0.01$), end-systolic volume (228 ± 58 ml vs. 133 ± 61 ml, $p<0.01$), LV mass (204 ± 49 g vs. 187 ± 53 g, $p<0.01$), and peak normalized wall stress (PNWS) (4.64 ± 0.98 vs. 3.72 ± 0.87, $p<0.01$). Increased sphericity index SI (0.57 ± 0.094 vs. 0.67 ± 0.13, $p<0.01$) and increased shape factor (S) (0.44 ± 0.085 vs. 0.54 ± 0.089, $p<0.01$) during end-systole indicates that the LV became more spherical after SVR.

	Pre-SVR			Post-SVR		
	End diastolic volume (ml)	End systolic volume (ml)	LV mass (g)	End diastolic volume (ml)	End systolic volume (ml)	LV mass (g)
Intraobserver						
Mean	318 ± 63	228 ± 58	204 ± 49	206 ± 59	133 ± 61	187 ± 53
Mean difference	1.1 ± 8.60	-1.4 ± 8.38	-5.1 ± 7.96	0.3 ± 6.94	-1.5 ± 3.20	0.6 ± 8.51
Correlation coefficient	0.99	0.99	0.99	0.99	0.99	0.99
t-test p	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
% variability	1.50 ± 1.99	2.13 ± 2.45	3.74 ± 3.71	2.42 ± 2.32	1.79 ± 1.77	4.09 ± 3.02
Interobserver						
Mean	318 ± 65	231 ± 61	206 ± 52	207 ± 60	135 ± 62	189 ± 51
Mean difference	0.3 ± 10	7.9 ± 10	7.6 ± 14	2.0 ± 8.9	5.5 ± 10	3.1 ± 9.8
Correlation coefficient	0.99	0.99	0.97	0.99	0.99	0.98
t-test p	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
% variability	2.73 ± 1.58	4.11 ± 2.93	5.90 ± 3.13	3.44 ± 2.94	5.87 ± 6.23	4.81 ± 2.49

N.S., not significant.
Data are mean ± SD.

Table 9. Reproducibility data in patients pre- and post-SVR. This table is related to our work in Ref [19].

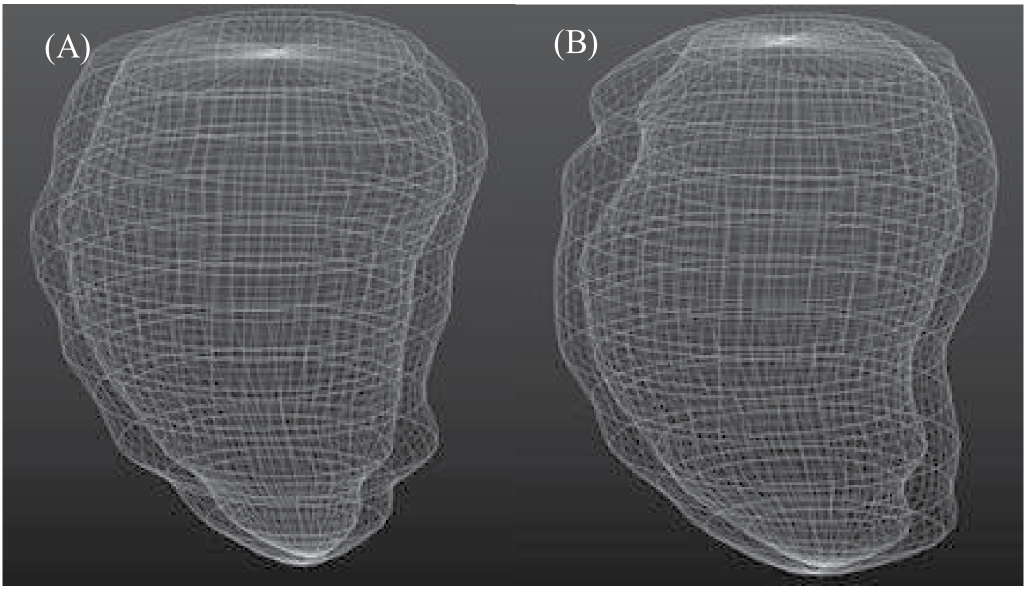
The prime effect of SVR may be viewed as: (i) effecting a decrease in myocardial oxygen consumption by reduction of LV peak normalized wall stress, resulting in improved functioning of LV, and (ii) augmentation of value of the contractility index $d\sigma^*/dt_{max}$ (2.69 ± 0.74 s^{-1} vs. 3.23 ± 0.73 s^{-1}, $p<0.01$). This improvement may be attributed to (i) increased maximal flow dV/dt_{max} with reduced LV mass, and (ii) improved regional contraction and contractility of the remote myocardium. The improvement in remote myocardial performance is likely due to reduced myocardial stress, along with effective and complete revascularization. This is because the SVR procedure reduces the volume by more dramatically reducing long-axis dimension compared with the short-axis dimension, and producing a more spherical ventricle.

Based on Table 10, increased LV contractile function $d\sigma^*/dt_{max}$ can be not only associated with increased maximal flow dV/dt_{max}, reduced LV mass, and also increased maximal change rate of shape factor dS/dt_{max} ($r=0.414$, $p<0.001$). There was also good correlation between $d\sigma^*/dt_{max}$ and LVEF ($r=0.69$, $p<0.001$, pre-SVR; $r=0.77$, $p<0.001$, post-SVR) (Figure 15).

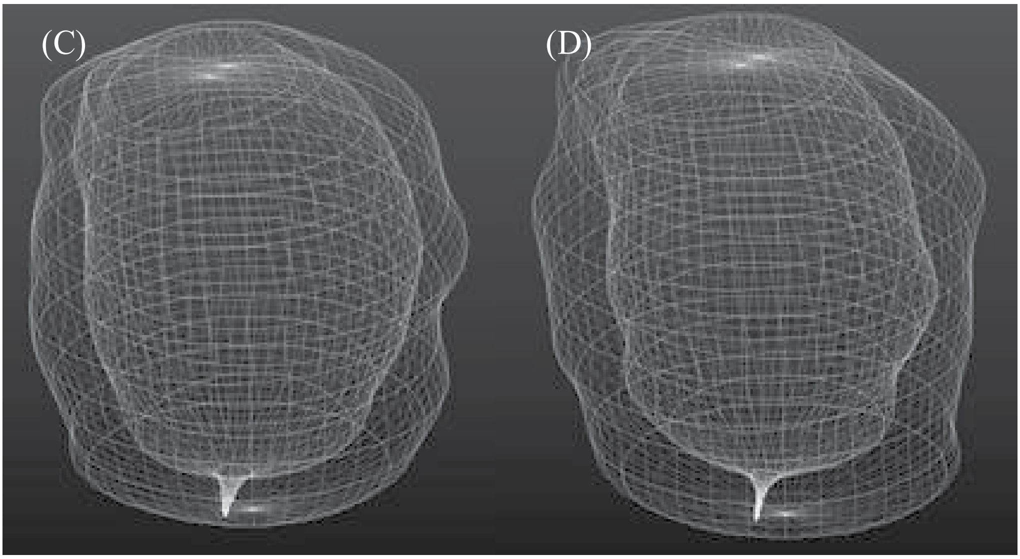
Variables	Pre SVR (n=40)	Post SVR (n=40)
Cardiac index (L/min/m^2)	2.84 ± 0.74	2.59 ± 0.74
Mean arterial pressure (mmHg)	85 ± 14	84 ± 8
Systolic blood pressure (mmHg)	115 ± 20	113 ± 10
Diastolic blood pressure (mmHg)	71 ± 12	70 ± 8
End diastolic volume index (ml/m^2)	156 ± 39	110 ± 33*
End systolic volume index (ml/ m^2)	117 ± 39	77 ± 31*
Stroke volume index (ml/ m^2)	39 ± 9	33 ± 8*
Left ventricular ejection fraction (%)	26 ± 7	31 ± 10*
LV mass index (g/m^2)	112 ± 25	101 ± 23*
End-diastolic long axis, BA_{ed}(cm)	10.89 ± 1.16	8.31 ± 1.00*
End-diastolic short axis, AP_{ed}(cm)	7.00 ± 0.80	6.64 ± 0.78*
End-systolic long axis, BA_{es}(cm)	10.37 ± 1.20	7.87 ± 1.05*
End-systolic short axis, AP_{es} (cm)	5.86 ± 0.98	5.23 ± 1.06*
End-diastolic sphericity Index, SI_{ed}	0.65 ± 0.087	0.81 ± 0.11*
End-systolic sphericity index, SI_{es}	0.57 ± 0.094	0.67 ± 0.13*
Difference between end-diastolic and end-systolic sphericity index, $SI_{ed} - SI_{es}$	0.077 ± 0.043	0.14 ± 0.059*
Long axis shortening (%)	4.8 ± 3.6	5.4 ± 4.4
Short axis shortening (%)	16.4 ± 6.8	22 ± 9.7*
dV/dt_{max} (ml/s)	364 ± 83	401 ± 81*
Pressure normalized wall stress	4.30 ± 0.95	3.31 ± 0.75*
Stroke work (mmHg·L)	6.61 ± 1.96	5.46 ± 1.64*
$d\sigma^*/dt_{max}$ (s^{-1})	2.69 ± 0.74	3.23 ± 0.73*

*$p < 0.05$.
Values are mean ± SD.

Table 10. Patients' data pre- and post-SVR. This table is related to our work in Ref [19].



(a)Pre-SVR ED & ES



(b) Post-SVR ED & ES

Fig. 13. 3-dimensional reconstructions during end-diastole (panels A and C) and end-systole (panels B and D) phases before (panels A and B) and after (panels C and D) SVR using LVtools. It is created from the endocardial and epicardial contours, which were drawn for calculations of ventricular volumes and function from the multiple short-axis cines (Figure 12). This figure is based on our work presented in Ref. 19.

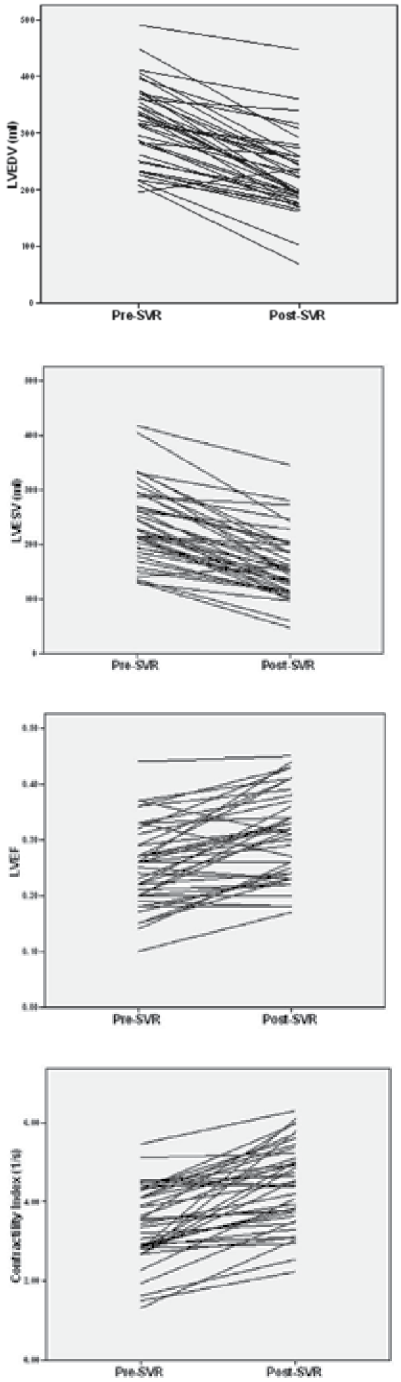


Fig. 14. Changes in end-diastolic volume (EDV), end-systolic volume (ESV), LVEF and contractility index $d\sigma^*/dtmax$ after SVR. This figure is adopted from our work in Ref. 19.

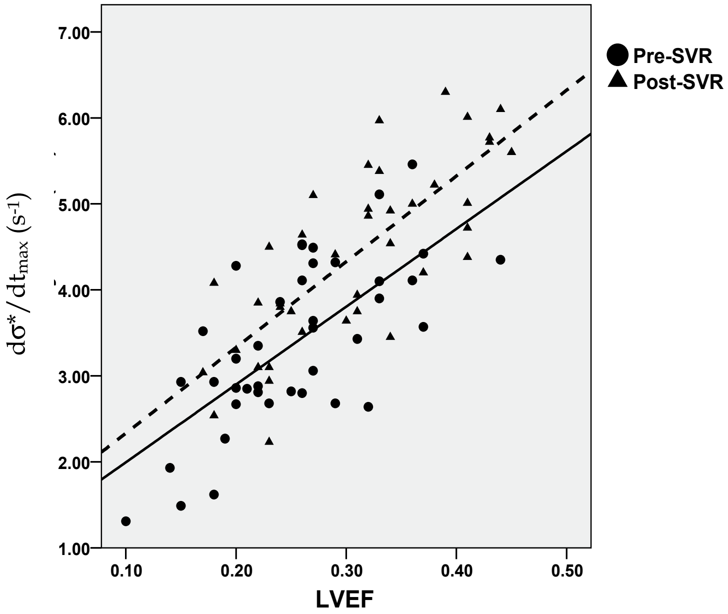


Fig. 15. Association between $d\sigma^*/dt_{max}$ and left ventricular (LV) ejection fraction (EF) pre- and post-SVR. (Solid line: $d\sigma^*/dt_{max}$=9.045×EF+1.091, r=0.69, p<0.001 for pre-SVR; dash line: $d\sigma^*/dt_{max}$=9.969×EF+1.337, r=0.77, p<0.001 for post-SVR). This figure is adopted from our work in Ref 19.

8. Conclusion

Myocardial infarct patients' hearts undergo substantial remodeling and loss of contractility, and can progress to heart failure. Echocardiographic texture analysis enables us to determine the percentage volume of the infarcted segments in the ventricular volume.

The ventricular remodelling is defined in terms of reduced alteration in curvedness index (%ΔC) from end-diastole to end-systole. The cardiac contractility is assessed in terms of the value of the index $d\sigma^*/dt_{max}$.

With the help of these three indices of percentage volume of MI segments, curvedness index and contractility index, we can track and assess LV progression to heart failure, and its recovery following CABG and SVR.

9. References

[1] Detection of myocardial scars in neonatal infants form computerised echocardiographic texture analysis, by M.V Kamath, R.C Way, D.N. Ghista, T.M. Srinivasan, C. Wu, S Smeenk, C. Manning, J. Cannon, *Engineering in Medicine*, Vol. 15, Vo.3, 1986

[2] A Geometrical Approach for Evaluating Left Ventricular Remodeling in Myocardial Infarct Patients,Yi Su, Liang Zhong, Chi-Wan Lim, Dhanjoo Ghista, Terrance Chua, Ru-San Tan, *Computer Methods and Programs in Biomedicine,* In Press, 2011

[3] Zhong L, Tan RS, Ghista DN, Ng EYK, Chua LP, Kassab GS (2007), Validation of a novel noninvasive cardiac index of left ventricular contractility in patients. *Am J Physiol Heart CircPhysiol,* 292(6):H2764-H2772.

[4] Zhong L, Ghista DN, Ng EYK, Lim ST.Passive and active ventricular elastances of the left ventricle. Biomed Eng Online4: 10, 2005.

[5] Shishido T, Hayashi K, Shigemi K, Sato T, Sugimachi M, Sunagawa K. Single-beat estimation of end-systolic elastance using bilinearly approximated time-varying elastance. *Circulation* 102: 1983–1989, 2000.

[6] Liang Zhong, Kian-Keong Poh, Li-Ching Lee, Thu-Thao Le, Ru-San Tan, Attenuation of Stress-based Ventricular Contractility in Patients with Heart Failure and Normal Ejection Fraction,Annals, Academy of Medicine, Singapore, April 2011, Vol 40, No 4

[7] Hogg K, Swedberg K, McMurray J. Heart failure with preserved left ventricular systolic function: epidemiology, clinical characteristics, and prognosis. *J Am Coll Cardiol.*2004;43:317-327.

[8] Little WC. The left ventricular (dP/dt)max-end-diastolic volume relation in closed-chest dogs. *Circ Res.* 1985;56:808-815.

[9] Vasan RS, Levy D. Defining diastolic heart failure: a call for standardized diagnostic criteria. *Circulation*. 2000;101:2118-2121.

[10] Devereux RB, Alonso DR, Lutas EM, Gottlieb GJ, Campo E, Sachs I, Reichek N. Echocardiographic assessment of left ventricular hypertrophy: comparison to necropsy findings. *Am J Cardiol*. 1986;57:450-458.

[11] Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E. Recommendations for chamber quantification. *J Am SocEchocardiogr*. 2005;18:1440-1463.

[12] Kitzman DW, Gardin JM, Gottdiener JS, et al. Importance of heart failure with preserved systolic function in patients > or =65 years of age. CHS Research Group Cardiovascular Health Study. *Am J Cardiol*. 2001;87:413-419.

[13] Banerjee P, Banerjee T, Khand A, et al. Diastolic heart failure: neglected or misdiagnosed? *J Am CollCardiol*. 2002;39:138-141.

[14] Burkhoff D, Maurer MS, Packer M. Heart failure with a normal ejection fraction: is it really a disorder of diastolic function? *Circulation*. 2003;107:656-658

[15] Zhang, J.M., Chua, L.P., Ghista, D.N., Yu, S.C.M. and Tan, Y.S., Numerical investigation and identification of susceptible site of atherosclerotic lesion formation in a complete coronary artery bypass model, *J. of Medical and Biological Engineering and Computing*, 2008, Vol. 46 (7), pp. 689-699.

[16] Sankaranarayanan, M., Chua L. P., Ghista D. N. and Tan Y. S., Computational model of blood flow in the aorta-coronary bypass graft, *BioMedical Engineering Online,*2005,4:14.

[17] Chua, L.P., Zhang, J.M., Yu, S.C.M., Ghista, D.N. and Tan, Y.S., Numerical study on the pulsatile flow characteristics of proximal anastomotic model, *J. of Engg. in Medicine, Proceedings of the Institution of Mechanical Engineers Part H,* 2005, Vol. 219, pp 361-379.

[18] Xiong, F.L., and Chong, C.K., A parametric numerical investigation on haemodynamics in distal coronary anastomoses, *Medical Engineering and Physics*, 2008, Vol. 30(3), pp. 311-320.

[19] Zhong L, Sola S, Tan RS, Ghista DN, Kurra V, Navia JL, Kassab GS. Effects of surgical ventricular restoration on left ventricular contractility assessed by a novel contractility index in patients with ischemic cardiomyopathy. *Am J Cardiol* 103: 674-679, 2009.

[20] Streeter Jr DD, Hanna WT. Engineering mechanics for successive states in canine left ventricular myocardium. I Cavity and wall geometry.*Circ Res*. 1973;33:639-655.

[21] Dhanjoo N. Ghista, *Applied Biomedical Engineering,* CRC Press, 2009

33

Renal Physiological Engineering – Optimization Aspects

David Chee-Eng Ng[1] and Dhanjoo N. Ghista[2]
[1]Department of Nuclear Medicine and PET, Singapore General Hospital,
[2]Department of Graduate and Continuing Education, Framingham State University,
Framingham, Massachusetts,
[1]Singapore
[2]USA

1. Introduction

Renal overall functional performance is characterized by excretory function of major end products of protein metabolism, regulation of ionic processes, maintenance of fluid balance and blood volume regulation. Minor functions include hormonal regulation of red cell production and stabilization of blood pressure. Although the kidneys comprise less than 0.5% of total body weight, they receive approximately 20% of the total cardiac output [1]. This consideration underscore the important role played by the kidneys.

The renal circulation has a unique sequence of vascular elements: a high-resistance afferent arteriole, a high-pressure glomerular filtration capillary structure, another high-resistance efferent arteriole and a series of tubular structures with unique absorption/excretion properties. The basic functional unit is the nephron. The nephron consists of a glomerular filtration structure and a tubular system, with its associated vascular elements.

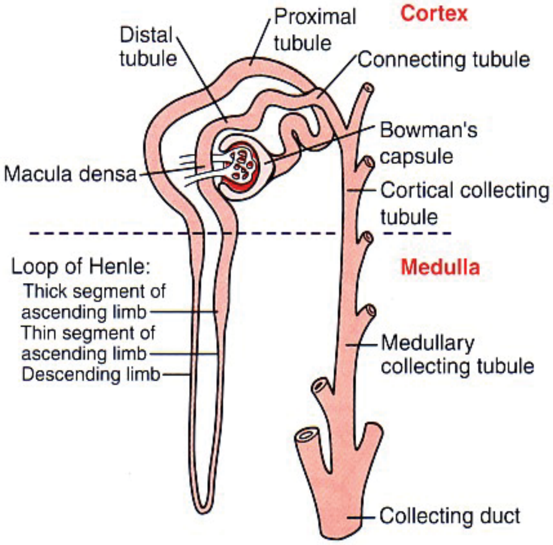

Proximal tubule
Cortex
Distal tubule
Connecting tubule
Bowman's capsule
Macula densa
Cortical collecting tubule
Loop of Henle:
Medulla
Thick segment of ascending limb
Thin segment of ascending limb
Descending limb
Medullary collecting tubule
Collecting duct

Fig. 1. Diagram of a nephron unit (adapted from [1]).

Structurally, there are 500,000 – 1,000,000 nephron units in the kidney [1,2]. Each nephron has a length of 40 mm and tubular diameter of 50 um. Microscopic puncture and perfusion techniques make it possible to measure single-nephron filtration rates, absorption and secretion rates. The single-nephron glomerular filtration rate (SNGFR) is approximately 30 nl/min.

Approximately 1200 ml/min of blood flow to the kidneys which represents 20% of total cardiac output. In the filtration mechanism, a total filtration surface of 1 m^2 is found. A global renal fitrate of 120 ml/min (180 L/day) is produced by both kidneys, of which 99.4% of water is reabsorbed to yield about 1 liter of urine/day. The concentration ability of the renal system is largely resides in the tubular system embedded within the renal medulla. Figure 1 demonstrates the disparate functionality of the renal nephron unit.

In this paper we will analyse the operating characteristics of the renal system and examine the optimal features of several aspects of renal physiological engineering mechanisms, particularly, the countercurrent multiplier mechanism for urine concentration and how optimal renal function in terms of renal clearance is maintained.

2. Kidney functional analysis

2.1 Countercurrent mechanisms and modelling of urine concentration

The concentration ability of the kidney is provided by a highly hyperosmotic renal medulla, which draws H_2O from the urinary filtrate within the collecting duct. The mechanism for providing and maintaining this highly hyperosmotic environment is found to be due to several mechanisms. The countercurrent multiplier process in the loop of Henle provides one of the most important mechanisms for this purpose. If the renal tubule is straight, it can provide a osmotic gradient within the renal medulla by about 300 mOsm/L, through an active Na^+ transport. However, for this gradient to continue, it requires rapid replacement against washout or dissipation of the osmotic environment in the renal medulla.

By looping the tubule in a parallel configuration and iterating the single effect of concentration, the kidney can generate and maintain the medullary concentration gradient up to 1200-1400 mOsm/L with lower energy costs. The two parallel tubes of the descending limb and the ascending limb of the loop of Henle are looped in close proximity in the hair-pin configuration (figure 1). The production of a chemical osmotic gradient is based on the active sodium reabsorption (requiring ATP) from the ascending loop; the maintenance of this gradient is crucial. The osmolality of interstitium in almost all parts of the body is about 300 mOsm/L. However, in the renal interstitium, the countercurrent mechanism provides increases up to 4 times from 300 mOsm/L to almost 1400 mOsm/L.

The hyperosmotic gradient provides the osmotic pressure to draw passive diffusion of water from the descending limb and the collecting duct. The countercurrent parallel design allows the descending limb to feedback to the ascending limb, forming a closed stable system of hyperosmolar environment and gradient, as shown in figure 2. It is the preservation of this hyperosmolar environment in steady state conditions that allows the urinary filtrate to be concentrated rapidly and efficiently.

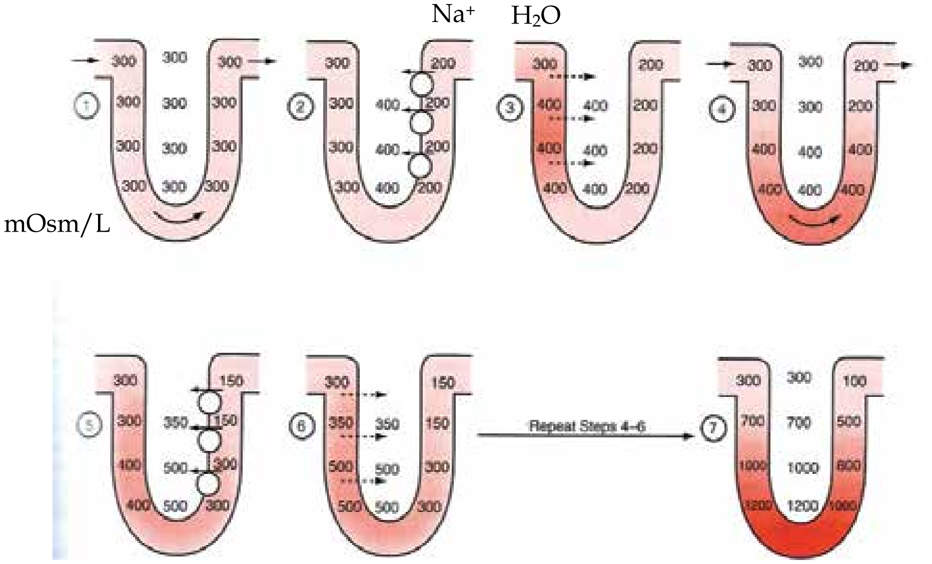
There is also a parallel system of renal vascular network of the vasa recta to prevent this hyperosmolar environment from dissipation. The contribution of the vasa recta into the concentrating mechanism is shown as follows: assume the interstitial tissue of cortex and glomerular filtrate are iso-osmotic to plasma. The tubular fluid entering the descending loop of Henle is also iso-osmotic to plasma. This fluid becomes progressively concentrated towards the bend. In the ascending loop it becomes less concentrated as it reaches the cortex. Under the influence of anti-diuretic hormone, blood flow through the vasa recta is decreased and osmotic equilibration of blood in the vasa with medullary interstitium is enhanced. In brief, the anatomical configuration of the vasa recta minimises but does not prevent solute loss from the medulla via the blood.

Other mechanisms in concentration include the role of urea re-circulated from the collecting duct, the role of vasopressin (antidiuretic hormone or ADH) acting on water transport cellular membrane water channel proteins aquasporin 1, 2, 3 and 4 regulating water permeability. Mutations of several aquasporin genes lead to loss of function and marked abnormalities of water balance, as documented in several reports involving AQP1 knock-out animals.

The contribution of urea to the concentration gradient in the renal medulla is also an important consideration. Diffusion of H_2O occurs from the tubular lumen into the interstitium. Active transport of Na^+ occurs from the tubular fluid. The withdrawal of H_2O from the collecting tubule leads to increased concentration of urea in the collecting tubule, causing a high gradient across the duct membrane, which favours diffusion of urea from the collecting duct into the medulla. From there, the urea diffuses into the descending loop of Henle and is re-circulated into the collecting duct. This contributes to the high urea content and osmolality of the medulla in the concentrating kidney.

2.1.1 Counter-current multiplier mechanism in the loop of Henle

The countercurrent mechanism in the loop of Henle is illustrated in figure 2:



Na^+
H_2O
mOsm/L
①
300 300 300
300 300 300
300 300 300
300 300 300
②
300 200
300 400 200
300 400 200
300 400 200
③
300 200
400 400 200
400 400 200
400 400 200
④
300 300 200
300 300 200
400 400 400
400 400 400
⑤
300 150
300 350 150
400 500 300
400 500 300
⑥
300 150
350 350 150
500 500 300
500 500 300
Repeat Steps 4-6
⑦
300 300 100
700 700 500
1000 1000 800
1200 1200 1000

Fig. 2. Countercurrent multiplier process in the loop of Henle, in creating osmotic gradient and urine concentration (adapted from [1]). Note Na^+ absorption from the ascending limb, and passive diffusion of H_2O from the descending limb.

For the loop of Henle shown in figure 2, as a start, it is assumed that the loop of Henle is filled with a fluid with a concentration of 300 mOsm/L. First, the sodium transport from the lumen of the ascending limb to the interstitium, which instantaneously equilibrates with the descending limb. The osmolality in the ascending limb decreases. Because of the higher osmolar interstitium, the fluid in the descending limb increases in osmolality as water is shifted out by passive osmosis. However, the hair-pin structure causes the flow of hyperosmolar fluid in the descending limb to enter the ascending limb. The steps are repeated over and over, with the effect that the process gradually traps sodium in the medulla and multiplies the concentration gradient until the osmolality of the fluid in the loop of Henle and the interstitium reaches 1200 to 1400 mOsm/L.

2.1.2 Concentration of urine in the medullary interstitium

As discussed, tubular fluid entering descending loop of Henle is iso-osmotic to plasma. This tubular fluid becomes progressively concentrated towards the bend. In the ascending loop, it becomes less concentrated as it rises to the cortex (decreasing from over 1000 to 100).

Within the medullary interstitium, other mechanisms co-operate to concentrate the urine. Under the influence of ADH (anti-diuretic hormone) blood flow through the vasa recta is decreased, and osmotic equilibration of blood in the vasa recta with medullary interstitium is enhanced. Solutes such as sodium, chloride and urea enter the descending blood vessels as they pass through the progressively higher osmolality of the interstitium and H_2O leaves the vessels. In the ascending limb, the opposite events take place and H_2O is reabsorbed into the blood vessels.

In brief, the anatomical configuration of the vasa recta minimises but does not prevent solute loss from the medulla via the blood supply. Because of diffusion of H_2O from the tubular lumen into the interstitium, there is equilibrium between fluid in the collecting tubule and that in the interstitium. The withdrawal of H_2O from the collecting tubule leads to increase in the concentration of urea in the collecting tubule causing a high gradient across the duct membrane, which favours diffusion of urea from the collecting duct into the interstitium. From there, urea diffuses into the descending limb of the loop of Henle and is recirculated into the ascending limb and back into the collecting duct, contributing to the high urea concentration in the medulla in the concentrating kidney.

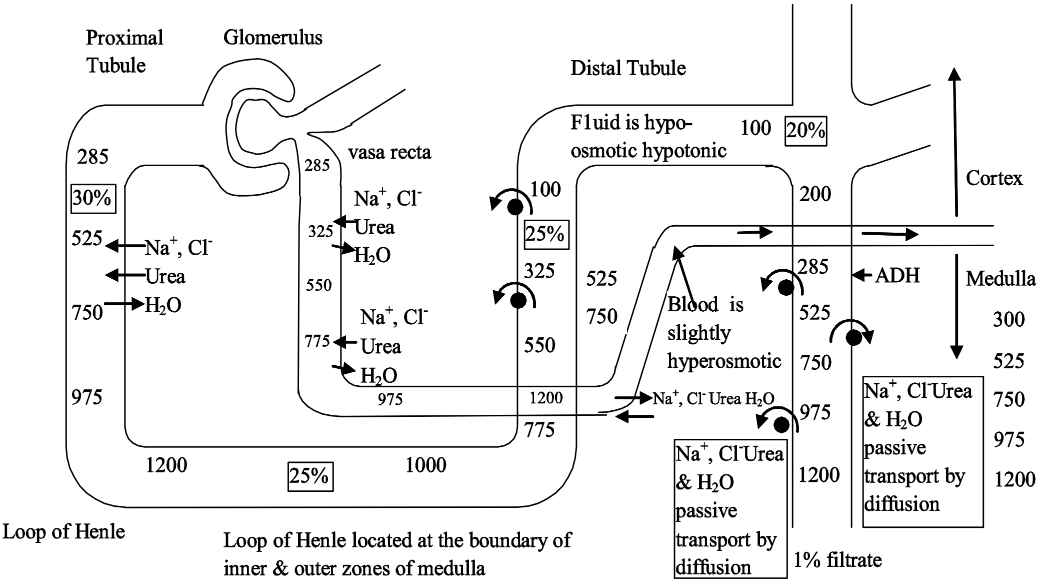

Proximal Tubule
Glomerulus
Distal Tubule
285
30%
525
Na⁺, Cl⁻
Urea
750
H₂O
975
1200
25%
1000
Loop of Henle
Loop of Henle located at the boundary of inner & outer zones of medulla
285
vasa recta
Na⁺, Cl⁻
325
Urea
H₂O
550
Na⁺, Cl⁻
775
Urea
H₂O
975
100
25%
325
550
1200
775
Fluid is hypo-osmotic hypotonic
100
20%
525
750
Blood is slightly hyperosmotic
Na⁺, Cl⁻ Urea H₂O
200
285
525
750
975
1200
Na⁺, Cl⁻Urea & H₂O passive transport by diffusion
1% filtrate
ADH
Na⁺, Cl⁻Urea & H₂O passive transport by diffusion
Cortex
Medulla
300
525
750
975
1200

Fig. 3. Schematic diagram showing both the countercurrent multiplier process in the loop of Henle and the vasa recta producing the osmotic gradient. Depicted in the figure are: (i) the passive and active exchanges of water and ions, (ii) concentrations of tubular urine and peritubular fluid in millimetres per litre, (iii) percentages of glomerular filtrate within the tubule at various levels.

In summary (please refer to figure 3),

1. In descending limb, Na^+ transports passively, Cl^- follows and H_2O transport by osmosis because the medullary region is hyperosmotic.
2. In ascending limb, Na^+ transports actively, Cl^- follows.
3. In distal tubule and collecting duct, presence of ADH makes water flow out osmotically (therefore the urine becomes very concentrated).

2.1.3 Linear coupled system of the loop of Henle and analytical solutions

In this model, the driving force for increasing the osmolality (largely contributed by the concentration of Na^+) of the descending limb is proportional to the difference in the osmotic gradient between it and renal interstitium. The renal interstitium itself has a osmotic concentration proportional to the ascending limb of the loop of Henle, due to active transport of Na^+ out of the ascending limb.

The osmolality (largely due to the concentration of Na^+) of the ascending limb is modelled on active sodium transport and hence the rate of fall is only related to its own concentration/osmolality.

The concentration of the interstitium is largely identical to the concentration in the descending tubule as there is passive movement of water through the descending tubule.

The schematic figure 4 illustrates the model of the renal tubule.

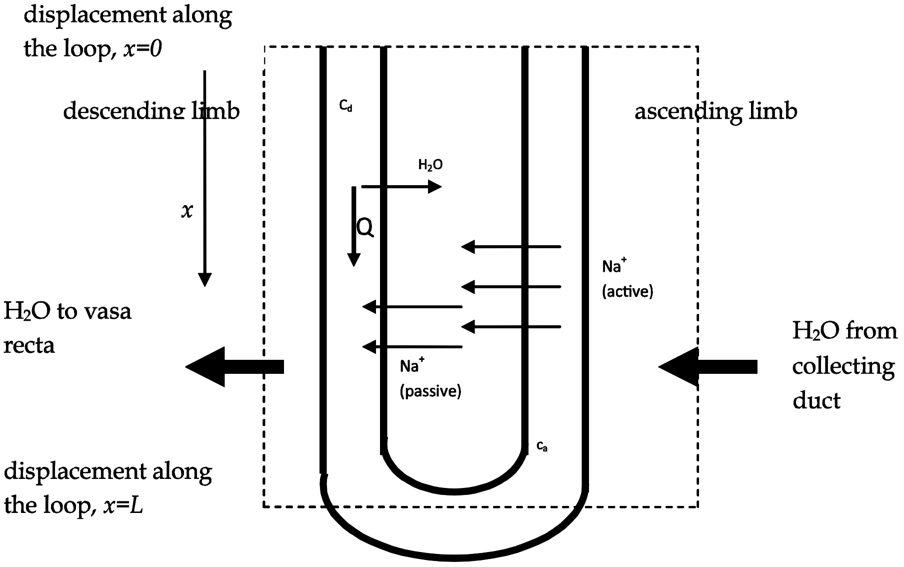


Fig. 4. Schematic diagram of the loop of Henle.

x the distance along the loop of Henle measured from the origin of the descending limb (in mm)

L the total actual length of the loop of Henle measured along one of the limb (in mm)

C_d the concentration of Na^+ in the descending limb (in mOsm/L)

C_a the concentration of Na^+ in the ascending limb (in mOsm/L)

C_i the concentration of Na^+ in the interstitium (assumed proportional to C_a i.e $= k_0 C_a$) (in mOsm/L)

k_d the transport coefficient of Na^+ ions into the descending limb (in ml/min.mm)

k_a the active transport coefficient of Na^+ out of the ascending limb due to Na^+ pump

Q is the tubular flow rate, assumed to be fairly constant in the first approximation (in mL/min)

The governing equations for the descending and ascending limbs of the loop of Henle are shown as a coupled system of linear first-order ODEs, with C_d and C_a the concentration/osmolality of the descending and ascending limbs of the loop of Henle respectively. In the descending limb, the change of concentration of Na⁺ is modelled as proportional to the concentration difference between the interstitium and the descending limb. In the ascending limb, the change of concentration of Na⁺ is modelled as directly proportional to the concentration in the ascending limb itself through active removal of Na⁺ by the Na⁺ pump. This leads to the following linear coupled system:

Na+ in the descending limb:
$$\frac{d(QC_d)}{dx} = k_d(C_i - C_d) = k_d k_0 C_a - k_d C_d$$
$$\frac{dC_d}{dx} = \frac{k_d}{Q}(C_i - C_d) = \frac{k_d k_0}{Q} C_a - \frac{k_d}{Q} C_d \tag{1}$$

Na⁺ in the ascending limb:
$$Q\frac{dC_a}{dx} = k_a C_a \tag{2}$$

with $k_d, k_a > 0$. The flow rate Q in the renal tubule is taken as constant in the first approximation. Expressed as matrix equation with upper triangular matrix,

$$\begin{pmatrix} C_d' \\ C_a' \end{pmatrix} = \begin{pmatrix} -\frac{k_d}{Q} & \frac{k_d k_0}{Q} \\ 0 & \frac{k_a}{Q} \end{pmatrix} \begin{pmatrix} C_d \\ C_a \end{pmatrix} \tag{3}$$

The eigenvalues are $-\frac{k_d}{Q}$ and $\frac{k_a}{Q}$ and the eigenvectors are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} k_d k_0 \\ k_d + k_a \end{bmatrix}$. The general solution of this system is given by:

$$\begin{pmatrix} C_d \\ C_a \end{pmatrix} = H_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{-\frac{k_d}{Q}x} + H_2 \begin{bmatrix} k_d k_0 \\ k_d + k_a \end{bmatrix} e^{\frac{k_a}{Q}x} \tag{4}$$

where H_1 and H_2 are constants of the solution.

Analytically in phase space, since the 2 eigenvalues are real and opposite in sign, the origin of the linear system is a saddle point, asymptotically unstable. Hence, it is unlikely that the system will remain in the state of zero concentration in the ascending and descending limbs. In fact, the solution (4) shows that the system will tend towards a state where an increasing concentration exists in the loop of Henle, because of the positive eigenvalue k_a / Q (representing active sodium transport in the ascending limb) for large x. This is consistent with the observation that it is the active sodium transport in the ascending limb that drives the production of the concentration gradient within the interstitium of the renal medulla and keeps the countercurrent mechanism operational, rather than the passive osmotic gradient as governed by $-k_d / Q$ which tends to dissipate the osmotic gradient.

At the loop end of the loop of Henle, the concentration/osmolality can reach extremely high levels, driven by active sodium transport. Indeed if the active Na⁺ transport $k_a = 0$, then the system decays to a baseline value through the exponential term associated with k_d. This shows that without active transport, the concentration gradient and the countercurrent multiplier mechanism will dissipate within the renal medulla.

If we take the boundary conditions provided by empirical data in figure 2:

$$C_d(0) = 300 \text{ mOsm/L}$$

$$C_a(0) = 100 \text{ mOsm/L}$$

and assuming trial values of $\frac{k_d}{Q} = \frac{k_a}{Q} = 1/\text{mm}$, the concentration within the ascending and descending limbs are obtained as:

$$C_d = 250e^{-x} + 50e^{x}$$
$$C_a = 100e^{x}$$

This is plotted in the following figures:

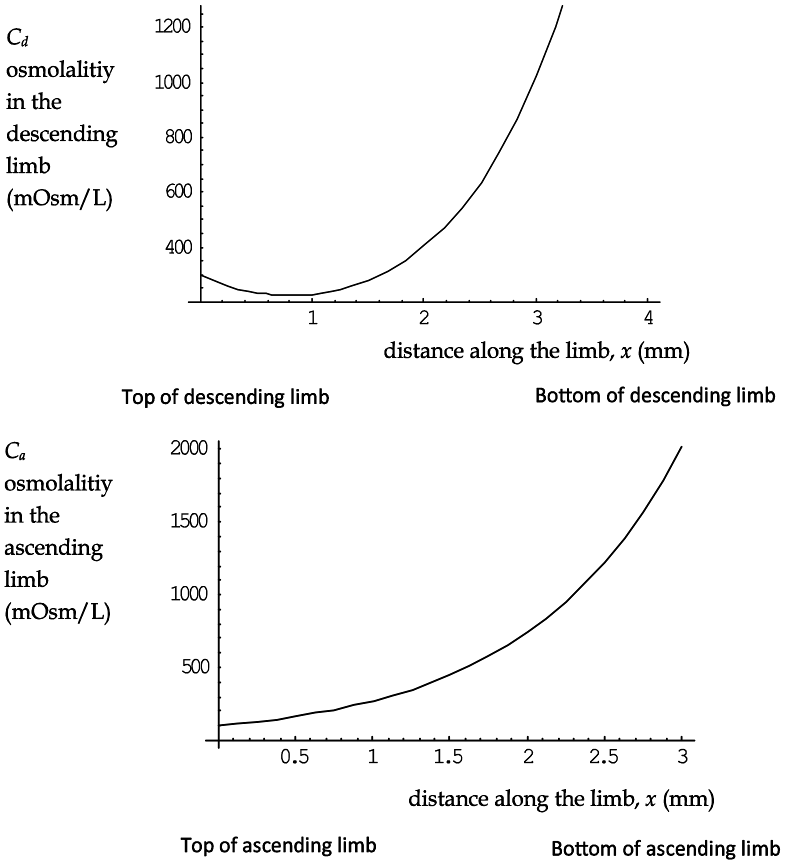


Fig. 5. Variation of Fluid Osmolality in the Descending and Ascending limbs of the Loop of Henle.

Although the values of kinetic transport coefficients are trial values, we can see how the model predicts the shape of the osmolality profile within the descending and ascending loops. The graphical representations in figure 5 demonstrate the highest level of osmolality achieved at the bottom end of the loop is about 1500 mOsm/L, reasonably consistent with empirical data. This model provides analytical solutions beyond that of Keener and Sneyd [3].

The limitations of this model is in the assumption of the values of transport coefficients of the descending and ascending tubules in the appropriate units, its linearity assumption and the disregard of its interaction with other modes or mechanisms of osmotic concentration. However, it can be seen that the analytical solutions provide a reasonable qualitative profile of the urinary concentration within the loop of Henle, under the assumptions made of its properties of its different segments and its "iterative" or "multiplier optimal design".

2.2 Single compartmental model of renal clearance kinetics – (1) single input

One of the important function of the kidney is excretion of metabolic waste products. How the kidney handles this excretory function has direct implications on clinical or physiological function. It is thus of interest to analyse the behaviour of the kidney as an excretory system.

Most accessible to analysis is the renal response to a single bolus of a metabolic substrate. Most renal clearance kinetics analyse the behaviour of the excretory function of the kidney in respect to endogenous or exogenous substrates. In some assessments, it involves the administration of a single bolus dose of an exogeneous substance into the blood circulation.

When a single bolus of such a substance is introduced into the human body system though an intravenous injection, the substance will initially spread out in the circulatory system and distribute into the extravascular body-fluid compartments of the body, while it is at the same time being removed by the kidney. Hence, if we represent the human body as a two-compartment system, then there will be 2 phases of decrease of the plasma concentration of this substance. The first phase represents the fall due to rapid distribution of the substance within the body from the blood circulation into the equilibrium body fluid compartments of the body, while it is at the same time being removed by the kidney. The second phase represents the fall due largely to the renal excretion of this substance.

However, in most cases, the first phase can be ignored and corrected for by empirical approximation so that only the second slower phase needs to be measured. Hence, a single late exponential function can be used to describe the fall in the plasma concentration of the substance. This principle is used in the physiological measurement of renal clearance or glomerular filtration rate (GFR) in human subjects.

2.2.1 Renal clearance analysis using a single-bolus model of renal tracer or substrate

Assume the amount of the tracer in the entire compartment is A (in mg or mmols). Let the concentration of the tracer in the compartment at time t be C_t (in mg/L or mmol/L) and the clearance be annotated as g (in L/min). By definition, $C = A/V$, where V is the total plasma volume or distribution volume, reasonably assumed constant in the body.

By the principle of mass conservation,

$$\frac{dA}{dt} = -g \cdot C_t \tag{5}$$

This is the governing first-order linear differential equation representing the kinetics of a one-compartment system.

Integrating over all time, the total tracer dose injected (*D*) is given by:

$$\int_0^\infty \frac{dA}{dt} dt = D = -\int_0^\infty g C_t dt = -g \int_0^\infty C_t dt \tag{6}$$

The absolute magnitude of the renal clearance *g* is:

$$|g| = \left| \frac{D}{\int_0^\infty C_t dt} \right| = \frac{\text{total dose of tracer injected}}{\text{area under the tracer concentration-time curve}} \tag{7}$$

We can show that the concentration of tracer in this compartment follows an exponential variation, by rewriting equation (5) as:

$$V \frac{dC_t}{dt} = -g \cdot C_t \tag{8}$$

Separating variables, we get:

$$\int_{C_0}^{C_t} \frac{dC_t}{C_t} = -\frac{g}{V} \int_{t_0}^{t} dt$$

We have a mono-exponential clearance scheme, as follows:

$$C_t = C_0 e^{-\frac{g}{V}(t - t_0)} \tag{9}$$

By taking logarithms of both sides, we get a linear relationship on the "semi-log" scale as:

$$\ln C_t = \ln C_0 - \frac{g}{V}(t - t_0) \tag{10}$$

Equation (10) is the basis of plotting the tracer concentration against time as a semi-log graph, so that (i) the absolute value of the gradient of the slope will be given by (renal clearance)/*V*, which is also called the clearance constant λ, and (ii) the y-intercept will be given by C_0 which is *D/V*.

So the initial volume of distribution, *V*, will be given by

$$V = \frac{D}{C_0} \tag{11}$$

Hence,

$$\text{Renal clearance} = V \times \lambda = \frac{D}{C_0} \times \lambda \tag{12}$$

Or, the estimated renal clearance is the

$$\text{Distribution volume} \times \text{Clearance constant} = \frac{\text{total dose injected}}{(\text{y-intercept of ln C vs t curve})} \times (\text{gradient of ln C vs t curve})$$

Historically, this methodology is often known as the indicator-dilution method or the Stewart-Hamilton method, although the origins of this method antedate the work of Stewart and Hamilton [4,5].

2.2.2 Physiological measurement of the Glomerular Filtration Rate (GFR)

If we can insert a microneedle with a flow gauge into each glomerulus in the kidney and measure the flow rate experimentally, this would constitute one way of measuring the GFR. This can be done in-vitro with micropuncture and microperfusion techniques. On a body system level, the global renal clearance has to be obtained by other ways.

Typically, GFR is deduced by measuring the renal filtered loss or clearance of a suitable substance from the plasma. In physiological tests, creatinine clearance and Cr-51 EDTA clearance are typically used. Dynamic renogram modelling of impulse tracer kinetics through the kidney had previously been performed [6].

In order to measure GFR accurately, the following properties must apply to the particular substance used to measure GFR :

1. the substance must be freely filtered through the glomerulus
2. it must not undergo renal tubular secretion or absorption
3. it must not bind to plasma proteins
4. it must not be lost through any other methods from the body
5. it must not be metabolised or changed chemically in the body.

These are severe restrictions and there are only some possible candidates for this substance, including :

1. endogenous creatinine, but this is less accurate in children
2. inulin
3. chromium-51 EDTA
4. technetium99m-DTPA

If the loss of these substances from the body can be measured, one can get a quantitative reflection of the excretory function of the kidney.

Of the four substances mentioned, only endogenous creatinine is found within the human body. The other substances have to be introduced into the human body. Inulin is a polysaccharide molecule with a molecular weight of 5200. Creatinine itself is a by-product of skeletal muscle metabolism, and it is present in the plasma at a relatively constant concentration and does not require intravenous infusion into the patient. However, creatinine is not a perfect marker for GFR because a small amount of it is secreted by the tubules and hence it is not a pure glomerular agent and tends to overestimate the GFR.

Incidentally, there is an agent, para-aminohippuric acid (PAH), which is not only filtered but also secreted to a large extent, so that it can be used to measure not the GFR but the effective renal plasma flow rate. Chromium-51 EDTA is a radiolabelled EDTA with the gamma emitter, chromium-51. This agent is very close to a purely glomerular filtered agent. Tc99m-DTPA is also a glomerular filtered substance, radiolabelled to the gamma emitter, Tc99m.

2.2.3 Continuous input of substrate model – relationship between steady-state serum creatinine concentration in the body and renal clearance

2.2.3.1 Theory and application

As opposed to the single-bolus renal kinetics, in the body, endogenous metabolic substrates are introduced into the blood circulation in a continuous way. Thus the model of renal clearance kinetics given above will have to be modified to take this continuous input into account. Analysing this continuous input model will be useful to evaluate the optimal and crucial renal handling of endogenous waste products in a typical human body.

The result of such a continuous input and renal excretion gives rise to a steady-state concentration of a renal-excreted substrate in the body. A typical endogenous substrate produced in a continuous fashion in the body and excreted by the renal route is creatinine. Figure 6 shows an inverse relationship between plasma creatinine concentration and GFR. The lower the renal clearance, the higher is the steady-state blood concentration of the substrate. However, this relationship is not linear but largely inverse rectangular hyperbolic (see figure 6).

To analyse this empirical relationship, further analysis can be performed using the single-compartment model but introducing a continuous input of substrate. This analysis follows from and extends the results obtained by Mazumdar [7].

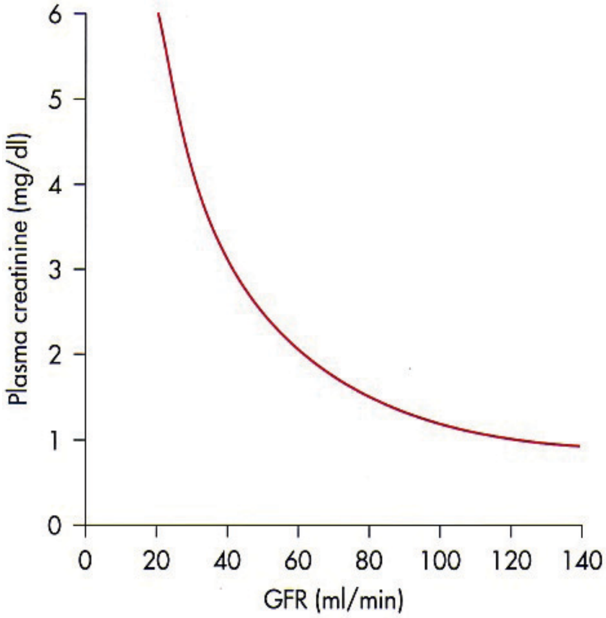


Fig. 6. Empirical relationship between blood creatinine levels and the renal clearance rate (adapted from [2]).

If instead of a single dose of tracer or substrate given as discussed in the previous section, constant doses of creatinine (substrate) are given at equal intervals of time ie, at intervals of period T, then using the single-bolus equation (7), the concentration of the substrate immediately after the second dose is given by:

$$C_1 = C_0 + \left(C_0 e^{-\frac{g}{V}T} \right) \tag{13}$$

Immediately after the third dose, the substrate concentration is given by:

$$\begin{aligned} C_2 &= C_0 + \left[C_0 + \left(C_0 e^{-\frac{g}{V}T} \right) \right] e^{-\frac{g}{V}T} \\ &= C_0 + C_0 e^{-\frac{g}{V}T} + C_0 e^{-\frac{2g}{V}T} \end{aligned} \tag{14}$$

Immediately after the *n*th dose, the substrate concentration is given by:

$$\begin{aligned} C_{n-1} &= C_0 + C_0 e^{-\frac{g}{V}T} + \ldots + C_0 e^{-(n-1)\frac{g}{V}T} \\ &= C_0 \left[1 + e^{-\frac{g}{V}T} + \ldots + e^{-(n-1)\frac{g}{V}T} \right] \\ &= C_0 \frac{1 - e^{-n\frac{g}{V}T}}{1 - e^{-\frac{g}{V}T}} \end{aligned} \tag{15}$$

As n tends to infinity, the creatinine concentration approaches an equilibrium value, given by:

$$C_\infty = \frac{C_0}{1 - e^{-\frac{g}{V}T}} \tag{16}$$

Linearization by Taylor's series approximation gives the equilibrium concentration of the creatinine in the blood to first order, as:

$$C_\infty = \frac{C_0}{\frac{g}{V}T - \frac{1}{2}\left(\frac{g}{V}T\right)^2 + \ldots} \approx \frac{(C_0 / T)V}{g} \tag{17}$$

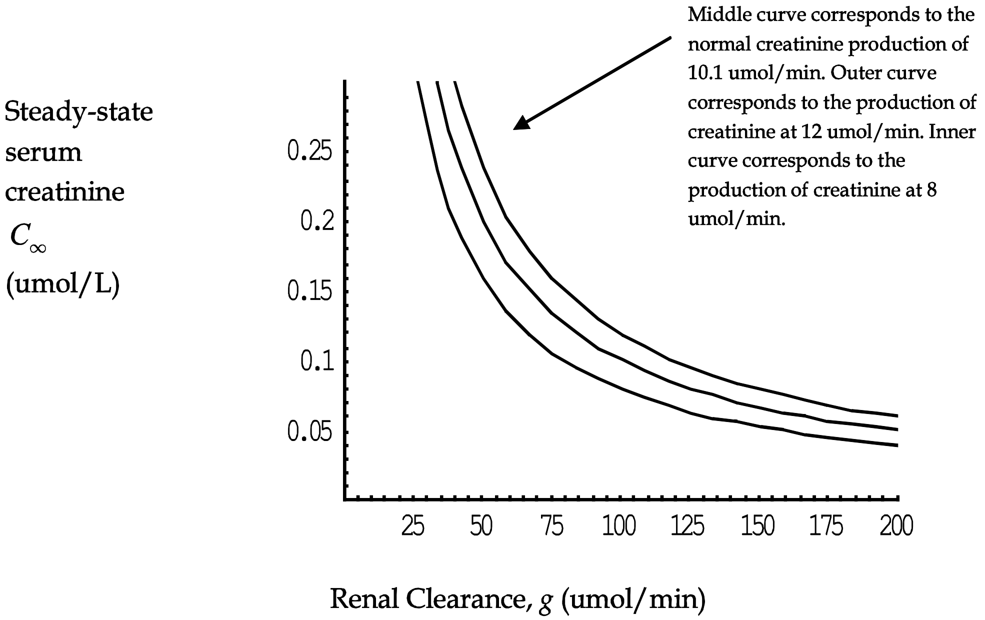
where $(C_0 / T)V$ is the amount of the creatinine introduced per unit time. Hence, using the parameter values in Table 1,

$$C_\infty \approx \frac{\text{amount of renal substrate introduced per unit time}}{\text{renal clearance}} = \frac{10.1}{\text{renal clearance}} \tag{18}$$

C_∞ is the steady-state concentration of creatinine, as it is produced and excreted continuously in the body. The relationship derived is the equilibrium or steady-state concentration of the substrate that is produced continuously in the body and excreted renally. The relationship is plotted in figure 6.

Rate of body production of creatinine metabolite, *A*	20-25 mg/kg body weight per day (approximately 1.5 g/day in a 70 kg man). In SI units, this would be 10.1 umol/min.
Human body renal clearance, *g*	120 ml/min (approximate).
Total volume of distribution of creatinine in a typical human of weight 70 kg	50,000 ml (approximate)

Table 1. Important physiological renal parameters (data from [2]).



Middle curve corresponds to the normal creatinine production of 10.1 umol/min. Outer curve corresponds to the production of creatinine at 12 umol/min. Inner curve corresponds to the production of creatinine at 8 umol/min.

Fig. 7. Model prediction of the relationship between blood creatinine levels and the renal clearance rate.

Graphically, equation (17) predicts a very close approximation to the empirical curve, which is an inverse rectangular hyperbolic relationship between serum creatinine levels and the renal clearance as shown in figure 7.

Depending on the rate of production of the metabolite creatinine in the human body, equation (17) also demonstrates that there is a series of iso-dose curves of renal clearance vs blood levels of renal substrate, similar to isothermal curves in ideal gas thermodynamics.

The application to human physiology can be seen as follows. It is well known that the serum creatinine levels in women is lower than in man. A typical muscular man producing a larger quantity of creatinine substrate due to muscle breakdown will, for the same renal clearance, demonstrate a higher blood concentration of creatinine for the same degree of renal clearance.

2.3 Renal clearance – convolution analysis

In general, the total amount of substrate in the body at time *t* is given by the convolution of the amount produced by the body per unit time, *A(t)*, which is a function of time and the biological clearance of that substance. In the case of pure renally-excreted substrate, such as creatinine, assuming a single-compartment clearance-kinetics as previously discussed, we have as follows:

Total amount of substrate in the body at time *t*: $A(t) * e^{-\frac{g}{V}t}$ (19)

If as above, the amount of substrate introduced into the blood compartment per unit time is constant, *A*, then the total amount of substrate at time *t* (accounting for renal clearance) is given by:

Total amount of substrate in the body at time *t*: $A * e^{-\frac{g}{V}t} = \int_{0}^{t} A du \cdot e^{-\frac{g}{V}(t-u)}$ (20)

The result takes a useful form for physical interpretation. Total amount of substrate in the body at time t is given by:

$$A * e^{-\frac{g}{V}t} = \int_{0}^{t} A du \cdot e^{-\frac{g}{V}(t-u)} = \frac{AV}{g}\left(1 - e^{-\frac{g}{V}t}\right) \tag{21}$$

Schematically, this relation is shown in figure 8, for blood creatinine levels:

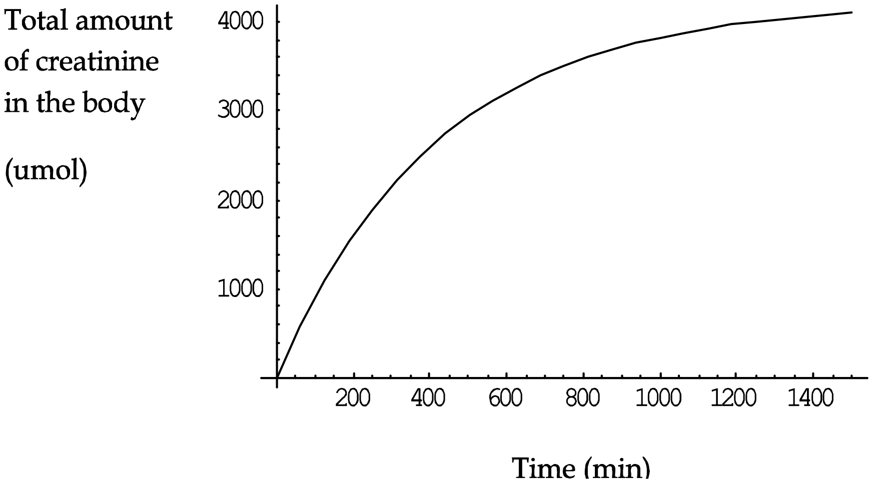


Fig. 8. Asymptotic steady-state concentration of blood creatinine levels, based on convolution analysis.

At $t \to \infty$, the equilibrium concentration of substrate in the blood compartment is:

$$C_\infty = \frac{1}{V}\left(\frac{AV}{g}\right) = \frac{\text{amount of substrate introduced per unit time}}{\text{renal clearance}} \tag{22}$$

consistent with the previous equation (18).

Application of the formula can be done using the physiological values in Table 1.

The body produces creatinine at the rate of 20-25 mg/kg body weight per day, which is approximately 1.5 g/day in a 70 kg man. In SI units, this is 10.1 umol/min. The renal clearance is approximately 120 ml/min. Hence, based on equation (22), the estimated steady-state serum creatinine level in the body based on the model would be predicted to be in the order of

10.1/120 = 0.084 umol/ml or 84 umol/L, as expected empirically.

Direct correlation with physiological parameters shows this convolution analysis to give reasonably close results.

3. Conclusion

The analytical model of the loop of Henle and the renal handling of metabolic substrates is aimed at showing to some extent how the renal system is optimized for filtration and regulation of urine concentration by the countercurrent mechanism in the loop of Henle and its medullary environment, which is largely physiologically engineered to increase and maintain at steady-state the high osmolality of the urine fluid to as high as 4 times normal blood osmolality.

The renal clearance of substrates is modelled as a single-compartment kinetic model, with single-input and more physiologically as continuous-input. The analytical solutions obtained from the continuous input of creatinine predict the body creatinine level to tend to asymptotically steady-state substrate blood concentration with time, in a relationship that is an inverse rectangular hyperbolic function to the renal clearance. This is close to the relationship found empirically. The relationship is found to be related to the amount substrate input per unit time divided by renal clearance. The same conclusion is obtained from convolution analysis of renal clearance. The formula predicts reasonable estimates for the actual serum creatinine levels in the body based on renal clearance and substrate input parameters.

4. References

[1] Guyton A C, Hall J E. Textbook of Medical Physiology. Tenth edition. W B Saunders 2000.
[2] Boron W F, Boulpaep E L. Medical Physiology. Updated Edition. Elsevier Saunders 2005.
[3] James Keener, Sneyd J. Mathematical Physiology. Springer-Verlag 1998.
[4] William Simon. Mathematical Techniques for Biology and Medicine. Dover Publications 1986.
[5] Rubinow, S I. Introduction to Mathematical Biology. Dover Publications 1975.

[6] Loh K M, Ng David, Ghista D N, Ridolph H. Quantitation of renal function based on two-compartmental modeling of renal pelvis. IEEE Conference 2005.

[7] Mazumdar J. An Introduction to Mathematical Physiology & Biology. Cambridge University Press. Second edition 1999.

34

Lung Ventilation Modeling for Assessment of Lung Status: Detection of Lung Disease and Indication for Extubation of Mechanically-Ventilated COPD Patients

Dhanjoo N. Ghista[1], Kah Meng Koh[2], Rohit Pasam[3] and Yi Su[4]
[1]*Department of Graduate and Continuing Education, Framingham State University,*
Framingham, Massachusetts,
[2]*VicWell BioMedical,*
[3]*Quodient, Inc,*
[4]*Institute of High Performance Computing, Agency for Science, Technology and Research,*
[1,3]*USA*
[2,4]*Singapore*

1. Introduction

In pulmonary medicine, it is important to detect lung diseases, such as chronic obstructive pulmonary disease (COPD), emphysema, lung fibrosis and asthma. These diseases are characterized in terms of lung compliance and resistance-to-airflow. Another important endeavour of pulmonary medicine is mechanical ventilation of COPD patients and determining when to wean off these patients from the mechanical ventilator. In both these medical domains, lung ventilation dynamics plays a key role.

So in this chapter, we develop the lung ventilation dynamics model in terms of monitored lung volume (V) and driving pressure (P_L), in the form of a differential equation with parameters of lung compliance (C) and resistance-to-airflow (R). We obtain the solution of this equation in the forms of lung volume (V) function of P_L, C and R. For the monitored lung volume V and pressure P_L data, we can evaluate C and R by matching the model solution expression with the monitored lung volume V and driving pressure P_N data. So what we have done here is to develop the method for determining lung compliance (C) and resistance-to-airflow (R) as average values of C (= C_a) and R (= R_a) during the ventilation cycle.

Now in some lung diseases such as in emphysema, the lung compliance (C) is high. In other lung diseases such as in asthma, the airway resistance (R) is high. So we need to determine the ranges of C and R for normal lung status as well as for lung disease states. Then, we can develop a 2-parameter R-C diagnostic coordinate plane, on which we can demarcate the zones for different diseases. Then, for any patient's (R, C) value, we can plot the (R, C) point in the R-C diagnostic coordinate plane, and based on its location designate the lung disease state of the patient. A more convenient way for detecting lung disease is to combine R and C along with some ventilator data (such as tidal volume and breathing rate) into a non-dimensional lung ventilator index (LVI). Then, we can determine the ranges of LVI for normal and disease states, and thereby employ the patient's computed values of LVI to designate a specific lung disease for the patient. The *LVI* concept for detecting lung disease is more convenient to adopt in clinical practice, because it enables detection of lung disease states in the form of just one lung-ventilation number.

Now, in this methodology, we need to monitor (i) lung volume, by means of a spirometer, and (ii) lung pressure (P_L) equal to P_{mo} (pressure at mouth) minus pleural pressure (P_p). The pleural pressure measurement involves placing a balloon catheter transducer through the nose into the esophagus, whereby the esophageal tube pressure is assumed to be equal to the pressure in the pleural space surrounding it. Now this procedure cannot be carried out non-traumatically and routinely in patients. Hence, for routineand noninvasive assessment of lung ventilation for detection of lung disease states, it is necessary to have a method for determining *R* and *C* from only lung volume data. So, then, we have shown how we can compute *R*, *C* and lung pressure values non-invasively from just lung volume measurement.

Finally, we have presented how the lung ventilation modeling can be applied to study the lung ventilation dynamics of COPD patients on mechanical ventilation. We have shown how a COPD patient's lung *C* and *R* can be evaluated in terms of the monitored lung volume and applied ventilatory pressure. We have also formulated another lung ventilator index to study and assess the lung status improvement of COPD patients on mechanical ventilation, and to decide when they can be weaned off mechanical ventilation.

2. Lung ventilation model

2.1 Scope

In this section, we have developed a lung ventilation model by modeling the lung volume response to mouth minus pleural driving pressure (by means of a first order differential equation) in terms of resistance-to-airflow (*R*) and the lung compliance (*C*). The lung volume solution of the differential equation is matched with the clinical volume data, to evaluate the parameters, *R* and *C*. These parameters' values can help us to distinguish lung disease states, such as obstructive lung and lung with stiffened parenchyma, asthma and emphysema.

2.2 Role of lung ventilation

Lung ventilation constitutes inhalation of appropriate air volume under driving pressure (= mouth pressure – pleuralpressure), so as to: (i) provide adequate alveolar O_2 amount at

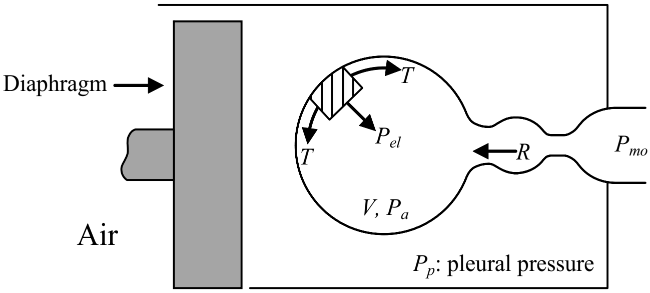


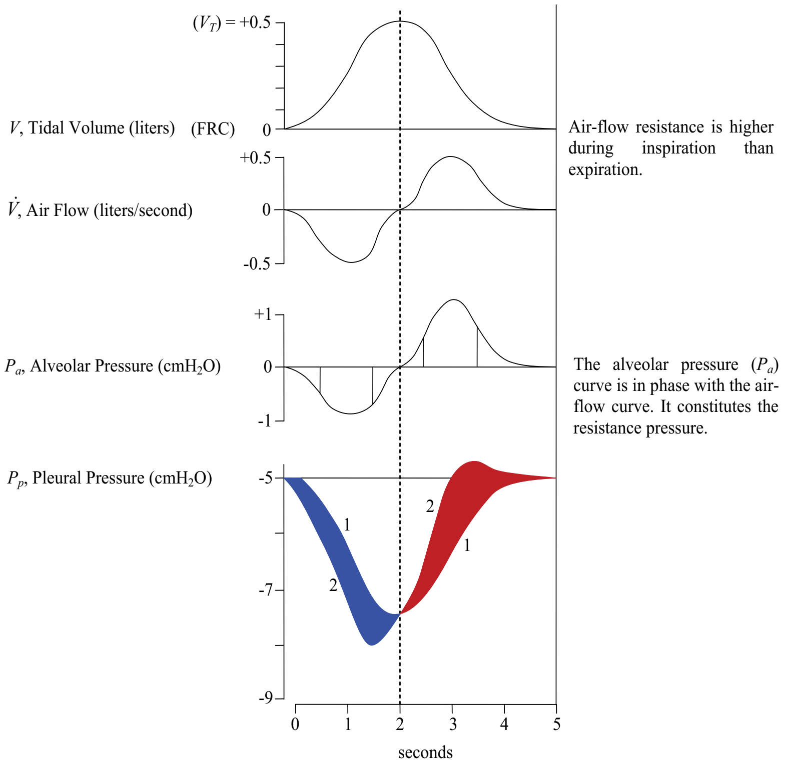
Fig. 1. Lumped Lobule Lung Model. In the figure, P_a is the alveolar pressure; P_{mo} is the pressure at the mouth; P_p is the pleural pressure; $P_{el} = P_a - P_p = 2h\sigma/r = 2T/r$, the lung elastic recoil pressure; *r* is the radius of the alveolar chamber and *h* is its wall thickness; *T* is the wall tension in the alveolar chamber; *V* is the lung volume; *R* is the resistance to airflow; and *C* is the lung compliance. This figure is adopted from our work in Ref [1].

appropriate partial pressure, (ii) oxygenate the pulmonary blood, and (iii) thereby provide adequate metabolic oxygen to the cells. Hence, ventilatory function and performance assessment entails determining how much air volume is provided to the alveoli, to make available adequate alveolar oxygen for blood oxygenation and cellular respiration.

In this lumped lobule lung model [1], we have (i) a lumped alveolar chamber of volume V and pressure P_a, and (ii) lumped airway having airflow resistance R. In this airway, the pressure varies from P_{mo} at the mouth to P_a in the alveolar chamber. The pleural pressure is P_p.

2.3 Lung ventilation analysis (using a linear first-order differential equation model)

We first analyze Lung Ventilation function by means of a model represented by a first-order differential equation (De_q) in lung-volume (V) dynamics in response to the driving pressure P_L (= mouth pressure – pleural pressure). In this model [2], the lung lobes and the alveoli are lumped into one lung lobule, as depicted in Figure 1. Figure 2 displays typical data of lung volume and flow, alveolar and pleural pressure.



Air-flow resistance is higher during inspiration than expiration.

The alveolar pressure (P_a) curve is in phase with the air-flow curve. It constitutes the resistance pressure.

Fig. 2. Lung ventilatory model and lung-volume and pleural-pressure data. In the bottom figure, Curve 1 represents the negative of P_{el}, the pressure required to overcome lung elastance (=1/C) plus elastic recoil pressure at the end of expiration. Curve 2 represents $P_p = -P_{el} + P_a$. Now, as can be noted from Figure 1, $P_a - P_p = P_{el}$. The lung driving pressure $P_L = P_{mo} - P_p$, and the net driving pressure $P_N(t)$ in Equation (1-b) equals P_L minus P_{el} at end-expiration. We define resistance-to-airflow (R) as $(P_{mo} - P_a)/\dot{V}$. We define lung compliance $C = V/(P_{el} - P_{el0}) = V/(P_a - P_p) - P_{el0}$. This figure is adopted from our work in Ref [1].

Based on Figures 1 and 2, we can put:

$$\left.\begin{array}{ll}\text{(i)} & (P_a - P_p) - P_{el} = 0 \\ \text{(ii)} & P_{el} = (2\sigma h) \;/\; R \;=\; 2\,T \;/\; R \;=\; V/C \;+\; P_{el0}\left(\text{at end} - \text{expiration}\right) \\ \text{(iii)} & (P_{mo} - P_a) = R\left(dV/dt\right) \\ \text{(iv)} & P_L = P_{mo} - P_p = \left(P_{mo} - P_a\right) + \left(P_a - P_p\right) \\ \text{(v)} & R\left(dV/dt\right) + V/C \;=\; P_L - P_{el0}(\text{lung elastic recoil pressure at end} - \text{expiration})\end{array}\right\} \text{(1-a)}$$

The lung ventilation model governing equation is hence formulated as:

$$R\dot{V} + \frac{V}{C} = P_L(t) - P_{el0} = P_N(t) \tag{1-b}$$

wherein:

i. the values of pressure$P_N(t)$ are obtained from the P_L (= P_{mo} – P_p) data relative to P_{el0}

ii. the parameters of this Governing De_q are lung compliance (C) and airflow-resistance (R); in the equation both R and C are instantaneous values

iii. $V = V(t) - V_0$ (the lung air volume at the end-expiration = lung air volume inspired and expired during a single breath)

iv. P_{el0} is the lung elastic-recoil pressure at the end of expiration, and

$$P_{el0} = P_{el} - \frac{V}{C} \tag{1-c}$$

v. At end-expiration when $\omega t = \omega T$, $P_L = P_{el0}$

Now, in order to evaluate the lung model parameters C and R, we need to simulate this governing equation to lung volume (V) and pressure (P_N) data. This clinical data is shown in Figure 3. The lung volume is measured by integrating the airflow velocity-time curve, where the airflow velocity can be measured by means of a ventilator pneumatograph; the lung volume can also be measured by means of a spirometer. Inhalation and exhalation pressures are measured by means of a pressure transducer connected to the ventilatory tubing; likewise, a pressure transducer can also be similarly connected to the spirometer tubing. The pleural pressure is measured by placing a balloon catheter transducer through the nose into the esophagus; it is assumed that the esophageal tube pressure equals the pressure in the pleural space surrounding it.

In Equation (1-a), we have put $P_N(t) = \sum_{i=1}^{3} P_i \sin(\omega_i t + c_i)$, expressed as a Fourier series. The governing equation (1-b) now becomes:

$$R\dot{V} + \frac{V}{C} = P_N(t) = \sum_{i=1}^{3} P_i \sin(\omega_i t + c_i) \tag{2-a}$$

where the right-hand side represents the net driving pressure minus pleural pressure: $P_N = (P_{mo} - P_p) - P_{el0}$. This P_N is in fact the driving pressure ($P_{mo} - P_p$) normalized with respect to its value at end-expiration. Equation (2-a) can be rewritten as follows:

$$\dot{V}+\frac{V}{RC}=\frac{1}{R}\sum_{i=1}^{3}P_i\sin(\omega_i t+c_i) \tag{2-b}$$

wherein the $P(t)$ clinical data (displayed in Figure 3) is represented by:

$$P(t)=\sum_{i=1}^{3}P_i\sin(\omega_i t+c_i) \tag{3}$$

P_1= 1.581 cmH$_2$O	P_2 = -5.534 cmH$_2$O	P_3 = 0.5523 cmH$_2$O
ω_1 = 1.214 rad/s	ω_2 = 0.001414 rad/s	ω_3 = 2.401 rad/s
c_1 = -0.3132 rad	c_2 = 3.297 rad	c_3 = -2.381 rad

The pressure curve (in Figure 3) represented by the above Equation (3) closely matches the pressure data of Figure 3. If, in Equation (1), we designate R_a and C_a as the average values (R and C) for the ventilatory cycle, then the solution of Equation (2) is given by:

$$V(t)=\sum_{i=1}^{3}\frac{P_iC_a[\sin(\omega_i t+c_i)-\omega_i R_a C_a\cos(\omega_i t+c_i)]}{1+\omega_i^2(R_aC_a)^2}-He^{-\frac{t}{R_aC_a}} \tag{4}$$

wherein the term (R_aC_a) is denoted by τ_a. We need to have $V = 0$ at $t = 0$. Hence, putting V (at $t = 0$) = 0, gives us:

$$H=\sum_{i=1}^{3}\frac{P_iC_a[\sin(c_i)-\omega_i R_a C_a\cos(c_i)]}{1+\omega_i^2(R_aC_a)^2} \tag{5}$$

Then from Equations (4) and (5), the overall expressions for $V(t)$ becomes

$$V(t)=\sum_{i=1}^{3}\frac{P_iC_a[\sin(\omega_i t+c_i)-\omega_i\tau_a\cos(\omega_i t+c_i)]}{1+\omega_i^2\tau_a^2}-\sum_{i=1}^{3}\frac{P_iC_a[\sin(c_i)-\omega_i\tau_a\cos(c_i)]}{1+\omega_i^2\tau_a^2}e^{-\frac{t}{\tau_a}} \tag{6}$$

We also want that dV/dt = 0 at t =0, implying no air-flow at the start of inspiration. So then by differentiating Equation (6), we get the expression for air-flow ($\dot{V}$), as:

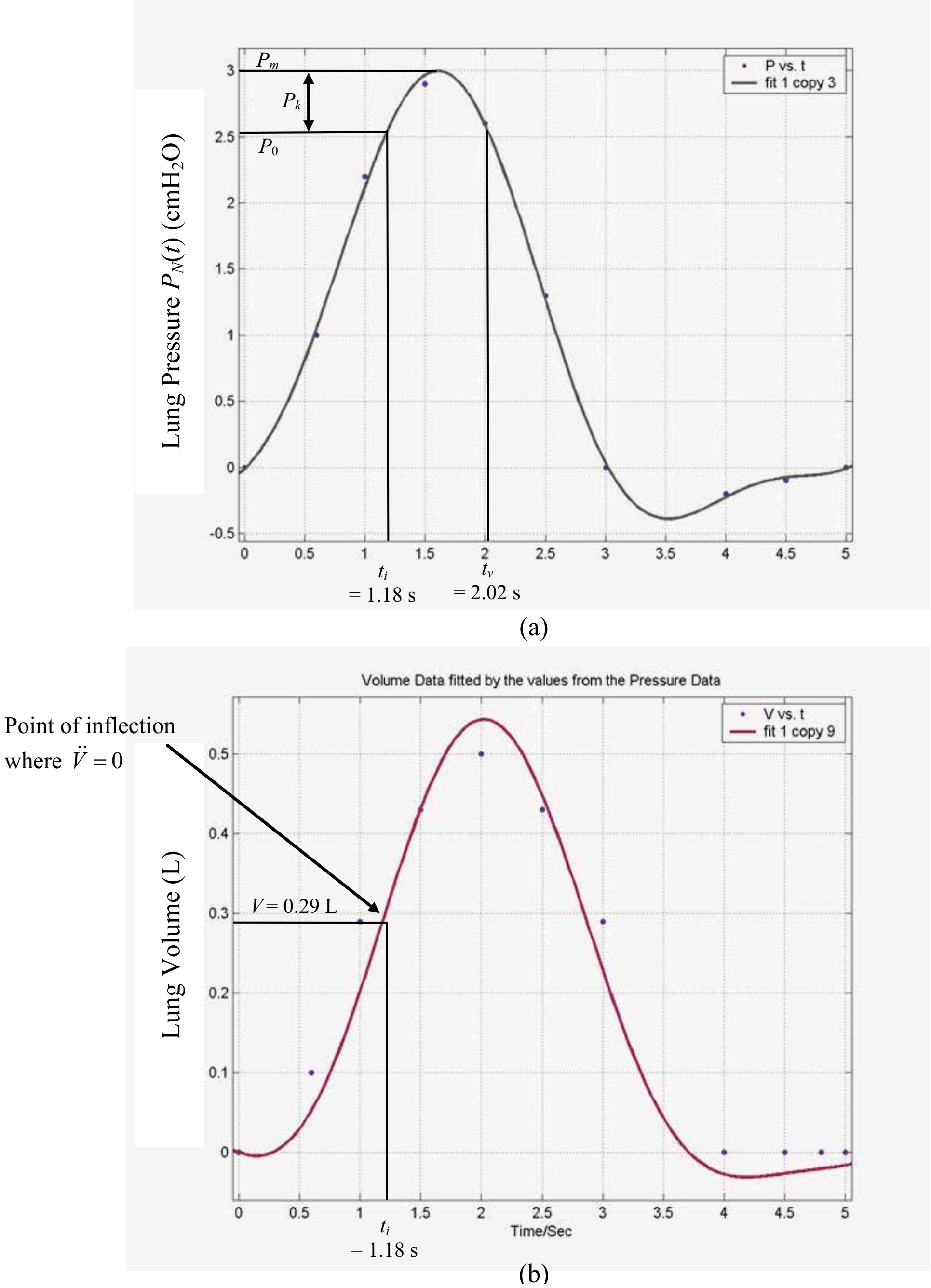
$$\dot{V}=\sum_{i=1}^{3}\frac{P_iC_a[\omega_i\cos(\omega_i t+c_i)+\omega_i^2\tau_a\sin(\omega_i t+c_i)]}{1+\omega_i^2\tau_a^2}+\sum_{i=1}^{3}\frac{P_iC_a[\sin(c_i)-\omega_i\tau_a\cos(c_i)]}{(1+\omega_i^2\tau_a^2)\tau_a}e^{-\frac{t}{\tau_a}} \tag{7}$$

For the above values of τ_a = 0.485 s and for ω_i and c_i given by Equation (3), we get $\dot{V}(t=0)=\sum_{i=1}^{3}(P_i/R_a)\sin(c_i)\approx 0$, to satisfy the initial condition.

Now by matching the above $V(t)$ expression in Equation (6) with the $V(t)$ data in Figure 3, and carrying out parameter-identification, we can determine the *in vivo* values of C_a, R_a and τ_a to be:

$$C_a = 0.218\ \mathrm{L}(\mathrm{cmH_2O})^{-1},\ R_a=2.275(\mathrm{cmH_2O})\mathrm{sL}^{-1},\ \tau_a = 0.485\ \mathrm{s} \tag{8}$$

The computed $V(t)$ curve, represented by Equation (6) for the above values of C_a and R_a, is shown in Figure 3.



(b)

Fig. 3. (a)The pressure curve represented by Equation (3) matched against the pressure data (represented by dots). (b) The volume curve represented by Equation (6), for C_a = 0.2132 L(cmH$_2$O)$^{-1}$ and R_a = 2.275 (cmH$_2$O)sL^{-1}, matched against the volume data represented by dots. In Figure 3(a), the terms P_0, P_m and P_k refer to Equation (11). At $t = t_v$, V is maximum and $\dot{V}$ is zero. This figure is adopted from our work in Ref [1].

Let us have some validation of the average values of C and R obtained by parameter-identification scheme, by determining the values of C and R at some specific time instants. For that purpose, we can put down from Equation (1-b),

$$R\ddot{V}+\frac{\dot{V}}{C}=\dot{P}_N(t) \tag{9}$$

Now the volume (V) curve in Figure 3(b) has an inflection point at $t = t_i = 1.18$ s, at which $\ddot{V}=0$. At $t = 1.18$ s, $V = 0.29$ L, $\dot{V}=0.48$ Ls^{-1}, $P_N = 2.53$ cmH$_2$O, and $\dot{P}_N = 1.66\,(\text{cmH}_2\text{O})\text{s}^{-1}$. Upon substituting these values into Equation (9), we get $C = 0.289$ L(cmH$_2$O)$^{-1}$. Then substituting this value of C along with the values of V, $\dot{V}$ and P_N into Equation (1-a), we get $R = 3.18$ (cmH$_2$O)sL^{-1}. These values of C and R at $t = t_i = 1.18$ s are of the same order of magnitude as the average values of C and R given by Equation (8). This provides us a measure of confidence to our parameter-identification scheme for obtaining the average values C_a and R_a.

Now since Lung disease will influence the values of R and C, these parameters can be employed to diagnose lung diseases. For instance in the case of emphysema, the destruction of lung tissue between the alveoli produces a more compliant lung, and hence results in a larger value of C. In asthma, there is increased airway resistance (R) due to contraction of the smooth muscle around the airways. In fibrosis of the lung, the membranes between the alveoli thicken and hence lung compliance (C) decreases. Thus by determining the normal and diseased ranges of the parameters R and C, we can employ this simple Lung-ventilation model for differential diagnosis.

3. Non-dimensional ventilatory index

For disease detection, it is more convenient to formulate and employ a non-dimensional number to serve as a ventilatory performance index LVI (to characterize ventilatory function), as:

$$LVI_1=\left[\left(R_aC_a\right)\left(\text{Ventilatory rate in s}^{-1}\right)60\right]^2=\tau_a{}^2\left(BR\right)^2 60^2 \tag{10}$$

where BR is the breathing rate.

Now, let us obtain its order-of-magnitude by adopting representative values of R_a and C_a in normal and disease states. Let us take the above computed values of $R_a = 2.275$ (cmH$_2$O)sL^{-1} and C_a =0.2132 L(cmH$_2$O)$^{-1}$ and $BR = 12$ m^{-1} or 0.2 s^{-1}, computed by matching Equation (6) to the data of Figure 3.

Then, in a supposed normal situation, the value of LVI_1 is of the order of 33.88. In the case of obstructive lung disease (with increased R_a), let us take $R_a = 5$ (cmH$_2$O)sL^{-1}, $C_a = 0.12$ L(cmH$_2$O)$^{-1}$ and $BR = 0.3$ s^{-1}; then we get $LVI_1 = 118.6$. For the case of emphysema (with enhanced C_a), let us take $R_a = 2.0$ (cmH$_2$O)sL^{-1}, $C_a = 0.5$ L(cmH$_2$O)$^{-1}$ and $BR = 0.2$ s^{-1}; then we obtain $LVI_1 = 144$. In the case of lung fibrosis (with decreased C_a), we take $R_a = 2.0$ (cmH$_2$O)sL^{-1}, $C_a = 0.08$ L(cmH$_2$O)$^{-1}$ and $BR = 0.2$ s^{-1}; then we obtain $LVI_1 = 3.7$.

We can, hence summarize that LVI_1 would be in the range of **2-5** in the case of fibrotic lung disease, **5-50** in normal persons, **50-150** in the case of obstructive lung disease and **150-200** for the case of emphysema,. This would of course be needed to be verified by analyzing a big patient population.

Now, all of this analysis requires pleural pressure data, for which the patient has to be intubated. If now we evaluate the patient in an outpatient clinic, in which we can only monitor lung volume and not the pleural pressure, then let us develop a non-invasive method for determining lung compliance (*C*), resistance-to-airflow (*R*) and ventilatory index.

4. For non-invasive assessment of lung status and determination of lung compliance and resistance-to-airflow

Our primary need is to be able to determine lung pressure $P_N(t)$ non-invasively. If we observe the $P_N(t)$ curve in Figure 3(a), we can note that, during the period of time from $t = t_i = 1.18$ s to $t = t_v = 2.02$ s, we can represent it as:

$$P = P_N = P_k \sin \omega_p (t - t_i) + P_0 \tag{11-a}$$

$$\cong P_k \sin \omega_p (t - 1.18) + 2.5 \tag{11-b}$$

where (i) $P_k = P_m - P_0$, and $P_k = 0.5$ cmH$_2$O in Figure 3(a), (ii) $t = t_i$, the inflection point on the lung volume curve (= 1.18 s in Figure 3(b)), and (iii) $P_0 = 2.5$ cmH$_2$O.

We can determine the value of ω_p, by invoking the condition that the pressure P becomes maximum (= P_m) at $t = t_m = (t_v + t_i)/2$. Hence, at $t = t_m$,

$$P_m = P_k \sin \omega_p (t_m - t_i) + P_0$$

But since $P_k = P_m - P_0$, we get

$$P_m - P_0 = (P_m - P_0) \sin \omega_p (t_m - t_i)$$

or

$$\sin \omega_p (t_m - t_i) = 1$$

wherein

$$t_m = (t_v + t_i) / 2 = (2.02 + 1.18) / 2 = 1.6$$

Therefore, we have

$$\omega_p (t_m - t_i) = \omega_p (1.6 - 1.18) = \pi / 2 = 1.57 \tag{12}$$

In Equation (12), $t_m = (t_v + t_i)/2$ and both t_v (= 2.02 s) and t_i(= 1.18 s) can be known from the lung volume curve. Hence, ω_p can be determined. For Figure 3 data, we get $\omega_p = 3.73$ rad/sec.

Hence, in our case of Figure 3 data, we can represent lung pressure P_N (= P) between t_i and t_v as:

$$P = P_k \sin \omega_p (t - t_i) + P_0 \tag{13-a}$$

$$= 0.5 \sin 3.73 (t - 1.18) + 2.5 \tag{13-b}$$

and

$$\dot{P} = 1.87\cos 3.73(t - 1.18) \tag{13-c}$$

So, in general, the parameters of the lung pressure curve are P_k and P_0, since t_i and ω_p can be determined in terms of t_v and t_i, as per Equation (12) and Figure 3. Likewise, we can also represent lung volume between t_i and t_v as:

$$V = V_T \sin \omega_v (t - t_i) + V_0 \tag{14-a}$$

whereV_T is the tidal volume and V_0 is the lung volume at $t = t_i$. Based on Figure 3(b) data, we can rewrite Equation (14-a) as follows:

$$V = 0.25 \sin \omega_v (t - 1.18) + 0.3$$

At $t = t_v = 2.02$ s, $V = V_T = 0.55$ L. Hence, $\omega_v(2.02 - 1.18) = 1.57$ (or $\pi/2$), so that $\omega_v = 1.87$ rad. So then Equation (14-a) can be written as:

$$V = 0.25 \sin 1.87(t - 1.18) + 0.3 \tag{14-b}$$

$$\dot{V} = 0.47 \cos 1.87(t - 1.18) \tag{14-c}$$

$$\ddot{V} = -0.88 \sin 1.87(t - 1.18) \tag{14-d}$$

Now based on Equation (13), we can represent the governing lung ventilation model Equation (1-b) as

$$R\dot{V} + \frac{V}{C} = P_K \sin 3.73(t - 1.18) + P_0 \tag{15-a}$$

or as,

$$R\dot{V} + \frac{V}{C} = 0.5 \sin 3.73(t - 1.18) + 2.5 \tag{15-b}$$

Let us employ this equation to determine the values of C and R at some specific points in the ventilation cycle. At $t = t_i$ (the inflection point) = 1.18 s, $\ddot{V} = 0$ Ls^{-2}, $\dot{V} = 0.48$ Ls^{-1} and $V = 0.3$ L. Now, we can differentiate Equation (15) as:

$$R\ddot{V} + \frac{\dot{V}}{C} = (P_K \omega_p) \cos 3.73(t - 1.18) \tag{16-a}$$

or as,

$$R\ddot{V} + \frac{\dot{V}}{C} = (0.5)(3.73) \cos 3.73(t - 1.18)$$

$$= 1.86 \cos 3.73(t - 1.18) \tag{16-b}$$

Hence, from this Equation (16), we get:

$$\frac{\dot{V}(=0.48)}{C} = 1.86 \text{, or } C = 0.258 \text{ L(cmH}_2\text{O)}^{-1}.$$

Then, upon substituting this value of C into Equation (15), we get:

$$R(0.48) + \frac{0.3}{0.258} = 2.5 \text{, or } R = 2.79 \text{ (cmH}_2\text{O)sL}^{-1}.$$

Hence at $t = t_i$ =1.18 s, C = 0.258 L(cmH$_2$O)$^{-1}$ and R = 2.79 (cmH$_2$O)sL^{-1}. (17)

Let us now evaluate C and R at $t = t_k$ =1.6 s, the time associated with the peak lung pressure. From Equations (15) and (16), we can put down:

$$R\dot{V} + \frac{V}{C} = 0.5\sin 3.73(1.6 - 1.18) + 2.5 = 3 \tag{18-a}$$

$$R\ddot{V} + \frac{\dot{V}}{C} = 1.86\cos 3.73(1.6 - 1.18) = 0 \tag{18-b}$$

At $t = t_k$ =1.6 s, we get from Equations (13) and (14), V = 0.48 L, $\dot{V} = 0.33$ Ls^{-1}, $\ddot{V} = -0.622$ Ls^{-2}, P = 3 cmH$_2$O and $\dot{P} = 0$ (cmH$_2$O)s^{-1}.

Substituting these values into Equations (18-a) and (18-b), we get:

$$0.33R + \frac{0.48}{C} = 3 \tag{19-a}$$

$$-0.622R + \frac{0.33}{C} = 0 \tag{19-b}$$

from which we obtain for $t = t_k$ =1.6 s,

$$C = 0.22\left(\text{cmH}_2\text{O}\right)\text{sL}^{-1}, R = 2.51\left(\text{cmH}_2\text{O}\right)\text{sL}^{-1} \tag{20}$$

Finally, let us evaluate C and R at $t = t_v$ =2.02 s. From Equation (15), we get:

$$R\dot{V} + \frac{V}{C} = 0.5\sin 3.73(2.02 - 1.18) + 2.5 \tag{21}$$

Now at $t = t_v$ =2.02 s, V = 0.55 L, $\dot{V} = 0$ Ls^{-1}. So then, from Equation (21), we obtain:

$$\frac{0.55}{C} = 2.5 \text{, or } C = 0.22 \text{ (cmH}_2\text{O)sL}^{-1} \tag{22}$$

It can be noted that the values of C and R given by Equations (17), (20) and (22) are similar to their average values $C_a = 0.218$ L(cmH2O)$^{-1}$ and $R_a = 2.275$ (cmH2O)sL^{-1}. This lends a measure of confidence to our Equation (15), for which V and $\dot{V}$ are given by Equations (14-b) and (14-c).

Let us now proceed to how we can determine the values of lung pressure function parameters P_k and P_0 along with R and C from the monitored values of lung volume.

At $t = t_i$ =1.18 s, we have from Equation (15-a)

$$R\dot{V} + \frac{V}{C} = P_K \sin 3.73(t - 1.18) + P_0 \tag{23}$$

so that by substituting $V = 0.3$ L, $\dot{V} = 0.48$ Ls^{-1}, we get:

$$R(0.48) + \frac{0.3}{C} = P_0 \tag{24}$$

Also, from Equation (16), we get by substituting the values of $\dot{V} = 0.48$ Ls^{-1} and $\ddot{V} = 0$ Ls^{-2}

$$\frac{0.48}{C} = 3.73 P_k \tag{25}$$

At $t = t_m$ =1.6 s, we have from Equations (15) and (16) as well as by substituting the values of $V = 0.48$ L, $\dot{V} = 0.33$ Ls^{-1} and $\ddot{V} = -0.622$ Ls^{-2}, we get:

$$0.33R + \frac{0.48}{C} = P_k + P_0 = P_m \tag{26}$$

$$-0.622R + \frac{0.33}{C} = 0 \text{, or } R = 0.53/C \tag{27}$$

Then at $t = t_v$ =2.02 s, we get from Equation (15), along with $V = 0.55$ L, $\dot{V} = 0$ Ls^{-1},

$$\frac{0.55}{C} = P_0 \tag{28}$$

We hence have Equations (24), (25), (26), (27) and (28) to solve and determine the best values of the four unknowns R, C, P_k and P_0. For this purpose, we define the ranges of these four terms, as:
R: 2.1, 2.2, 2.3 (cmH_2O)sL^{-1}; C: 0.20, 0.21, 0.22 L(cmH_2O)$^{-1}$; P_k = 0.4, 0.5, 0.6 (cmH_2O); P_0 = 2.4, 2.5, 2.6 (cmH_2O)
Then, in order to satisfy these equations, we obtain for the best values of R, C, P_k and P_0, based on the solution in Appendix, as

$$R = 2.29(\mathrm{cmH_2O})\mathrm{sL}^{-1},\ C = 0.22\,\mathrm{L}(\mathrm{cmH_2O})^{-1},\ P_k = 0.58(\mathrm{cmH_2O}),\ P_0 = 2.47(\mathrm{cmH_2O}) \tag{29}$$

As can be noted, these values of C and R correspond to the average values of C and R given by Equation (8). This then lends credibility to our procedure for non-invasive determination of C and R, for lung disease detection. This procedure enables us to in fact determine lung pressure toward evaluation of C and R.
Now since this procedure enables us to determine maximum lung driving pressure $P_m = P_k + P_0$, we can also formulate the non-dimensional lung ventilatory index as:

$$LVI_2 = \frac{R}{C}\frac{(TV)^2}{(P_m)^2}(BR)(60)^2 \tag{30}$$

wherein BR is in s^{-1}. For our case, R = 2.275 (cmH_2O)sL^{-1}, C = 0.2132 L(cmH_2O)$^{-1}$, P_m = 3 cmH_2O, and TV = 0.55 L. This gives LVI_2 = 25.8. By using this LVI_2 index, we can expect its value to be of the order of 30 for normal subjects, 300 for COPD patients, 5 for emphysema patients, and 100 in the case of lung fibrosis.

Here again, we need to determine ***LVI*** for normal lung states as well as for different lung disease states. We can then compare which of the formulas (10) or (30) enable better separation of lung disease states from the normal state.

Comments related to values of the ranges of the parameters: Before proceeding to the next section, let us address the basis of providing the above indicated ranges of parameters. The lung ventilation volume and driving pressure curves in Figure 3 are for a normal case. By carrying out this procedure for other normal subjects, we can define and confirm the above mentioned normal ranges for these parameters, for obtaining their best values.

Now then how do we distinguish subjects with disorders, such as obstructive lung disease (with increased value of R_a), emphysema (with enhanced value of C_a), lung fibrosis (with decreased value of C_a)? This can be made out from the shape and values of the lung ventilation volume curve. So then by repeating this procedure for subjects with these disorders, we will be able to characterize the shapes of the lung ventilation curves for normal subjects and for prescribing appropriate ranges of the parameters, for obtaining the best values of these parameters.

5. Lung-status evaluation and indicators for extubation of mechanically-ventilated COPD patients

5.1 Introduction

In mechanically ventilated patients with chronic-obstructive-pulmonary-disease (COPD), elevated airway resistance and decreased lung compliance (i.e., stiffer lung) are observed with rapid breathing. The need for accurate predictive indicators of lung-status improvement is essential for ventilator discontinuation through stepwise reduction in mechanical support, as and when patients are increasingly able to support their own breathing, followed by trials of unassisted breathing preceding extubation, and ending with extubation.

For this reason, we have developed an easy-to-employ lung ventilatory index (*LVI*), involving the intrinsic parameters of a lung ventilatory model, represented by a first-order differential equation in lung-volume response to ventilator driving pressure. The *LVI* is then employed for evaluating lung-status of chronic-obstructive-pulmonary-disease (COPD) patients requiring mechanical ventilation because of acute respiratory failure.

5.2 Scope and methodology

We recruited 13 mechanically ventilated patients with chronic obstructive pulmonary disease (COPD) in acute respiratory failure [3]. All patients met the diagnostic criterion of COPD. The first attempt of discontinuation (or weaning off the ventilator) for every patient was made within a short duration (not exceeding 88 hours). The patients in the study were between the ages of 54-83 years. All the patients were on synchronized intermittent mandatory ventilation (SIMV) mode with mandatory ventilation at initial intubation. Based on the physician's judgment, the modes were changed for eventual discontinuation of mechanical ventilation. The time period for recording observations was one hour. For all purposes in this study, a successful ventilator discontinuation is defined as the toleration to extubation for 24 hours or longer and a failed ventilator discontinuation is defined as either a distress when ventilator support is withdrawn or the need for reintubation. Our LVI was then employed to distinguish patients who could be successfully weaned off the mechanical ventilator.

Hence, the scope of this section is that we have developed a lung-ventilatory index (*LVI*), based on a lung-model represented by a first-order differential equation in lung-volume dynamics to assess lung function and efficiency in the case of chronic-obstructive-pulmonary-disease (COPD) patients requiring mechanical ventilation because of acute respiratory failure. Herein, we have attempted to evaluate the efficacy of the *LVI* in identifying improving or deteriorating lung condition in such mechanically ventilated chronic-obstructive-pulmonary-disease (COPD) patients, and consequently if *LVI* can be used as a potential indicator to predict ventilator discontinuation. In our bioengineering study of 13 COPD patients who were mechanically ventilated because of acute respiratory failure, when their *LVI* was evaluated, it provided clear separation between patients with improving and deteriorating lung condition. Finally, we formulated a lung improvement index (*LII*) representative of the overall lung response to treatment and medication, and a parameter *m* that corresponds to the rate of lung improvement and reflects the stability of lung-status with time. This chapter is based on our previous chapter [3] in the book on Human Respiration (edited by V. Kulish and published by WIT Press) and other works on this subject [4-9].

5.3 Lung ventilation model

From a ventilatory mechanics viewpoint, the lungs can be considered analogous to a balloon, which can be inflated and deflated (passively). The gradient between the mouth-pressure (P_{mo}) and the alveolar pressure (P_{al}) causes respiration to occur. During inspiration, $P_m > P_{al}$, which causes air to enter the lungs. During expiration P_{al} increases, and is greater than P_{mo}; this causes the air to be expelled out of the lungs passively. These pressure differentials provide a force driving the gas flow. The pressure difference between the alveolar pressure (P_{al}) and pleural pressure (P_p) counter balances the elastic recoil. Thus the assessment of respiratory mechanics involves the measurement of flows, volumes (flow integrated over time) and pressure-gradients. The lung ventilation model (shown in Figure 1) is based on the same dynamic-equilibrium differential equation (Equation 1-b), expressing lung volume response to pressure across the lung, as:

$$R\dot{V} + \frac{V}{C} = P_L(t) - P_{el0} = B\sin(\omega t) \tag{31}$$

wherein:

i. the total positive pressure across the lungs, $P_L = P_{mo} - P_p$, wherein P_p is determined by intubating the patient, and assuming that the pressure in the relaxed esophageal tube equals the pressure in the pleural space surrounding it.
ii. the parameters of the governing Equation (31) are lung compliance (C) and airflow-resistance (R), with both R and C being instantaneous values
iii. $V = V(t) - V_e$ (wherein V_e is the end-expiratory lung volume)
iv. P_{el0} is the end-expiratory pressure
v. the net driving pressure $P_N = P_L - P_{el0}$

Let B be the amplitude of the net pressure wave form applied by the ventilator, C_a be the averaged dynamic lung compliance, R_a the averaged dynamic resistance to airflow, the driving pressure P_L be given as $P_L = P_{el0} + B\sin(\omega t)$, and the net pressure P_N be given by $P_N = B\sin(\omega t)$, as depicted in Figure 4. The governing equation (31) then becomes:

$$R_a\dot{V} + \frac{V}{C_a} = P_N = B\sin(\omega t) \tag{32}$$

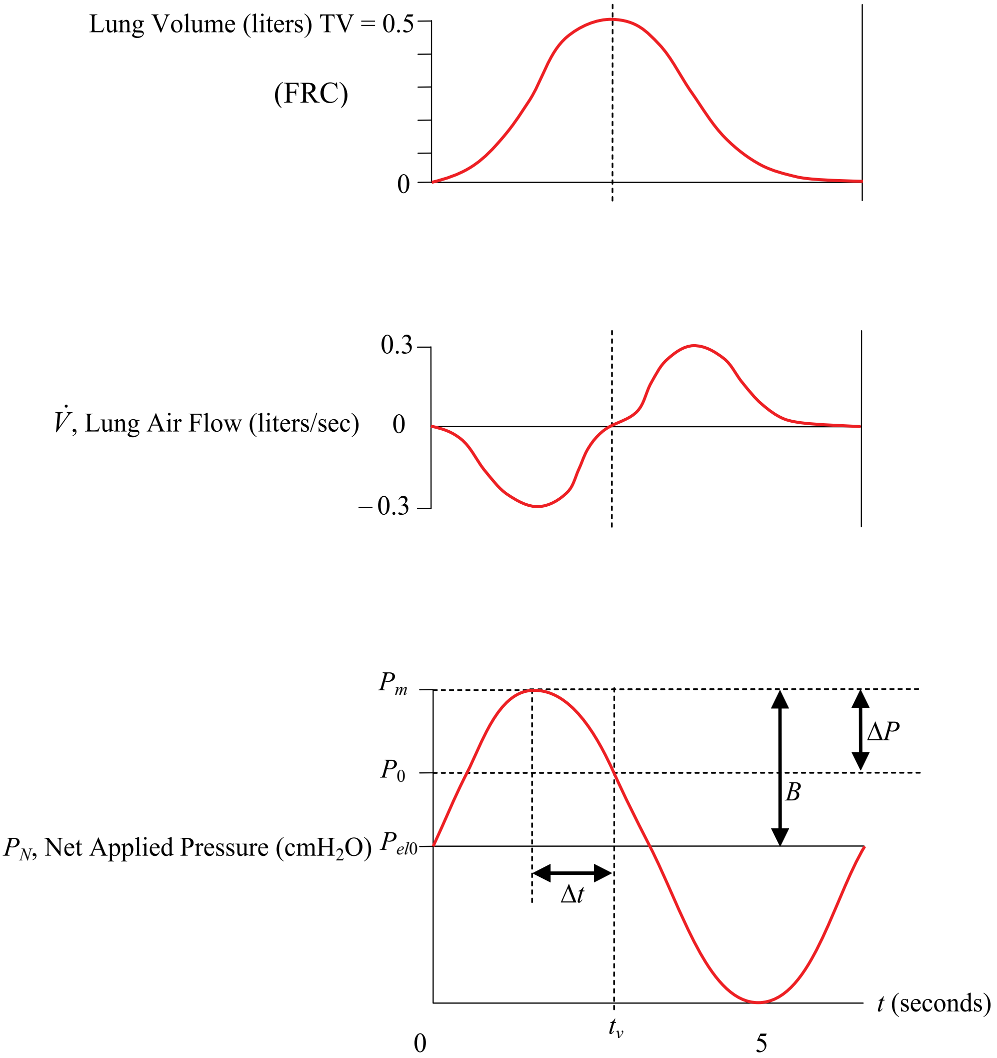


Fig. 4. Lung ventilatory model data shows air-flow ($\dot{V}$) and volume (V) and net pressure (P_N). Pause pressure (P_0) occurs at t_v, at which the volume is maximum (TV = tidal volume). Δt is the phase difference between the time of maximum volume and peak pressure (P_m). It is also the time lag between the peak and pause pressures. B is the amplitude of the net pressure waveform P_N applied by the ventilator. This P_N oscillates about P_{el0} with amplitude of B. The difference between peak pressure P_m and pause pressure P_0 is Δp. This figure is adopted from our work in Ref [3].

The volume response to P_N (the solution to Equation (32)) is given by:

$$V(t) = \frac{BC_a[\sin(\omega t) - \omega k_a \cos(\omega t)]}{1 + \omega^2 k_a{}^2} + He^{\left(\frac{-t}{k_a}\right)} \tag{33}$$

wherein:

i. k_a (= $R_a C_a$) is the averaged time constant,
ii. the integration constant H is determined from the initial conditions,
iii. the model parameters are C_a and k_a (i.e., C_a and R_a), and
iv. $\omega$is the frequency of the oscillating pressure profile applied by the ventilator

An essential condition is that the flow rate is zero at the beginning of inspiration and end of expiration. Hence, the flow rate $dV/dt = 0$ at $t = 0$. Applying this initial condition to our differential Equation (33), the constant H is obtained as:

$$H = \frac{BC_a \omega k_a}{1 + \omega^2 k_a{}^2} \tag{34}$$

Then, from Equations (33) and (34), we obtain:

$$V(t) = \frac{BC_a[\sin(\omega t) - \omega k_a \cos(\omega t)]}{1 + \omega^2 k_a{}^2} + \frac{BC_a \omega k_a}{1 + \omega^2 k_a{}^2} e^{\left(\frac{-t}{k_a}\right)} \tag{35}$$

For $t = t_v$, $V(t)$ is maximum and equal to the tidal volume (TV). Now in a normal person, k_a is of the order of 0.1 and 0.5 in ventilated patients with respiratory disorders, which is relevant to our study of COPD patients. At $t = t_v$ at which the lung volume is maximum, we note from Figure 4 that t_v is of the order of 2 s. Hence t_v/k_a is of the order 20-4, so that e^{-t/k_a} is of the order of e^{-20} to e^{-4}, which is very small and hence negligible. Hence, in Equation (35), we can neglect the exponential term so that,

$$V(t) = \frac{BC_a[\sin(\omega t) - \omega k_a \cos(\omega t)]}{1 + \omega^2 k_a{}^2} \tag{36}$$

Figure 4 illustrates a typical data of V, $\dot{V}$ and P_N. For evaluating the parameter k_a, we will determine the time at which $V(t)$ is maximum and equal to the tidal volume (TV), Hence, putting $dV/dt = 0$ in Equation (36), we obtain:

$$\cos(\omega t) + \omega k_a \sin(\omega t) = e^{\left(\frac{-t}{k_a}\right)}, \text{ for } t = t_v \tag{37}$$

Hence from Equation (37), we obtain the following expression for k_a:

$$\tan(\omega t) = -1 / \omega k_a, \text{ for } t = t_v \tag{38-a}$$

or,

$$k_a = -(1/\omega)\tan(\omega t_v) \tag{38-b}$$

Since both ω and t_v are known, we can evaluate k_a from Equation (38-b).
Now from Equation (38-b), we can put down:

$$\mathrm{TV} = \frac{BC_a[\sin(\omega t_v) - \omega k_a \cos(\omega t_v)]}{1 + \omega^2 {k_a}^2} \tag{39}$$

Since ω, t_v and k_a are known, we can now determine C_a in terms of TV and applied pressure amplitude B.
Then knowing k_a and C_a, we can determine

$$R_a = k_a / C_a \tag{40}$$

For our COPD patients, the ranges of the computed values of these parameters are:

$$R_a = 9 - 43\left(\mathrm{cmH_2 0}\right)\mathrm{sL}^{-1}; C_a = 0.020 - 0.080\ \mathrm{L}\left(\mathrm{cmH_2 0}\right)^{-1} \tag{41}$$

Now that we have determined the expressions for the parameters R_a and C_a, the next step is to develop an integrated index lung ventilatory incorporating these parameters.

5.4 Formulating a Lung Ventilatory Index (*LVI*) incorporating R_a and C_a

We believe that the correlations between average airflow-resistance (R_a), average lung-compliance (C_a), tidal volume (TV), respiratory rate (RF), and maximum inspiratory pressure or peak pressure (P_m) can be used as a possible indicator for determining lung-status in a mechanically ventilated COPD patient with acute respiratory failure. We hence propose that a composite index (LVI), incorporating these isolated parameters, can have a higher predictive power for assessing lung status and determining when a patient on a mechanical ventilator.
For this purpose, we note that COPD patients have higher R_a, lower C_a, lower TV, higher P_m and higher respiratory rate (or breathing frequency) RF. If we want the non-dimensional lung-ventilatory index (LVI) to have a high value for a COPD patient, further increasing LVI for deteriorating lung-status and decreasing LVI for improving lung-status in a mechanically ventilated COPD patient in acute respiratory failure, then the non-dimensional lung-ventilatory index (LVI) can be expressed, as given by Equation (30):

$$LVI_2 = \left[\frac{R_a(\mathrm{TV})^2(RF)}{C_a(P_m)^2}\right] \times (60)^2 \tag{42}$$

where RF is the respiratory-rate frequency.
Let us obtain the order-of-magnitude values of this LVI_2 index for a mechanically ventilated COPD patient in acute respiratory failure (by using representative computed values of the parameters R_a, C_a, RF, TV, and P_m), in order to verify that the formula for LVI_2 (given by Equation (42)) can enable distinct separation of COPD patients in acute respiratory failure from patients ready to be weaned off the respirator. For an intubated COPD patient, we have

$$LVI_2 \text{ (Intubated COPD)} = \frac{[15(\text{cmH}_2\text{O})\text{sL}^{-1}][0.5\text{ L}]^2[0.33\text{ s}^{-1}]}{[0.035\text{ L}(\text{cmH}_2\text{O})^{-1}][20\text{ cmH}_2\text{O}]^2} \times (60)^2 = 318 \tag{43}$$

wherein $R_a = 15$ (cmH$_2$O)sL^{-1}, $C_a = 0.035$ L(cmH$_2$O)$^{-1}$, $RF = 0.33$ s^{-1}, TV = 0.5 L and $P_m = 20$ cmH$_2$O.

Now, let us obtain the order-of-magnitude of *LVI* (by using representative computed values of R_a, C_a, *RF*, TV, and P_m) for a COPD patient with improving lung-status just before successful discontinuation. For a successfully weaned COPD patient (examined in an outpatient clinic), we have

$$LVI_2 \text{ (Outpatient COPD)} = \frac{[5(\text{cmH}_2\text{O})\text{sL}^{-1}][0.35\text{ L}]^2[0.33\text{ s}^{-1}]}{[0.10\text{ L}(\text{cmH}_2\text{O})^{-1}][12\text{ cmH}_2\text{O}]^2} \times (60)^2 = 50.5 \tag{44}$$

wherein $R_a = 5$ (cmH$_2$O)sL^{-1}, $C_a = 0.10$ L(cmH$_2$O)$^{-1}$, $RF = 0.33$ s^{-1}, TV = 0.35 L and $P_m = 12$ cmH$_2$O.

Hence for LVI_2 to reflect lung status improvement in a mechanically ventilated COPD patient in acute respiratory failure, there should be a pronounced decrease in the value of LVI_2. This shows that the *LVI* given by Equation (42) can enable effective decision making to wean off a COPD patient from mechanical ventilator.

Appendix:

Solution procedure to obtain the best values of R, C, P_k and P_0 provided in Equation (29).

For the four unknowns R, C, P_k and P_0, where

R: 2.1, 2.2, 2.3 (cmH$_2$O)sL^{-1};
C: 0.20, 0.21, 0.22 L(cmH$_2$O)$^{-1}$;
P_k: 0.4, 0.5, 0.6 (cmH$_2$O);
P_0: 2.4, 2.5, 2.6 (cmH$_2$O);

we want to find the best values of R, C, P_k and P_0 such that they satisfy the following equations:

$$R(0.48) + \frac{0.3}{C} = P_0 \tag{24}$$

$$\frac{0.48}{C} = 3.73P_k \tag{25}$$

$$0.33R + \frac{0.48}{C} = P_k + P_0 \tag{26}$$

$$-0.622R + \frac{0.33}{C} = 0 \tag{27}$$

$$\frac{0.55}{C} = P_0 \tag{28}$$

Solution:

We can rewrite Equations (26) to (30) so that the terms are collected at the LHS, i.e.,

$$R(0.48)+\frac{0.3}{C}-P_0=0 \tag{A-1}$$

$$\frac{0.48}{C}-3.73P_k=0 \tag{A-2}$$

$$0.33R+\frac{0.48}{C}-P_k-P_0=0 \tag{A-3}$$

$$-0.622R+\frac{0.33}{C}=0 \tag{A-4}$$

$$\frac{0.55}{C}-P_0=0 \tag{A-5}$$

Since we are trying to find the best values of R, C, P_k and P_0, the RHS of Equations (A-1) to (A-5) can be replaced by an error term so that

$$R(0.48)+\frac{0.3}{C}-P_0=e_1 \tag{A-6}$$

$$\frac{0.48}{C}-3.73P_k=e_2 \tag{A-7}$$

$$0.33R+\frac{0.48}{C}-P_k-P_0=e_3 \tag{A-8}$$

$$-0.622R+\frac{0.33}{C}=e_4 \tag{A-9}$$

$$\frac{0.55}{C}-P_0=e_5 \tag{A-10}$$

In general, we can represent the overall error incurred at any values of R, C, P_k and P_0 by an objective function F so that

$$F(R,C,P_k,P_0)=\sum_{i=1}^{5}\|e_i\| \tag{A-11}$$

We can then find the best values of R, C, P_k and P_0 but solving Equation (A-11) as an optimization problem with the aim to minimize F, subjected to the bounded constraints:

$2.1 \le R \le 2.3$
$0.20 \le C \le 0.22$
$2.4 \le P_0 \le 2.6$
$0.4 \le P_k \le 0.6$

The optimal solution is obtained at $F = 0.239539$, and the associated best values of R, C, P_k and P_0 are

$$R = 2.29425\ \mathrm{cmH_2O)sL^{-1}}$$
$$C = 0.219758\ \mathrm{L(cmH_2O)^{-1}}$$
$$P_k = 0.57926\ (\mathrm{cmH_2O})$$
$$P_0 = 2.46702\ (\mathrm{cmH_2O}) \qquad \text{(A-12)}$$

6. References

[1] Ghista D N, Loh K M, Lung Ventilation Modelling and assessment, in *Human Respiration: Anatomy and Physiology, Mathematical Modelling and Applications*, ed by Vladimir Kulish, WIT Press, 2006.

[2] Ghista D N, Sankaranarayanan M, Lung Ventilation Modeling for Lung Disease Diagnosis, in *Applied Biomedical Engineering Mechanics*, ed by Dhanjoo N. Ghista, CRC Press (Taylor & Francis Group), Boca Raton, Florida (US), 2009.

[3] Ghista D N, Pasam R, Vasudev S B, Bandi P, Kumar R V, Indicator for Lung-status in mechanically-ventilated COPD patients using Lung Ventilation Modeling and Assessment, in *Human Respiration: Anatomy and Physiology, Mathematical Modeling*, Numerical Simulation and Application, ed byVladimir Kulish, WIT Press (Southampton, UK), 2005.

[4] MacIntyre N R., Evidence-based Guidelines for Weaning and Discontinuing Ventilatory support by A Collective task Force of the American College of Chest Physicians, *Respiratory Care*, 47(1), 69-90, 2002.

[5] Menzies R, Gibbons W, Goldberg P., Determinants of weaning and survival among patients with COPD who require mechanical ventilation for acute respiratory failure, *Chest*, 95(2), 398-405, 1989.

[6] Yang KL, Tobin MJ., N Engl, A prospective study of indexes predicting the outcome of trials of weaning from mechanical ventilation, *J Med.*, 324(21), 1445-50, 1991.

[7] Alvisi R, Volta CA., Righini ER., Capuzzo M, Ragazzi R, Verri M, Candini G, Gritti G, Milic-Emili J, Predictors of weaning outcome in chronic obstructive pulmonary disease patients, *Eur Respir J.*, 15(4), 656-62, 2000.

[8] Pauwels RA., Buist AS., Calverley PM., Jenkins CR., Hurd SS., Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease, NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary, *Am J Respir Crit Care Med.*, 163(5), 1256-76, 2001.

[9] Meade M, Guyatt G, Cook D, Griffith L, Sinuff T, Kergl C, Mancebo J, Esteban A, Epstein S, Predicting success in weaning from mechanical ventilation, *Chest.*, 120(6) Suppl, 400S-24S, 2001.

[10] Ubran A, Van de Graaff WB, Tobin MJ, Variability of patient-ventilator interaction with pressure support ventilation in patients with chronic obstructive pulmonary disease, *Am J Respiratory Crit Care Med.*, 152(1), 129-136, 1995.

35

Physiological Nondimensional Indices in Medical Assessment: For Quantifying Physiological Systems and Analysing Medical Tests' Data

Dhanjoo N. Ghista
*Department of Graduate and Continuing Education,
Framingham State University, Framingham, Massachusetts,
USA*

1. Introduction

1.1 Concept of Non-Dimensional Physiological Indices (NDPIs) or Physiological Numbers (PHYNs)

In this chapter, we are providing a new concept in physiological systems analysis (or organ systems analysis), in terms of nondimensional physiological indices (or physiological system numbers), for qualifying patient health and disease status as well as patient improvement.

The concept of a Nondimensional Physiological Index (or NDPI) is quite new, and has been adopted from Engineering, wherein nondimensional numbers (made up of several terms) are employed to characterize disturbance phenomena. For example, in a cardiovascular fluid-flow regime, the Reynold's number

$$N_{re} = \rho VD/\mu \tag{1}$$

is employed to characterize the conditions when N_{re}exceeds a certain critical value, at which laminar blood flow changes to turbulent flow. This can occur in the ascending aorta when either the aortic valve is stenotic (giving rise to murmurs) or in the case of anaemia (decreased blood viscosity).

In physiological medicine, the use of nondimensional indices or numbers can provide a generalized approach by which unification or integration of a number of isolated but related events into one nondimensional physiological index (NDPI)can help to characterize an abnormal state associated with a particular organ or physiological system or an anatomical structure. The evaluation of the distribution of the values of such NDPI(s), in a big patient-population, can then enable us to designate normal and disordered ranges of NDPI, with a critical value of NDPI separating these two ranges. In this way, NDPI(s) can help us to formulate patient-health indices (PHIs), not only to facilitate differential diagnosis of patients but to assess the severity of the disease or disorder as well[1,2].

This chapter is based on the author's paper: Nondimensional Indices for Medical Assessment, in Mechanics in Medicine and Biology, Vol 9, No 4, 2009, World Scientific Publishers.

In medicine, assessment tests are carried out to (i) determine the functional performance of an organ (such as the heart) or a physiological system (such as the glucose-insulin regulatory system), and (ii) diagnose an anatomical structure's pathological condition, such as a calcified mitral valve or osteoporosis. In many cases, the medical tests do not quantifiably assess the concerned oral or physiological system and do not quantifiably diagnose the pathological condition of the anatomical structures.

So for some conventional tests (such as the Treadmill test to assess heart function, and Oral glucose tolerance test to diagnose diabetes), we have developed NDPI (s) made up of parameters that (i) are associated with the methodology of the tests, and (ii) characterise the function and disease states of organs (such as the heart) and physiological systems (such as the glucose-insulin regulatory system). We have also developed some new medical tests to detect anatomical structures' pathology (such as arteriosclerosis and mitral valve calcification) in terms of NDPI (s) to characterize their pathological state. In this chapter, we have formulated nine medical tests and their associated NDPI s).

We would like that the NDPI (s) developed in this chapter, for both conventional tests as well as newly formulated tests, can be applied clinically to set the stage for more accurate medical assessment. These medical tests and their associated NDPI (s) need to be applied to large patient population, to determine the normal and abnormal (or pathological) ranges of these NDPI (s). This will enable incorporation of our newly formulated NDPI (s) into medical practice.

2. Cardiac contractility index

Let us provide an example of one of our NDPI(s) which has been clinically employed. In cardiology, the index $(dP/dt)_{max}$ (of maximum rate-of-increase of left ventricular chamber pressure) has been traditionally employed as a measure of cardiac contractility. Diminished cardiac contractility affects cardiac output and can lead to heart failure. Hence, this is an important index of left ventricle (LV) functional capability. However, this index requires the invasive measurement of LV chamber pressure by catheterization. So, we developed an alternative cardiac contractibility index in terms of the normalised wall stress of the LV with respect to LV chamber pressure, $\sigma^* = \sigma / P$. Now, corresponding to $(dP/dt)_{max}$, we have formulated the cardiac contractility index (*CCI*) of $d\sigma^*/dt_{max}$, which does not require the measurement of LV chamber pressure. This contractility index can be conveniently expressed in terms of LV chamber cavity volume (V) and myocardial volume (V_m), as indicated in Ref 4. as well as in this section 5 of chapter 34 on how cardiac disease states cause decreased contractility and how surgical ventricular restoration improves contractility. By employing a thick-walled spherical LV model with inner and outer radii $r_i$and r_e, we can express LV pressure-normalised wall stress (σ^*) as [3,4]:

$$\sigma^*(\mathrm{r} = \mathrm{r_i}) = \left(\frac{3V}{2V_m} + \frac{1}{2} \right) \tag{2}$$

where $V(=4\pi r_i^3/3)$ denotes LV volume, and $V_m(=4\pi(r_e^3 - r_i^3)/3)$ denotes LV myocardial volume. Then by differentiating σ^* with respect to time, we get the expression for CCI as:

$$CCI, d\sigma^*/dt_{\max} = \left| \frac{d(\sigma_\theta / p)}{dt} \right|_{\max} = \frac{3}{2V_m} \left| \left(\frac{dV}{dt} \right) \right|_{\max} \tag{3a}$$

At the National Heart Center (in Singapore Gerneral Hospital), we have validated $d\sigma^*/dt_{max}$ against dP/dt_{max} in subjects with disparate ventricular function (as illustrated by figure 6 in chapter 34) and demonstrated the index's load independence. For normal subjects, this index value range is 4 – 4.5 s^{-1}. We have also successfully employed this $(d\sigma^*/dt)_{max}$ contractility index to assess (i) reduced LV contractility in ischemic cardiomyopathy patients in the range of 2. 64 ± 0.74s^{-1} and (ii) improved contractility in the range of 3.32 ± 0.73 s^{-1} following surgical ventricular restoration [5].

We could even divide this *CCI* index by the heart rate, and make it nondimensional, as:

$$CCI-2= \left(100\right)\, [d\sigma^*/dt_{max}] \;/\; \mathrm{HR}\left(\mathrm{in}s^{-1}\right) \tag{3b}$$

For normal hearts (with HR in the range of 60-80 per min), this index would give values in the range of 300-500, while for ischemic cardiomyopathic hearts (and HR greater than 100 per min) this index would be in the range of 200 and below. This *CCI-2* index would even more reliably represent cardiac contractility, by more distinctly differentiating ischemic and infracted hearts from normal hearts.

3. Assessing cardiac fitness and heart function

In this section, we show how the conventional Treadmill test can be formulated in biomedical engineering terms, to derive a cardiac fitness index to detect a malfunctioning heart due to, for instance, an infarcted heart caused by coronary occlusion.

The Treadmill test's cardiac fitness model consists of a first-order differential-equation system models, describing the heart rate (*HR*) response (*y*) to exertion (exercise, jogging, etc.) monitored in terms of a constant work-load or power (*W*) , where y is defined as follows [6,2]

$$y=\frac{HR(t)-HR(rest)}{HR(rest)} \tag{4}$$

In a Treadmill test, the subject is asked to exercise on the treadmill for a period of time, t_e(min). During this period, the *HR*(*t*) (and hence *y*) is monitored. Now we develop a model to simulate (i) the *HR*(*t*) response to during exercise, i.e. $t < t_e$ and (ii) thereafter for *HR*(*t*) decay after the termination of exercise (as illustrated in figure 1).

The DEq for y response to exercise on the treadmill at a constant work-load or power exerted (W) is given by

$$\frac{dy}{dt}+k_1y=C_0W \text{ , for } t \le t_e$$

$$\frac{dy}{dt}+k_2y=0 \text{ , for } t \ge t_e \tag{5}$$

where (1) k_1 and k_2 are the model parameters, and (2) C_0 is a conversion factor to express *W* in the same units as the other terms of the equation. The *y* solutions to equations (5) are represented by:

$$y=\frac{C_0W}{k_1}\left(1-e^{-k_1t}\right) \text{, for } t \le t_e \text{ during the exercise} = \frac{y_e\left(1-e^{-k_1t}\right)}{\left(1-e^{-k_1t_e}\right)} \tag{6}$$

$$y = y_e e^{-k_2(t-t_e)} \text{, for } t \geq t_e \text{ during the recovery period when } W=0 \tag{7}$$

where $y_e = y(t=t_e)$, and k_1 and k_2 are the model parameters which can serve as cardiac fitness parameters (in units of min^{-1}).

We now carry out parametric simulation to the monitored HR data on treadmill, by making equations (6 and 7) fit the HR data. The parameters k_1 and k_2 can be combined into a non-dimensional fitness index CFI given by:

$$CFI = k_1 k_2 t_e^2 \tag{8}$$

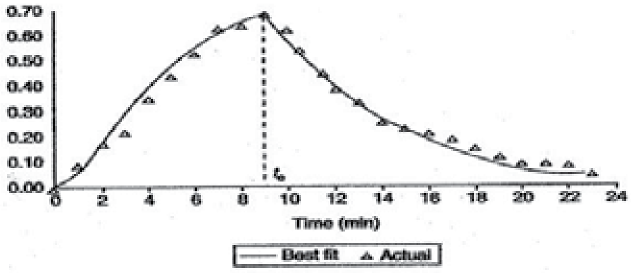


Fig. 1. Sample subject's monitored *y* versus *t* data (Adopted from Ref 6: Lim GeokHian, Dhanjoo N. Ghista, Koo TseYoong, John Tan Cher Chat, Philip EngTiew& Loo Chian Min; *International Journal of Computer Application in Technology: Biomedical Engineering & Computing Special Issue,* Vol 21, No 1/2, 2004.).

According to this formulation of *CFI*, for subjects exercised at identical workloads, a healthier subjects would have (1) greater k_1 (i.e., slower rate of increase of HR during exercise), (2) greater k_2 (i.e., faster rate of decrease of HR following exercise), (3) greater t_e(i.e., exercise endurance); and hence (4) higher value of *CFI.*

We need to evaluate *CFI* for a big spectrum of patients, and then compute its distribution curve, to determine the efficacy of this index, in order to yield distinct separation of *CFI* ranges for healthy subjects and cardiac patients. This *CFI* can then also be employed to assess improvement in cardiac fitness following cardiac rehabilitation regime.

4. Lung ventilation Index to detect lung disorders

Herein, we are formulating a new test involving: (i) monitoring of lung volume by means of a spirometer; (ii) the biomedical engineering model of the lung volume response to lung inflation driving pressure, which is equal to mouth pressure minus pleural pressure monitored by placing a balloon catheter transducer through the nose into the esophagus; (iii) derivation of the lung ventilation index made up of the parameters of the lung volume response model, and its employment to detect lung disorders.

The differential equation of lung ventilation volume (V) response to lung inflation-pressure (P_L) is given (as illustrated in figure 2) by [1, 2]:

$$\begin{aligned} R\frac{dV}{dt} + \frac{V}{C} &= P_D = \left(P_0 - P_p\right) - \left|\left(P_p\,@end-\exp iration\right)\right| \\ &= P_1 - P_1\cos\omega t + P_2\cos\omega t \end{aligned} \tag{9}$$

wherein *V* is the lung volume in litres (L), P_0 is the presence at the mouth(in cm H_20) and P_p is the pleural pressure (in cm H_20). The right-hand side terms constitute the fourier series representation of the lung driving pressure (P_D) in cm H_20, *R* is the resistance to airflow (in cm $H_2O \cdot S \cdot L^{-1}$), *C* is the lung compliance (in L/cm H_2O), and P_1 and P_2 are the magnitudes of fourier series terms of the lung driving (oscillatory) pressure P_D [= (mouth-pressure minus pleural-pressure), with respect to the absolute value of end-expiratory pleural pressure].
For a typical P_D cyclic pressure profile (Fig.2), represented by

$$P_1 = 1.84 \text{cm } H_20, \quad P_2 = 3.16 \text{cm } H_20, \quad \omega = 0.5\pi s^{-1},$$

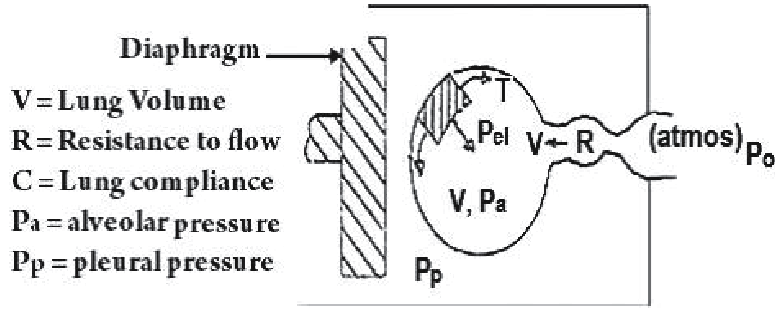
the solution to the above Eq. (9) is given by

$$V = P_1C\left(1-e^{-t/\tau}\right) - P_1C\frac{\left(\cos\omega t + \omega\tau\sin\omega\tau\right)}{1+\omega^2\tau^2} + P_2C\frac{\left(\sin\omega t + \omega\tau\cos\omega\tau\right)}{1+\omega^2\tau^2} + \frac{e^{-t/\tau}}{\left(1+\omega^2\tau^2\right)}\left[P_1C\left(1+2\omega^2\tau^2\right) + P_2C\left(\omega\tau\right)\right] \tag{10}$$

wherein τ = RC. By fitting this lung volume solution to the clinically monitored lung-volume parameters, we get: $R = 1.24(\text{cm } H_20)\ sL^{-1}$, $C = 0.21L(\text{cm } H_20)^{-1}$.
Now, then, let us formulate the nondimensional lung ventilator performance index (LVPI-1) given by:

$$\text{LVPI-1} = RC\ (\text{BR per min}), \tag{11}$$

wherein BR, the breathing rate =30ω/π per min = 15/min or 0.25/sec for the data provided in Fig.2. For our case study, the value of *LVPI* is 3.9.



$$\begin{aligned} P_L &= P_o - P_p \\ &= (P_o - P_a) + (P_a - P_p) = (P_o - P_a) + P_{el} \\ &= \left(R\frac{dV}{dt}\right) + \left(\frac{V}{C} + P_{el} @ t = T\right) \\ &= \left(R\frac{dV}{dt}\right) + \left(\frac{V}{C} - P_p @ t = T\right) \end{aligned}$$

$$\text{Driving Pressure, } P_D = P_L - \left|P_p @ t = T\right|$$

Fig. 2. Lung Ventilation Lumped parameter model.

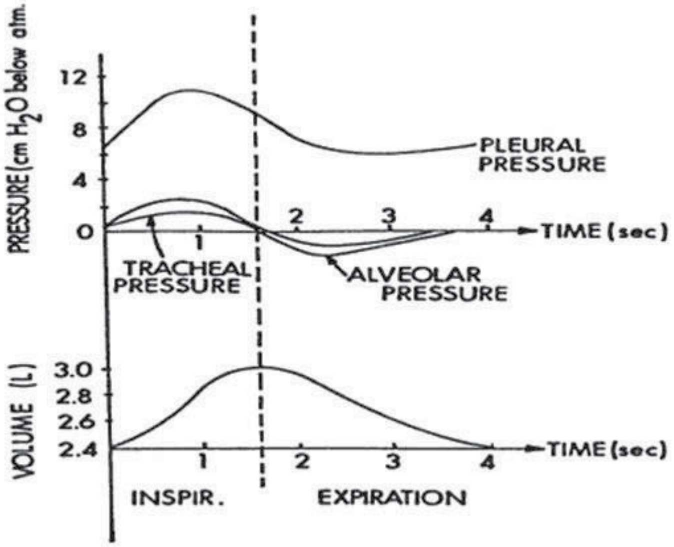


Fig. 3. Lung pressure and volume as functions for normal breathing; note that the pressure extremes occur before the volume extremes.

Let us now see how lung disease will influence *R, C* and hence *LVPI*. For instance in emphysema, the destruction of lung tissue will produce a more compliant lung and hence a larger value of C = 0.5L (cm H_20)$^{-1}$, say, yielding a value of *LVPI* of about 10. In asthma, there is increased airway resistance (due to contraction of the smooth muscles around the airways) to say R=5(cm H_20) sL^{-1}. The breathing rate can go also go up to BR = 20/min, say. Hence, the value of *LVPI* can go up to 20. In the case of lung-congestion, due to mitral-valve disease, it would be important to determine *LVPI*, so as to serve as an indicator for determining cardiac condition (in end-stage heart-disease). By determining the distribution of a big patient population, we can determine the *LVPI* ranges for normal and disease states.

Now this procedure refers to monitoring of pleural pressure (P_P) by placing a balloon catheter transducer through the nose into the esophagus (assuming that the esophageal tube pressure equals the pressure in the pleural space surrounding it). So, let us now develop a procedure whereby we do not need to monitor pleural pressure and only need to monitor lung volume by means of a spirometer. For this purpose, we now identify three model parameters (P_1C), (P_2C) and τ . These parameters can be determined by having equation (10) match or stimulate the lung volume data in figure 3. We now formulate a non-invasively determinable nondimensional ventilator index (*LVPI-2*), as

$$LVPI-2=\frac{(BR)\tau(TV)^2}{(P_1C)(P_2C)}\text{, } TV\text{= tidal volume} \qquad (12)$$

Upon evaluating *LVPI-2* for a number of patients, we can determine its ranges for normal and disease states, to employ it diagnostically. We can expect that subjects with chronic obstructive lung disease (COPD) and asthma subjects (with a high value of *R*) will have a high value of *LVPI -2*, while emphysema subjects (with high value of *C*) will have a low value of *LVPI* -2. So in the distribution curve of *LVPI* -2, emphysema subjects will be at the low end of distribution curve, COPD and asthma subjects will be at the high end of the distribution curve, while normal subjects will be in the middle of the distribution curve.

Now for noninvasive assessment of lung disease state in terms of lung compliance (C) and resistance-to-flow (R), we need to be able to determine lung pressure (P_D) function noninvasively. In section 4 of chapter 36 on lung ventilation modeling for assessment of lung station we have shown how we can determine the lung driving pressure functional parameters along with C and R in termo of the monitored values of lung volume.

We have also formulated a non-dimensional lung ventilatory index (equation 30) in terms of R, C, tidal volume (TV), lung pressure value at TV and breating rate. After this index is evaluated for different disease states, it will enable reliable noninvasure assessment of lung status.

5. Diabetes diagnosis from oral-glucose-tolerance test by means of a diabetes index (OGTT)

In this section, we have developed a biomedical engineering model for OGTT and demonstrated how it can be applied to OGTT data in (figure 4) to formulate and evaluate a diabetic NDPI, to more reliably diagnose diabetes.

For oral-glucose tolerance test simulation (entailing digestive and blood-pool chambers), the differential equation, and governing blood-glucose response (y) to oral ingestion of glucose bolus (G, gm $L^{-1}hr^{-1}$), given by[7]:

$$y'' + 2Ay' + \omega_n^2 y = G\delta(t)\text{ ; } y \text{ in gL}^{-1}\text{, } G \text{ in gL}^{-1}\text{hr}^{-1}$$

$$y'' + \lambda T_d y' + \lambda y = G\delta(t) \tag{13}$$

where ω_n (= $\lambda^{1/2}$) is the natural oscillation-frequency of the system, A is the attenuation or damping constant of the system, ω = $(\omega_n^2 - A^2)^{1/2}$ is the (angular) frequency of damped oscillation of the system, λ $(2A/T_d = \omega_n^2)$ is the parameter representing regulation proportional to rate-of–change of glucose concentration (*y*), and λT_d is the parameter representing regulation proportional to rate-of-change of glucose concentration (y').

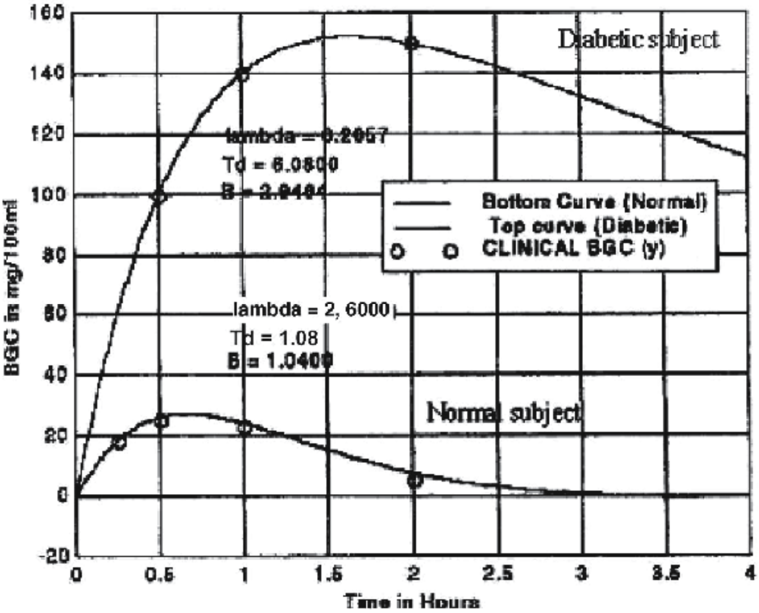
Figure 4 illustrates the OGTT data for typical normal and diabetic subjects. For an impulse glucose ingestion input, we can simulate a normal patient's blood-glucose concentration databy means of the solution of the Oral glucoseregulatory (second-order system) model, as an under-dampedglucose-concentration response curve, given by:

$$y(t) = \left(\frac{G}{\omega}\right) e^{-At} \sin \omega t, \tag{14}$$

When this solution is made to simulate the normal subjects OGTT data, we get A=1.4 hr^{-1}, ω = 0.775 rad/hr, G = 1.04 gL^{-1} hr^{-1}, λ=2.6 hr^{-2}, T_d= 1.08 hr. The simulated curve is also depicted in figure 4.

For a potential diabetic subject, we adopt the solution of the above Differential equation model, as an over-damped response function:

$$y(t) = \left(\frac{B}{\omega}\right) e^{-At} \sinh \omega t. \tag{15}$$



$$y(t) = \frac{G}{\omega} e^{-At} \sinh \omega t$$

$$(AT_d = 4.9)$$

($A = 0.808\ \mathrm{hr}^{-1}, \lambda = 0.2657\ \mathrm{hr}^{-2}$,
$T_d = 6.08$ hr,
$G = 2.9464\ \mathrm{gL}^{-1}\ \mathrm{hr}^{-1}$)

$$y(t) = \frac{G}{\omega} e^{-At} \sin \omega t$$

$$(AT_d = 1.3)$$

($A = 1.4\ \mathrm{hr}^{-1}, \lambda = 2.6\ \mathrm{hr}^{-2}$
$T_d = 1.08$ hr,
$G = 1.04\ \mathrm{gL}^{-1}\ \mathrm{hr}^{-1}$

Fig. 4. OGTT Response Curves: $A = 1.4\mathrm{hr}^{-1}$ (i.e. higher damping-coefficient value) for the normal subject for the diabetic patient, $A = 0.808\ \mathrm{hr}^{-1}$. Also, for the normal subject, the regulation parameters λ and λT_d are 2.6 hr^{-2} and 2.8 hr^{-1} respectively, which are greater than their values of 0.26 hr^{-2} and 1.62 hr^{-1} for the diabetic subject. Further, the non-dimensional number for the normal subject is 1.3, compared to 4.9 for the diabetic subject.

For this OGTT data simulated function (figure 4), the parameters values are: $A = 0.808\ \mathrm{hr}^{-1}$, $\omega = 0.622$ rad hr^{-1}, $G = 2.95\ \mathrm{gL}^{-1}\ \mathrm{hr}^{-1}$, $\lambda = 0.266\ \mathrm{hr}^{-2}$, $T_d = 6.08$ hr.
Now, we come to the interesting part of this model, by formulating the nondimensional Diabetes index (*DBI*) as:

$$DBI = AT_d = \frac{2A^2}{\lambda} = \frac{2A^2}{\omega_n^2}. \tag{16}$$

The value of *DBI* for the normal subject is found to be 1.5, whereas that for the diabetic subject is 4.9. It is further seen (in our initial clinical tests) that *DBI* for normal subjects is < 1.6, while the *DBI* for diabetic patients is > 4.5. This is a testimony of the efficacy of the model, and especially for the nondimensional*DBI*.
Now between these two cases of under-damped and over-damped responses, we have the case of a critically-damped response, for which the solution of the OGTT model differential equation (13) is given by

$$y(t) = Gte^{-At}, \text{ for which } \omega = 0, \text{ and } \mathrm{A}^2 = \lambda = \omega_n^2$$

This critically-damped response corresponds to cases of subjects who are not distinctly normal or diabetic but are at the risk of becoming diabetic. It can be seen that *DBI* for the critically-damped response is 2. So, in the distribution curve of *DBI* (to be obtained by applying this method to a large patient population), the *DBI* range of less than 1.6 would refer to normal subjects, the range of greater than 4.5 would refer to diabetic subjects, and range of 2-4 would refer to subjects at risk of becoming diabetic. This would make the use of the model and the *DBI* to be so convenient for the physician.

6. Characterization of arterial stiffness or arteriosclerosis by means of NDI

In this section, we are formulating a new test to noninvasively determine the arterial constitutive property so as to be able to diagnose arteriosclerosis.
For a circular cylindrical arterial tube of radius *a* and wall-thickness *h*, we can express the stress $\sigma$and elastic-modulus *E*, as follows:

$$\sigma = \frac{Pa}{h} = \frac{130Pa}{h} N / m^2; \; E = \frac{2(PWV)^2 a\rho}{h}; \; E = E_0 + m\sigma \tag{17}$$

in terms of (i) the arterial dimensions *a* and *h*, the auscultatory (or automatedly) measurable diastolic pressure (*P*) and pulse-wave velocity (*PWV*) determined by ultrasound [8]. The table below then depicts the computed values of σ and *E* at four independent times.

P (mmHg)	PMV (m/s)	A (mm)	h (mm)	$E\left[\frac{N}{m^2}\right]$	$\sigma\left[\frac{N}{m^2}\right]$
80	5.3	4.1	1.10	2.13×10^5	3.38×10^4
85	5.4	4.5	1	2.6×10^5	4.97×10^4
90	5.42	5.0	0.90	3.01×10^5	5.97×10^4
95	5.5	5.0	0.90	3.38×10^5	6.68×10^4

Table 1.

$$\text{Result}: \; E\left(N/m^2\right) = 4.2\sigma + 0.5\text{X}10^5\left(N/m^2\right) = mo + E_0. \tag{18}$$

We will now define the arteriosclerotic non-dimensional index

$$ART - NDI = mE_0/(\text{mean diastolic pressure}) \tag{19}$$

For the above patient, the value of the *ART – NDI* is

$$ART - NDI = \frac{(4.2)\left(0.5\text{X}10^5 \, N/m^2\right)}{\left(87\text{X}137 \, N/m^2\right)} = 17.6 \tag{20}$$

and will be much higher for arteriosclerotic patients, which we will determine by conducting clinical tests-applications of this analysis. This *ART-NDI* to detect arteriosclerosis requires echocardiographic determination of arterial dimensions and PWV[8], and auscultatory diastolic pressure.

7. To non-invasively determine aortic elasticity (m), peripheral resistance (R), aortic NDI, and aortic pressure profile

Herein, we have developed the analysis to noninvasively determine the aortic pressure profile, which can have significant diagnostic applications. This analysis is also employed to determine (i) aortic volume elasticity parameter *m (=dP/dV),* (ii) periphal resistance parameters R=P(pressure)/Q(flow rate), and (iii) the aortic property NDPI, given by aortic number. Based on the aorta fluid mechanics model (figure 5), we obtain:

$$\frac{dV}{dt} = I(t) - Q(t) = I(t) - P(t)/R \tag{21-a}$$

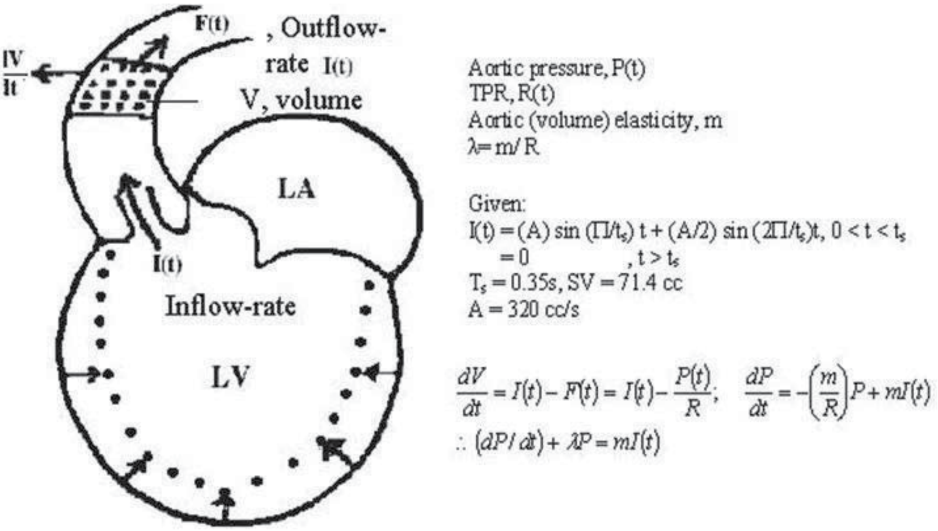
$$\frac{dP}{dt} = \frac{dP}{dV}\frac{dV}{dt} = m\frac{dV}{dt} \tag{21-b}$$

We can then put down the aortic pressure (P) response to aortic inflow-rate or LV outflow-rate *I(t)* as follows [9,1] :

$$\frac{dP}{dt} + \lambda P = mI(t); \tag{22}$$

wherein, $m =$ Volume elasticity of aorta (in Pa/m³), and $\lambda = (m/R)$ in s^{-1}.
The LV outflow-rate is represented as:

$$\begin{aligned} I(t) &= (A)\sin(\pi/t_s)t + (A/2)\sin(2\pi/t_s)t, \quad 0 < t < t_s \ (systole) \\ &= 0; \quad t > t_s \ (diastole) \end{aligned} \tag{23}$$



Aortic pressure, P(t)
TPR, R(t)
Aortic (volume) elasticity, m
$\lambda = m/R$

Given:
$I(t) = (A)\sin(\Pi/t_s)t + (A/2)\sin(2\Pi/t_s)t, \ 0 < t < t_s$
$= 0 \qquad , t > t_s$
$T_s = 0.35\mathrm{s}$, $SV = 71.4$ cc
$A = 320$ cc/s

$$\frac{dV}{dt} = I(t) - F(t) = I(t) - \frac{P(t)}{R}; \quad \frac{dP}{dt} = -\left(\frac{m}{R}\right)P + mI(t)$$
$$\therefore (dP/dt) + \lambda P = mI(t)$$

Fig. 5. To derive the equation for Aortic-pressure response to the stroke-volume or LV output rate *I(t)*.

If t_s= 0.35s, and Stroke vol(*SV*) is known (from, say, echocardiography), then we have

$$\int_0^{t_s}\left[(A)\sin(\pi/t_s)t + (A/2)\sin(2\pi/t_s)t\right] = SV \tag{24}$$

wherein $A = \pi(SV)/2t_s$

$$\text{So if} \quad SV = 71.4cc, \quad \text{then} \quad A = 320cc/s. \tag{25}$$

The solutions of Eq. (22) for the aortic diastolic and systolic periods are obtained as follows [9,1]:

Aortic Diastolic Pressure expression (fig 6):

$$\begin{aligned} &P_d(t) = P_z e^{-\lambda(t-t_s)};\ P_2 = \ \textit{pressure at start of diastole} \\ &= P_1(at\, t = T) = \ \textit{pressure at end of systolic phase} \\ &\therefore P_d(t) = P_1 e^{\lambda(T-t)}, \quad \textit{wherein} \ \ T = 0.8s. \end{aligned} \tag{26}$$

This $P_d(t)$ function is equal to P_2 at t= t_s (at the end of systolic phase) and P_1 at t=T (end of diastotic phase).

Aortic Systolic Pressure expression (fig 6):

$$\begin{aligned} P_s(t) = &\left(P_1 + \frac{Am\omega}{\lambda^2 + \omega^2} + \frac{2Am\omega}{\lambda^2 + 4\omega^2} \right) e^{-\lambda t} + mA\left(\frac{\lambda \sin \omega t - \omega \cos \omega t}{\lambda^2 + \omega^2} \right) \\ &+ \frac{mA}{2}\left(\frac{\lambda \sin 2\omega t - 2\omega \cos 2\omega t}{\lambda^2 + 4\omega^2} \right); \quad \omega = \frac{\pi}{t_s}. \end{aligned} \tag{27}$$

This $P_s(t)$ function is maximum at t=t_m, and equal to the monitored systolic auscultatory pressure P_3 = 118mmHg.

We now (i) incorporate into Eqs. (26) and (27) the auscultatory data of P_1(80mmHg) and P_3 (118mmHg) with T = 0.8 s, and us=0.35s, as well as (ii) invoke continuityin diastolic and systolic pressure expressions, to (iii) put down and solve the following three equations (in three unknowns: m, $\lambda$and t_m which P_s= P_2):

$$P_d(\text{at}_s = 0.35\text{s}) = P_s(\text{at}_s = 0.35\text{s}) \tag{28}$$

$$\frac{dP_s(\text{at}t_m)}{dt} = 0 \tag{29}$$

$$P_s(t = t_m) = P_3(= 118\text{mmHg}) \tag{30}$$

to obtain: λ=0.66 s^{-1}, m = 0.78 mmHg cm^{-3}, R = 1.18mmHg cm^{-3} s, t_m= 0.25 s, for T = 0.8 s.

We now formulate the Aortic number (or index):

$$\text{Aortic number} = \lambda T = \text{mT/R} \tag{31}$$

whereinλ= m/R in the governing differential equation (22), m =103 x 10^6 Pa m^{-3} , R=157x 10^6 Pa m^{-3} s, and λ =*0.66*s^{-1}

We thereby obtain the $\textit{Aortic Number} = \lambda\text{T} = \left(0.66\text{s}^{-1}\right)(0.8\text{s}) = 0.52.$ (32)

In order to have a more convenient order-of-magnitude value of the Aortic number index (equation 31), we could employ the *Aortic number* = 100 (λT). In the distribution of *Aortic* *number* (obtained by applying this methodology to a large patient population), the low range of Aortic Number will correspond to patients with vasoconstriction, the high range of Aortic Number will associated with arteriosclerotic patients, and patients with normal healthy aorta will be in the middle of the distribution.

Finally with the help of the evaluated parameters m and R, we can now construct the aortic pressure profile based on equations (26 and 27), as illustrated in figure 6. This aortic pressure profile can have significant diagnostic implications. As we know, in Ayurvedic medicine and Chinese Traditional medicine, the physician feels the pulsation of the patients brachial artery (just proximal to the wrist), and based on it provides diagnosis of a wide spectrum of diseases. Essentially, the physician is feeling the magnitude and shape of the arterial pressure pulse.

Now, we have shown that we can in fact noninvasively determine the aortic pressure profile, which is more diagnostically indicative than the pressure pulse in the more distally located brachial artery. Hence, we can bring to bear this medical inferential and experiential knowledge to firstly characterise the magnitude and shape of the aortic pressure profile (by Fourier analysis), and then correlate the Fourier series parameters to the information about the associated disease states available from Ayurvedic and Chinese medicine systems.

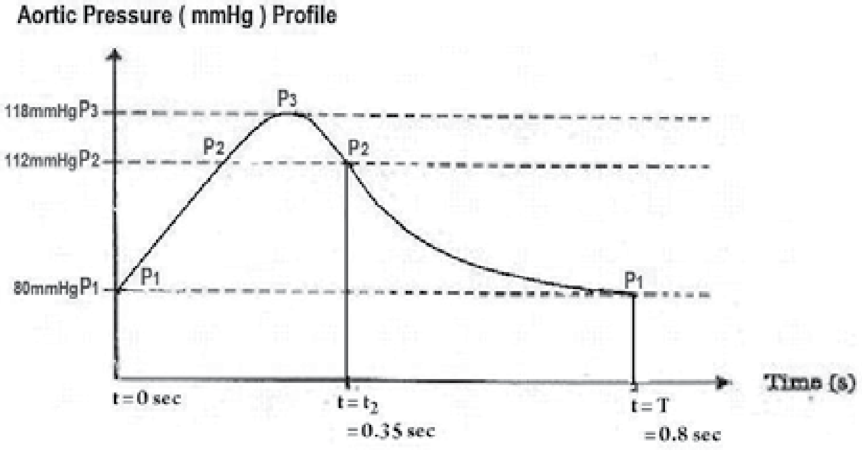


Fig. 6. Illustration of the Aortic pressure profile, based on the analysis. The systolic phase is from t = 0 to t = ts = 0.35s. The diastolic phase is from t = ts = 0.35s to t = T= 0.8s.

8. Mitral-Valve (MV) property determination for its pathology characterization (to provide interventional guidelines)

Determining the in-vivo constitutive property of the mitral valve (for a quantifiable estimate of its calcification and degeneration) constitutes another example combining “clinical-data monitoring and processing” with “modelling-for-clinical diagnosis”.

The mitral valve opens at the start of the diastolic phase when the blood from the left atrium fills the left ventricle (LV). At the end of the diastolic phase and at the initiation of LV contraction phase, the rising LV pressure and the blood flow pattern in the LV chamber brings the valve cusps together to close the MV, and set its cusps into vibratory motion, which is monitored as the First Heart sound (FHS).

From a biomedical engineering consideration, the mitral valve in its closed position (at the end of the diastolic phase) can be modelled as a semi-circular membrane, which is fixed along its circular edge to the heart chamber wall and supported along its straight edge by the chordae tendineae (as depicted in figure 7), so that its deflection is zero along its edges [10,11]. In this configuration, the MV vibrates after its cusps come together to close the valve. The frequency of MV vibration (f_{mv}) (as obtained from the FHS frequency spectrum) can be expressed in terms of the MV constitutive parameter property (E *vs*σ) , which conveys information about its health state and pathology. This methodology provide a more reliable and quantitative approach for detecting a pathological MV (such as owing to its leaflets calcification) than by merely listening to the First Heart sound (aspractised clinically).

To this end, we provide the expressions for determining MV stress (σ) and its elastic modulus (E) from the physiological data of the MV vibration its closed configuration. We then develop expressions for mitral valves modulus-based property E^* and stress-based property σ *, and propose that the E^* *vs*σ * relationship be employed to characterise mitral valve pathology. Alternatively, we can also track mitral valve pathological deterioration, by monitoring the changes in valve of m (= E/σ) in terms of the changes in f_{mv} as the valve pathology progresses, and determine the time for intervention of replacing the pathological MV by means of a prosthetic MV.

We make use of:

- echocardiography (to determine the mitra-valve geometry) and spectral phonocardiography (of the first-heart sound associated with MV vibration), to determine the second-peak frequency (f_2) of the first heart-sound spectrum
- along with static and dynamic (vibration) analyses of the semi-circular mitral valve leaflet model (held along its circular boundary), as illustrated in Fig. 7, to obtain the following expressions [10 & 11], for

Stress (σ) in the leaflet membrane

$$= \frac{\pi_2 f_2^2 a^2 \rho}{\left(K_{nm}/2\right)^2}, \tag{33}$$

wherein a is the radius of the semi-circular leaflet; ρ is the leaflet membrane density per unit area; K_{nm} is the mth zero of the nth order Bessel function $J_n(k)$, m (number of nodal circles = 1, and (number of nodal diameters)=1, and k_{11} = 3.832.

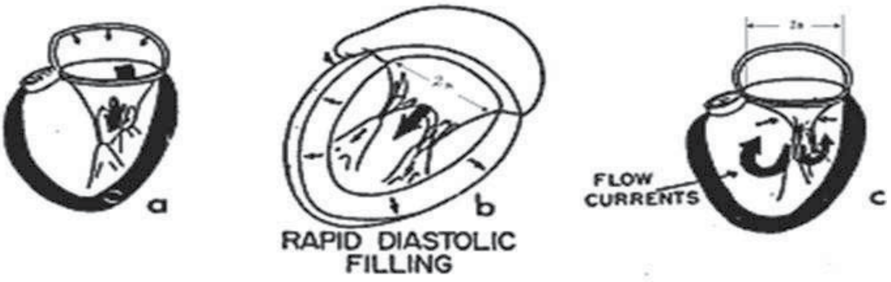


Fig. 7. The mechanism of MV closure and subsequent vibration, in the genesis of the first-heart sound. Functional mechanics of the Mitral valve: (a) mitral valve opening at start of left centricular diastole; (b) as the filling left ventricle distends, traction is applied through the chordae tendineae to the valve cusps, pulling them together; (c) at the start of LV systole, the valve cusps are sealed together by the high internal pressure and the flow pattern in the ventricular chamber. It is at this point in time that the mitral valve starts vibrating.

Modulus (E) of the leaflet membrane $= \dfrac{\pi^8 f_2^2 \rho^3 t^2 a^4 (1-\nu)}{\left(K_{11}/2\right)^6 q_0^2 S_n},$ (34)

Wherein t = leaflet thickness, ν is the Poisson's ratio, q_o is the pressure difference across the leaflet at time of occurrence of the closed Mitral-valve (MV) Vibration, and S_n (the summation of a series) = 0.105.

Based on eqs. (33) and (34), the nondimensional constitutive parameter (m) of the MV, is given by

$$m = \frac{E}{\sigma} = \frac{3\pi^6 f_2^4 \rho^2 t^2 a^2 (1-\nu)}{q_0^2 S_n \left(K_{11}/2\right)^4} \tag{35}$$

As a matter of interest, for the data: $f_2 = f_{mv}$ = 100Hz, q_0 = 2 mm Hg, ρ= 1.02 gm/cm^3, a = 1 cm, t = 0.5 mm, υ = 0.5, and evaluating S_n (= 0 · 0234, Eq 2 · 16, Ref.2), we get $\sigma = 2.75 \times 10^3$ N/m^2 and E=1.6x10^5 N/m^2.

Now, changes in MV pathology will affect its density(ρ) and thickness (t), as well as its modulus (E) vs stress (σ) property which we want to determine by combining FHS power-spectrum analysis (to determine f_{mv}) and 2-d echocardiographic analysis (to determine the size parameter a).

We now designate a new stress-based property (σ*) of MV (from equation 33), as

$$\sigma^* = \frac{\sigma}{\rho} = \frac{\pi^2 f_{mv}^2 a^2}{(1.916)^2} \tag{36}$$

as well as a new modulus –based property (E*) of MV (from equation 34), as

$$E^* = \frac{E q_0^2}{\rho^3 t^2} == \frac{\pi^8 f_{mv}^6 a^4 (1-\nu)}{(1.916)^2 S_n}, \tag{37}$$

We can now employ this E^* *vs.* σ^* relationship as a constitutive property of MV, to characterize and track its degeneration for timely intervention purpose.

This technology and methodology can provide the basis for timely surgical and/ or replacement intervention for a diseased MV. In order to apply this analysis, we can determine the valvular leaflet size parameter (a) from 2-D echocardiograms. The valvular leaflet vibrational frequency (f_{mv}) can be obtained from the frequency spectra of the FHS phonocardiographic signal associated with MV movement.

We can study a number of patients and determine the in vivo (E^*,σ^*) values of their valves, at a regular intervals during their degeneration process. We can also simultaneously and regularly monitor cardiac symptoms and chamber sizes and correlate them with the valcular constitutive E^*- σ^* property. By means of these correlations, we can determine the critical (E^*- σ^*) boundary at which intervention will have to be made to replace the degenerated natural valve by means of a prosthetic flexible leaflet MV.

In an alternative somewhat simpler approach, the mitral valve constitutive property parameter m (equation 35) can be employed diagnostically to track the deterioration due to calcification of the MV, in terms of Δmaccording to the relationship:

$$\Delta m = \left(\partial m/\partial f_{mn}\right)\Delta f_{mn} + \left(\partial m/\partial q_0\right)\Delta q_0; \text{ where in } f_{mv} = f_2 \tag{38}$$

$$\text{so that: } \frac{\Delta m}{m} = 4\left(\frac{\Delta f_{mn}}{f_{mn}}\right) - 2\left(\frac{\Delta q_0}{q_0}\right), \tag{39}$$

$$\text{or, } \frac{m'(= m + \Delta m)}{m} = 1 + 4\left(\frac{\Delta f_{mn}}{f_{mn}}\right) - 2\left(\frac{\Delta q_0}{q_0}\right) \tag{40}$$

Now at the time of occurrence of the first heart sound (FHS), the differential pressure or loading (q_0) across the mitral valve is very small. Hence the change in pressure loading valve (Δq_0), over the period of time of patient-tracking, will also be small compared to Δf_{mv}, and hence can be neglected in equation (40).
Hence from Eq. (39), we can compute

$$\frac{\Delta m}{m} = \left(\frac{4\Delta f_m}{f_m}\right) \tag{41}$$

to represent the change (Δm) in the parameter (m), by merely monitoring the change in frequency (Δf_{mv}) with respect to its earlier value (f_{mv}).

9. Noninvasive determination of bone osteoporosis index (in terms of bone flexural stiffness) for osteoporosis detection

Osteoporosis is a metabolic bone disease that is characterised by decreased bone mineral content and associated decreased in its mechanical strength. Thus, the osteoporotic bone is more prone to fracture.
Noninvasive measurement methods for osteoporosis detection include single and dual beam photon absorptiometry and a comparatively low cost low-frequency mechanical vibration (resonance and impedance) method [1, 12]. The low-frequency impedance response curve of (the first bending mode of) ulna yields the resonant frequency (f_r) value, which can be formulated in terms of the mechanical properties of the ulna bone, namely its bending stiffness (*EI)* and mass (*M*). It has been found that the difference between normal and osteoporotic bone is 20% in resonant frequency (f_r) and 80% in bending stiffness *EI* [13]. This is because f_r is the ratio of bone stiffness (EI) to mass (M), and in the pathologic osteoporotic condition both stiffness and mass decrease. Also, it has been shown that in fresh canine bone, the bending moment causing fracture has a correlation with EI of r =0.96 and with bone mineral content of r=0.90. Thus based on these results it appears that the ulna bending stiffness EI is a good indicator of bone fracture strength, which is diminished in osteoporosis. Now, both EI and M are contained in the expression for the natural frequency (f_r) of ulna vibrations, which in turn can be obtain from its resonance frequency [14]. In order to determine the resonance frequency of the ulna bone, it can be simply supported at its extremities and a vibrating probe pressed against the skin at the center of the forearm (as carried out by Steel and Gordon [14] and schematised in Fig. 8). The resonance frequency f_r (= natural frequency f_n of vibration of ulna) is obtained from the recording of the acceleration response as a function of the frequency.
If the bone is vibrating at an angular frequency p, the weight of the ulna bone per unit length is w radius of the ulna is R, and its length is ℓ, then the natural frequency f_n of the vibrating ulna beam, with its mass concentrated in the middle is given by :

$$f_n = p / 2\pi = \sqrt{(g / \delta st)} / 2\pi \tag{42}$$

where δ_{st}, the maximum central deflection of the simply supported ulna bone, is given by

$$\delta_{\mathrm{st}} = \mathrm{w}\ell^4 / (77\mathrm{EI}) \tag{43}$$

wherein w is the ulna weight per unit length.
Hence, from (42) and (43), we get

$$f_n = \frac{1}{2\pi}\left(\frac{77gEI}{wl^4}\right)^{1/2} \tag{44}$$

By putting $w = \rho A g$ ρ: bone density, A: cross-sectional area, we get

$$f_n = \frac{1}{2\pi}\left(\frac{77EI}{\rho A l^4}\right)^{1/2} = 1.4\left(\frac{EI}{Ml^3}\right)^{1/2}, \tag{45}$$

where M is the mass of the ulna bone.
By altering the frequency of the vibrating probe, we set the ulna into resonance, and the resonance frequency will be equal to the natural frequency. For f_r resonance frequency = f_n = 400Hz, $A = 50 \times 10^{-4}\mathrm{m}^2$, $I = 3 \times 10^{-8}\mathrm{m}^4$, length ($\ell$) = 0.17m, $\rho = 1.8 \times 10^3$ kg/m^3, we get $E = 20 \times 10^9$ N/m^2 and EI = 30Nm^2.
Thus, from equation (45), by modeling the ulna bone as a simply-supported vibrating beam, and determining its natural transverse-vibrational frequency (equal to its measured resonance frequency f_r), we can measure its flexural stiffness EI, to detect osteoporosis.

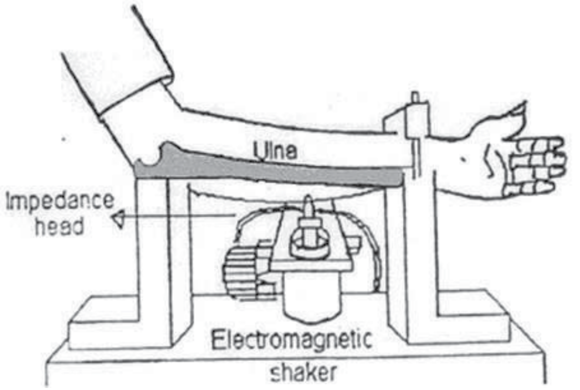


Fig. 8. Set-up used by Steel and Gordon [14] to determine the impedance of ulna. In this set up, the impedance head is attached to the moving element in the shaker. The probe, which contracts the skin, is attached to the impedance head.

10. Cardiac assessment based on Myocardial infarct detection and Intra-ventricular flow and pressure determination

In cardiology, a primary disorder is that of a heart with infarcted myocardium. This infarcted myocardial wall mitigates adequate contraction of the wall. So, the end-result of an infarcted left ventricle (LV) is poor intra-LV velocity distribution and pressure-gradient distribution, causing impaired outflow from the LV into the aorta.

In the infarcted myocardial segments, the myocardial infrastructure of actin and myosin filaments (and their cross – bridges) is disrupted, and hence there is no contraction within these infarcted myocardial segments. Figure 9 [15] illustrates a myocardial sarcomere segment's bioengineering model, composed of two symmetrical myocardial structural units (MSUs). In these MSU(s), the contractile elements represent the actin-myosin contractile components of the sarcomere segment.

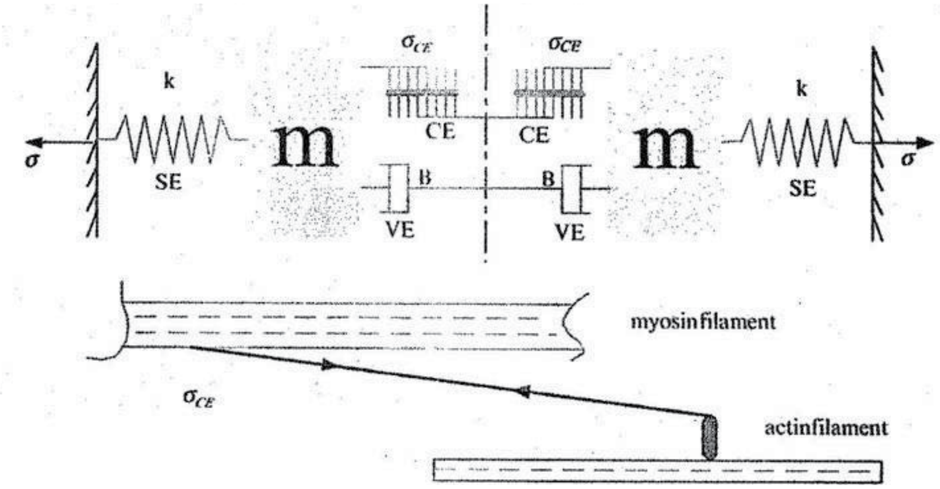


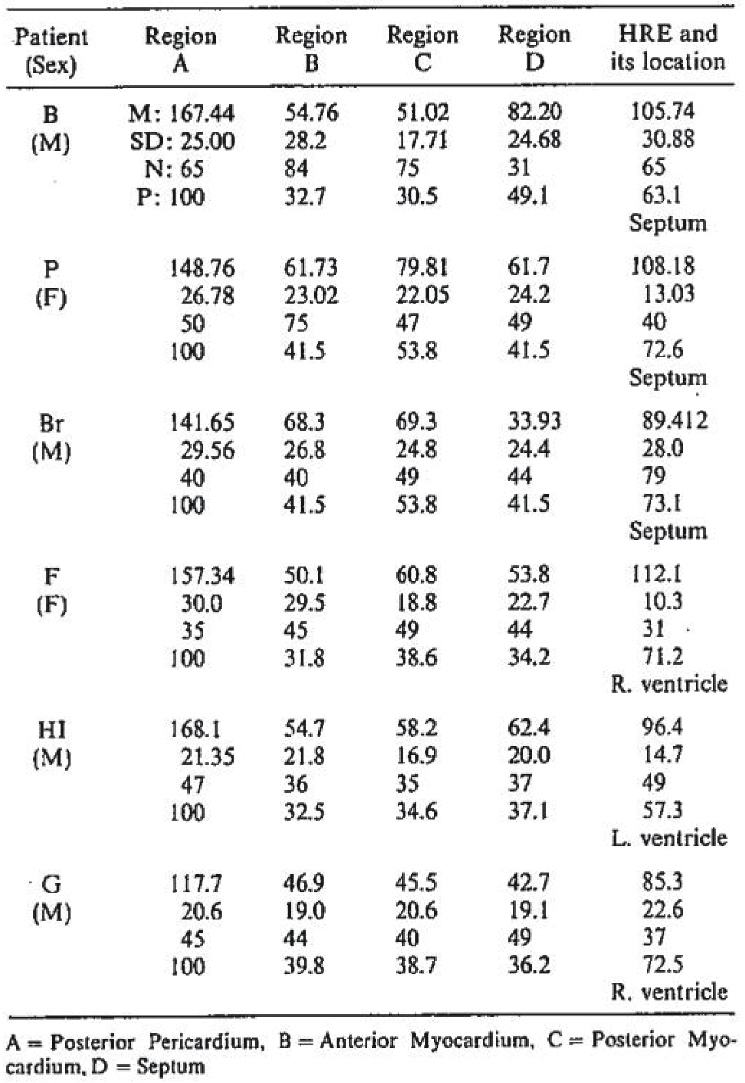
Fig. 9. Based on the conventional Hill three-element model and Huxley cross bridge theory, we have developed a myocardial model involving the LV myocardial mass, series-elastic element (CE). In this figure we have linked the anatomical associations of these myocardial model elements with microscopic structure of the heart muscle. This figure illustrates the sarcomere element contractile model, involving: the effective mass (*m*) of the muscle tissue that is accelerated; elastic parameter k of the series element stress σ_{SE} (k = elastic modulus of the sarcomere) viscous damping parameter B of the stress σ_{VE} in the parallel viscous element VE, the generated contractile stress σ_{CE} between myosin (thick) and actin (thin) filaments.

The disruptions of these contractile elements impairs the contractile capability of that sarcomere segment. Hence, a LV with infarcted myocardial segments will have diminished contractility, inadequate and improper intra-LV flow, and poor ejection.

Detection of myocardial infarcted segments: Now, infarcted myocardial segments can be detected as highly reflectile echo zones (HREZs) in 2- dimensional B-scan echocardiograms. In this context, we have shown earlier [16] how infarcted myocardial segments can be detected (in shape and size), by echo-texture analysis, as highly reflectile echo zones or HREZs. Now, each tissue component of the heart generates a grey scale pattern or texture related to the tissue density and fibrous content, and hence tissue stiffness.

In diseased states (such as myocardial ischemia, myocardial fibrosis, and infiltrative diseases), changes in myocardial tissue stiffness have been recognised by employing echo intensity and mean grey level of pixel as the basis for recognition of such myocardial disorders. It was found that hyper- reflectile echoes (HREs) correlated well with diseased cardiac muscle, and that myocardial tissue containing HREs corresponded with foci of sub endorcardial necrosis and even calcification.

In our earlier study [15], in order to determine highly reflectile echo zones (HREZs), echocardiograms were recorded; each image was made up of 256 x 256 pixels, with each pixel having a resolution of 0-256 grey scales. The echocardiographic images were digitised into 256 grey scales. Then, echo intensity levels from normal infants were used to delineate the range of echo intensities for normal tissues. The upper bound of the echo intensity was set to 100 per cent in each normal infant, and the intensities from the rest of the image was referenced to this level. Normally, pericardium had the highest intensity level. It was found that the upper- bound of the echo intensity value for healthy tissue (expressed as a percentage of the pericardial echo intensity value) was 54.2.



Patient (Sex)	Region A	Region B	Region C	Region D	HRE and its location
B	M: 167.44	54.76	51.02	82.20	105.74
(M)	SD: 25.00	28.2	17.71	24.68	30.88
	N: 65	84	75	31	65
	P: 100	32.7	30.5	49.1	63.1
					Septum
P	148.76	61.73	79.81	61.7	108.18
(F)	26.78	23.02	22.05	24.2	13.03
	50	75	47	49	40
	100	41.5	53.8	41.5	72.6
					Septum
Br	141.65	68.3	69.3	33.93	89.412
(M)	29.56	26.8	24.8	24.4	28.0
	40	40	49	44	79
	100	41.5	53.8	41.5	73.1
					Septum
F	157.34	50.1	60.8	53.8	112.1
(F)	30.0	29.5	18.8	22.7	10.3
	35	45	49	44	31
	100	31.8	38.6	34.2	71.2
					R. ventricle
HI	168.1	54.7	58.2	62.4	96.4
(M)	21.35	21.8	16.9	20.0	14.7
	47	36	35	37	49
	100	32.5	34.6	37.1	57.3
					L. ventricle
G	117.7	46.9	45.5	42.7	85.3
(M)	20.6	19.0	20.6	19.1	22.6
	45	44	40	49	37
	100	39.8	38.7	36.2	72.5
					R. ventricle

A = Posterior Pericardium, B = Anterior Myocardium, C = Posterior Myocardium, D = Septum

Table 1. Echo intensity values for various anatomic regions of diseased pediatric hearts (based on long axis view). The numbers in the four rows represent Mean (M), Standard Deviation (SD), Number of Pixels (N), Percentage of Posterior Pericardial Intensity (P).

For patients whose echo-texture analysis showed presence of HREs, it was found that the echocardiographic intensities of the HREs from these patients intensities), were distinctly higher than the echo intensity range of normal tissue (as depicted in Table 1).

Figure 10(a) depicts an echo image of an infant with visible scars regions 1 and 2, while figures 10(b) depict printouts of the echo intensities from these two regions, wherein the infarcted segments are depicted in dark colour.

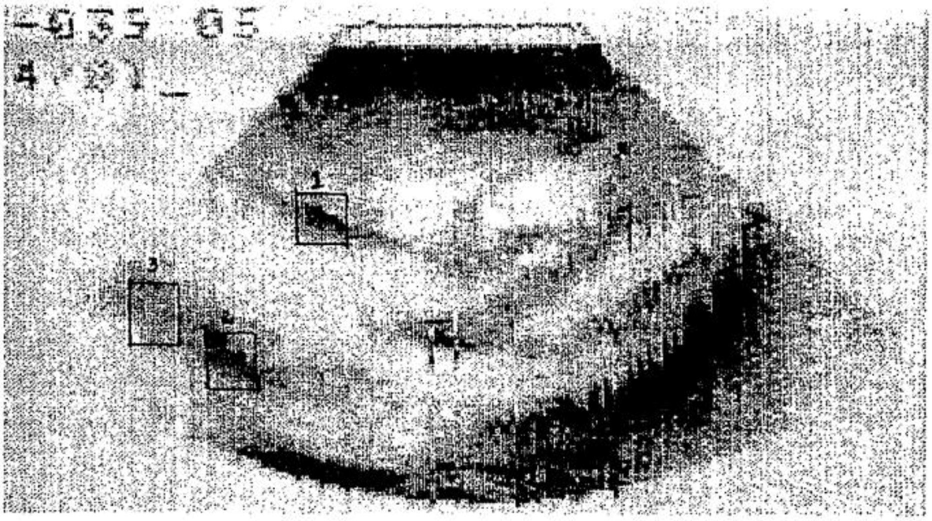
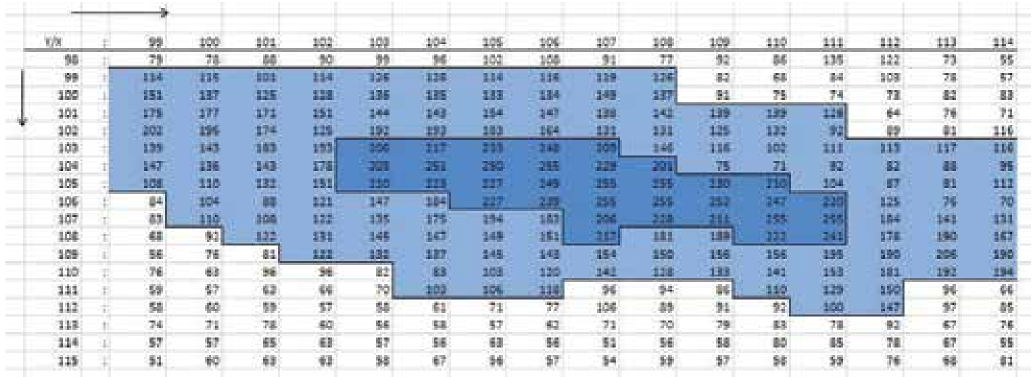


Fig. 10. (a) Long axis view of a pediatric patient's heart showing HRE regions 1 and 2 and a healthy region 3.(Adopted from the author's paper Ref 16).



Y/X	:	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114
98	:	79	78	88	90	99	96	102	108	91	77	92	86	135	122	73	55
99	:	114	115	101	114	126	128	114	116	119	126	82	68	84	103	78	57
100	:	151	137	125	128	136	135	133	134	149	137	91	75	74	73	82	83
101	:	179	177	171	151	144	143	154	147	138	142	139	139	128	64	79	71
102	:	202	196	174	125	192	193	183	164	131	131	125	132	92	89	81	116
103	:	139	143	183	193	206	217	233	248	209	146	116	102	111	113	117	116
104	:	147	136	143	178	209	251	250	255	229	201	75	71	82	82	88	95
105	:	108	110	132	151	230	223	237	249	255	255	230	210	104	87	81	112
106	:	84	104	88	121	147	184	227	239	255	255	252	247	220	125	76	70
107	:	83	110	108	122	135	175	194	183	206	228	211	155	255	184	141	131
108	:	68	93	122	131	146	147	149	151	217	181	188	222	241	178	190	167
109	:	56	76	81	122	132	137	145	143	154	150	156	156	195	190	206	190
110	:	76	63	98	96	82	83	103	120	142	128	133	141	153	181	192	194
111	:	59	57	62	66	70	103	106	118	96	94	86	110	129	150	96	66
112	:	58	60	59	57	58	61	71	77	106	89	91	92	100	147	97	85
113	:	74	71	78	60	56	58	57	62	71	70	79	83	78	92	67	76
114	:	57	57	65	63	57	56	63	56	51	56	58	80	85	78	67	55
115	:	51	60	63	63	58	67	56	57	54	59	57	58	59	76	68	81

Fig. 10. (b) Pixel values corresponding to highly reflectile echo region 1. The central region having echo-intensity values greater than 200 is infarcted. (Adopted from the author's paper, Ref 16).

Myocardial tissue pixels having echo-intensity values greater than 200 were designated to be infarcted. This infarcted sub-region is seen to be surrounded by an ischemic sub-region whose pixels have echo intensity values between 100 and 200. The surrounding healthy tissue has echo intensity less than 100.

In this way, in each highly reflectile echo zone (HREZ) made up of , say, N number of pixels, we can determine the number (I) of infarcted pixels. The ration I/N represents the infarcted potion of that HREZ myocardial segment. The total number of all the infarcted pixels in all the HREZs provides an indication of the amount of infarcted myocardium of the heart or of the LV.

Intra-LV Blood Flow velocity and pressure distribution: Now, let us come to the outcome of an infarcted heart and LV. Figure 11 illustrates this outcome in the form of intra-LV blood-flow velocity and pressure (or pressure-gradient) distributions [17]. During LV diastole, from the monitored entrance velocity of blood at the mitral valve and the wall motion of the expanding LV, we can compute the intra-LV blood-flow velocity and pressure distributions, by computational fluid dynamics (CFD). During systole, from the monitored exit velocity or the aortic valve and the wall motion of the contracting LV, we can compute the intra-LV blood-flow velocity and pressure distribution by CFD.

Figure 11 illustrates the computed intra-LV blood-flow velocity and pressure distribution of a patient with an infarcted myocardium, before and after administration of nitroglycerin to determine the viability of the myocardial wall following bypass surgery. Referring to Fig 11, for the patient (with a myocardial infarct), Figs. 11(a1) depict super-imposed LV outlines at known equal intervals during diastole and systole before nitroglycerin administration, and Figs 11 (a2) depict super-imposed LV outlines at known equal intervals during diastole and systole after nitroglycerin administration; nitroglycerin is a myocardial perfusing agent, and hence a quasi-simulator of coronary bypass surgery or coronary angioplasty. From these images, we can determine the instantaneous wall displacements and hence the wall velocities at these time instants.

This data, along with the monitored entrance and exit velocities of blood into and from the LV, constitutes the data for our CFD analysis. For computational purposes, the intra-LV flow is determined from the boundary condition of LV wall-motion velocity and inlet/outlet blood flow velocity to the standard potential-flow equation $\nabla^2\Phi = 0$. The intra-LV pressure gradient can then be computed from the Bernoulli equation for unsteady potential flow.

Figures 11(b1) and 11 (c1) depict intra-LV blood-flow velocity distributions during diastole and systole, before nitroglycerin administrations.

Figures 11(b2) and 11 (c2) depict intra-LV blood-flow velocity distributions during diastole and systole, after nitroglycerin administrations

Figures 11(d1) and 11 (e1) depict intra-LV pressure distributions during diastole and systole, before nitroglycerin administrations

Figures 11(d2) and 11 (e2) depict intra-LV pressure distributions during diastole and systole, after nitroglycerin administrations

In this patient, the poor motion of the infarcted LV wall offers resistance to proper filling of the LV (Fig 11-b1). However, it can be noted that following the administration of nitroglycerin, there is improved filling of the LV (Fig. 11-b2). During systolic ejection phase, the infarcted LV wall segments do not contract, and this results in inadequate intra-LV flow velocity distribution, which mitigates adequate emptying of the LV (Fig 11-c1). Following nitroglycerin administration, there is improved outflow velocity distribution. Likewise, figures (11-d1 and 11-e1) demonstrate adverse intra-LV pressure gradients during filling and ejection phases, which are improved after administration of nitroglycerin (Eq 11-d2 and 11-e2). This has provided the basis for advocating coronary revascularization by coronary bypass surgery, for this patient.

The computed intra-LV blood-flow velocity and pressure distributions provide illustrative and quantitative outcome of an infarcted LV to the physician, which enables more distinct assessment of LV dysfunction. The cause of this LV dysfunction is provided by the echo-texture analysis of 2-d B-scan echocardiograms of HREZ(s), in terms of the amount (or number of echocardiogram image pixels) of the infarcted myocardial wall. Together, these two methodologies provide reliable and quantitative assessment of (i) how much of the LV myocardium is infarcted and its effect on the intra-LV blood flow and pressure- gradient, and (ii) intra-LV distributions of blood-flow velocity and pressure distributions, to assess candidacy for coronary bypass surgery.

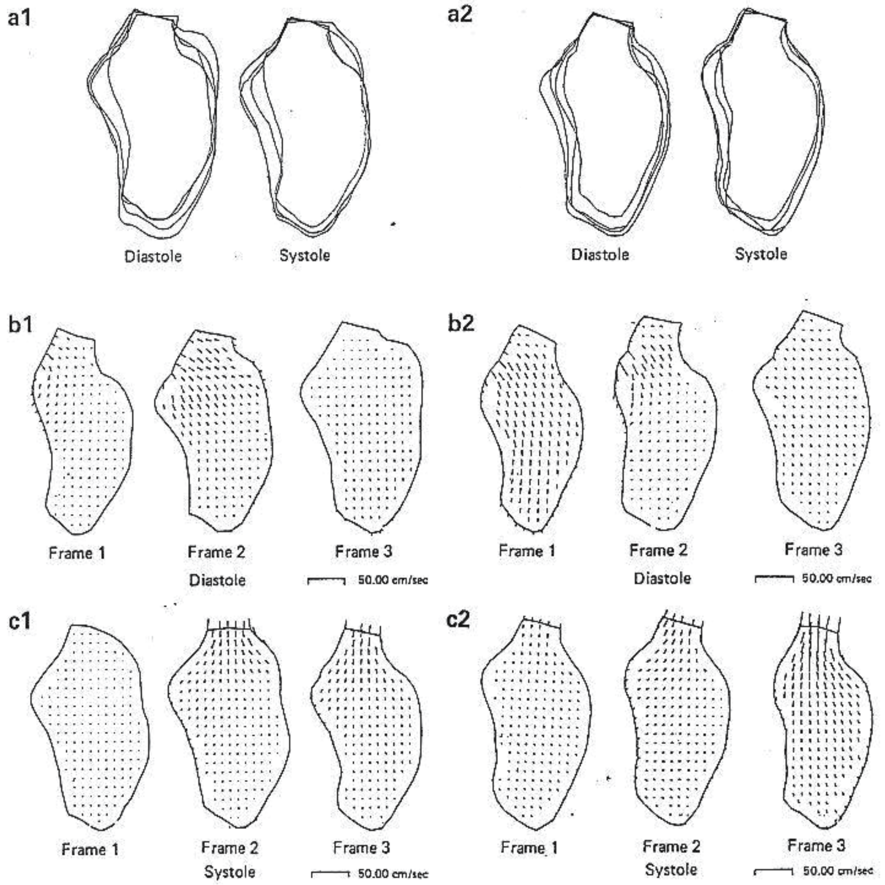

a1
Diastole
Systole
a2
Diastole
Systole
b1
Frame 1
Frame 2
Frame 3
Diastole
50.00 cm/sec
b2
Frame 1
Frame 2
Frame 3
Diastole
50.00 cm/sec
c1
Frame 1
Frame 2
Frame 3
Systole
50.00 cm/sec
c2
Frame 1
Frame 2
Frame 3
Systole
50.00 cm/sec

Fig. 11. (a,b,c) Patient TURN: (a) Superimposed sequential diastolic snd systolic endocardial frames (whose aortic valves centres and the long axis are matched) before (1) and after (2) administration of nitroglycerin. (b) Instantaneous intra-LV distributions of velocity during diastole, before (1) and after (2) administration of nitroglycerin. (c) Instantaneous LV distributions of velocity during ejection phase, before (1) and after (2) administration of nitroglycerin. (d) Instantaneous intra-LV distributions of pressure-differentials during diastole, before (1) and after (2) administration of nitroglycerin. (e) Instantaneous intra-LV distributions (pressure-differential) during ejection phase, before (1) and after (2) administration of nitroglycerin.

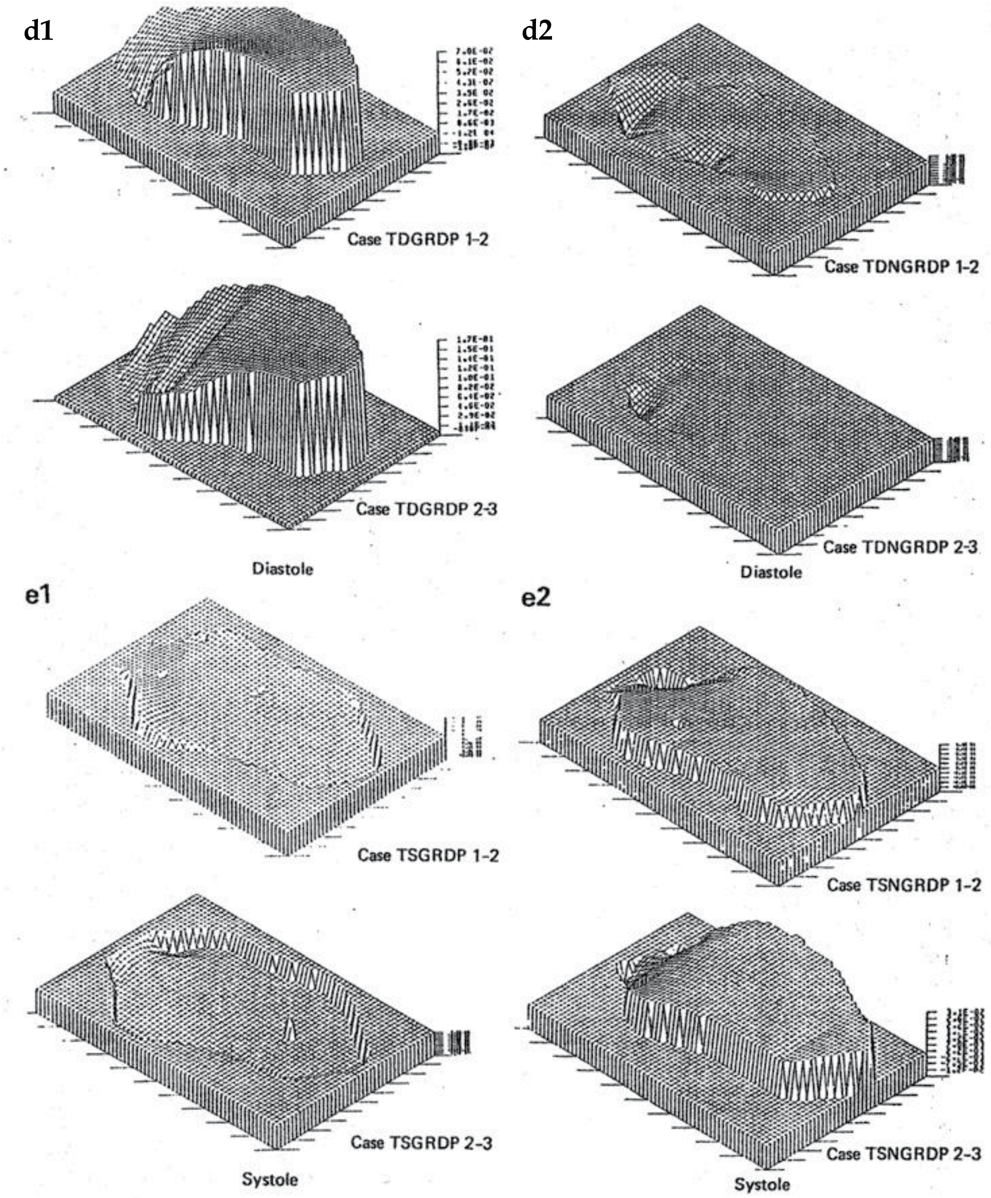


Fig. 11. (d,e) Patient TURN: (a) Superimposed sequential diastolic snd systolic endocardial frames(whose aortic valves centresand the long axis are matched) before (1) and after (2) administration of nitroglycerin. (b)Instantaneous intra-LV distributions of velocity during diastole, before (1) and after (2) administration of nitroglycerin. (c) Instantaneous LV distributions of velocity during ejection phase, before (1)and after (2) administration of nitroglycerin. (d) Instantaneous intra-LV distributions of pressure-differentials during diastole, before (1) and after (2) administration of nitroglycerin. (e) Instantaneous intra-LV distributions (pressure-differential) during ejection phase, before (1) and after (2) administration of nitroglycerin.

11. Biomedical engineering concept of Heart Failure

Heart Failure, in biomedical engineering (BME) terms, can imply failure of the heart to:

i. develop adequate contractility, due to a sizable amount of non-contractile infarcted myocardium (systolic heart failure)
ii. generate appropriate intra-LV pressure gradient during the ejection phase, to produce adequate LV outflow rate, stroke volume and cardiac output;
iii. produce adequate pressure increase during isovolumic contraction and ejection phases, to overcome the aortic systolic pressure;
iv. effect adequate stroke volume, due to poor contractility (systolic heart failure) or poor filling due to diseased and stiff myocardium (diastolic heart failure).

So the factors causing systolic heart failure may be summarized to be:

1. excessive percentage amount of infarcted myocardium *PMI* (as determined by the procedure in section 10),
2. resulting incapacity of the LV to produce appropriate intra-LV pressure gradient, for adequate LV outflow velocity and flow rate *dV/dt* (as determined by the methodology in section 10), as manifested by the contractility index *CCI* (formulated in section 2).

Now the resultant *dV/dt* factor in item 2 is incorporated in the formula for CCI (equation 3). No doubt, the *PMI* affects *CCI*, but there is no direct formulation connecting these two indices. So we can state that *PMI* and *CCI* are the two parameters that can be attributed to the occurrence of heart failure.

We can then define the Systolic Heart Failure index, as

$$HFIN = PMI\left(\text{in }\%\right)\ \text{x}HR\left(\text{in s}^{-1}\right) / CCI\left(\text{in s}^{-1}\right) \tag{46}$$

Now, in order to assess the terminal value of *HFIN*, we need to determine the terminal values of *PMI*, *CCI*, and *HR*, by studying normal subjects (as these indices are noninvasively determinable) as well as patients in different stages of heart failure.

Now, based on our studies (in Refs 4 & 5), let us (for the time being) adopt (i) the minimum acceptable value of *CCI* to be 3 s^{-1}, (ii) the maximum acceptable value of *PMI* to be 15 %, and (iii) the maximum resting *HR* to be 120/min or 2 s^{-1}. Substituting these values into equation (46) gives the terminal value of *HFIN* to be 10. In other words, if the value of *HFIN* exceeds the value 10, we can designate the patient to be in heart failure.

12. Concluding remarks

In this chapter, we have developed and presented noninvasive medical tests methodologies and associated NDPI(s), to make the case for reliable medical assessment of organ performance, physiological system function and dysfunction, anatomical structural property and pathology. The following tests and associated NDPI(s) have been presented:

3. Determination of cardiac contractility, from measurement of LV chamber volume and myocardial volume, in terms of the cardiac contractility index *CCI* of $(d\sigma^*/dt)_{max}$, given by equation (3).
4. Treadmill test to assess cardiac fitness and heart function, by means of the cardiac fitness index *CFI*, given by equation (8).

5. Lung Ventilation test, by monitoring lung volume by spirometry, for assessing lung ventilation and diagnosing lung disorders by means of lung ventilation index *LVPI-2* (equation 12).
6. Oral Glucose Tolerance test, to more reliably diagnose diabetes, by determining the diabetes index *DBI*, given by equation (16).
7. Noninvasive determination of arterial stiffness to detect arteriosclerosis, by ultrasound measurement of arterial dimensions and pulse wave velocity and auscultatory diastolic pressure, by means of the arteriosclerosis index *ART-NPI*, given by equation (20).
8. Noninvasive determination of (i) Aortic Pressure profile and (ii) Aortic normal vs disease state property in terms of the *Aortic number* given by equation (31), by monitoring the left ventricular outflow into the aorta and auscultatory diastolic and systolic pressures.
9. Noninvasive determination of Mitral valve pathology, by (i) monitoring its vibrational frequency from the first heart sound spectrum, and its size parameter from 2-d echocardiogram, and (ii) employing this data to structure its E^* *vs.* σ^* constitutive property (equations 36 and 37), and determine the alteration in the value of the constitutive index m given by equation (40).
10. Characterization of Osteoporosis, by determining the ulna bone vibratory resonance frequency, in terms of its flexural stiffness *EI*, given by equation (45).
11. Quantitative determination of (i) the amount of infarcted myocardial segment of the heart from echo-texture analysis of 2-d echo cardiograms (figure 10), and (ii) associated outcome of LV dysfunction in terms of the intra-LV blood-flow velocity and pressure distributions (figure 11).

Together these tests and their associated NDPI (s) can provide more reliable medical assessment. What now needs to be done is (i) application of these tests to large patient populations, and (ii) determination of the ranges of NDPI (s) for normal and abnormal states of organs, physiological systems and anatomical structures.

All of these tests can be employed in tertiary patient case, through the department of biomedical engineering (BME) in a tertiary-cave medical center. This makes a strong case for the institution of BME departments in tertiary case medical centers, which will revolution therapy health care.

13. References

[1] "Physiological Systems' Numbers in Medical Diagnosis and Hosipital Cost-effective Operation", by Dhanjoo N. Ghista, in *Journal of Mechanics in Medicine & Biology* 2005, vol 4, No.4.

[2] *Applied Biomedical Engineering Mechanics*, by Dhanjoo N. Ghista, CRC press (Taylor & Francis Group) Baton Rouge Florida 334872-2742, ISNBN 978-0-8247-5831-8,2008

[3] Measures and indices of intrinsic characterization of cardiac dysfunction during filling & systolic ejection," by Liang Zhong, Dhanjoo N. Ghista,Eddie Y-K. Ng, Lim SooTeik and Chua Siang Jin, in *Journal of Mechanics in Medicine & Biology* 2005, 5(2):37-322.

[4] "Validation of a novel noninvasive characterization of cardiac index of left ventricle contractility in patients", by Zhong L, Tan RS, Ghista DN, Ng E. Y-K, Chua LP, Kassab GS, *Am J Physiol Heart CircPhysiol*2007, 292:H2764-2772.

[5] "Effects of Surgical Ventricular Restoration on LV Contractility assessed by a novel Contractility index in patients with Ischemic cardiomyopathy", by Zhong L, Sola S, Tan RS, Le TT, Ghista DN, Kurra V, Navia JL, Kassab G.; in *Am J Cardiology*, 2009;103(5):674-679.

[6] Cardiac Fitness mathematical Model of Heart-rate response to V02 during and after Stress-Testing", Lim GeokHian, Dhanjoo N. Ghista, Koo TseYoong, John Tan Cher Chat, Philip EngTiew& Loo Chian Min; *International Journal of Computer Application in Technology(Biomedical Engineering & Computing Special Issue),* Vol 21, No 1/2, 2004.

[7] "Glucose Tolerance Test Modeling & Patient-Simulation for Diagnosis", by Sarma Dittakavi & Dhanjoo N. Ghista, *Journal of Mechanics in Medicine & Biology,* Vol. 1, No.2, Oct.2001.

[8] "Determination of the In-vivo Elasticity of Blood Vessel and Detection of Arterial Disease", by D. N. Ghista, *Automedica*, Vol. 1, No. 3, 1974.

[9] "Determination of Aortic Pressure-time Profile , Along with Aortic Stiffness and Peripheral Resistance", by Liang Zhong, Dhanjoo N. Ghista, Eddie Y-K. Ng, Lim SooTeik and Chua Siang Jin, in *Journal of Mechanics in Medicine & Biology* 2004, 4(4):499-509.

[10] "Mechanics of the Mitral Valve: Stresses in the Membrane, Indirect Determination of the Instantaneous Modulus of the Membrane", by D.N. Ghista; *Journal of Biomechanis,* Vol. 5, No. 3, 1972.

[11]"Mitral Valve Mechanics – Stress/Strain Characteristics of Excise Leaflets, Analysis of its Functional Mechanics and its Medical Applications", by D. N. Ghista and A. P. Rao: *Medical and Biological Engineering,* Vol. 11, No. 6, 1973.

[12] In Vivo Measurement of the Dynamic Response of Bone, by J.M. Jurist, H.D. Hoeks, D.A. Blackketter, R. K. Snyder and E.R. Gardner, in *Orthopaedic Mechanics: Procedures and Devices (Volume III),*ed by Dhanjoo N. Ghista and Robert Roaf, Academic Press (London), 1981.

[13] Noninvasive determination of Ulna stiffness from Mechanical Response---In Vivo comparison of stiffness and Bone mineral content in humans, by C.R. Steele, L.J. Zhou, D.Guido, R. Marcus, W. L. Heinrichs and C. Cheema, in *Journal of Biomechanical Engineering,* Vol. 10 (87), 1988.

[14] Preliminary Clinical results using SOBSA for noninvasive determination of ulna bending stiffness, by C.R. Steele and A.F. Gordon, in *Advances in Bioengineering (ASME),* edited by R. C Eberhand and A.H. Burstein, 1978, pp 85-87.

[15] Measures and Indices for Intrinsic Characterization of Cardiac Dysfunction during Filling and Systolic Ejection, by Liang Zhong, Dhanjoo N. Ghista, Eddie Y. Ng, Lim SooTeik, and Chua Siang Jin, in *Journal of Mechanics in Medicine and Biology*,Vol 5, No. 2, 2005.

[16] Detection of myocardial scars in neonatal infants form computerised echocardiographic texture analysis, by M.V Kamath, R.C Way, D.N. Ghista, T.M. Srinivasan, C. Wu, S Smeenk, C. Manning, J. Cannon, *Engineering in Medicine*, Vol. 15, Vo.3, 1986
[17] Intrinsic Indices of the Left Ventricle as a Blood Pump in Normal and Infarcted Left Ventricle, by K. Subbaraj, D.N. Ghista, and E. L. Fallen, in *J of Biomedical Engineering*, Vol 9, July issue, 1987

Edited by Dhanjoo N. Ghista

This innovative book integrates the disciplines of biomedical science, biomedical engineering, biotechnology, physiological engineering, and hospital management technology. Herein, Biomedical science covers topics on disease pathways, models and treatment mechanisms, and the roles of red palm oil and phytomedicinal plants in reducing HIV and diabetes complications by enhancing antioxidant activity. Biomedical engineering coves topics of biomaterials (biodegradable polymers and magnetic nanomaterials), coronary stents, contact lenses, modelling of flows through tubes of varying cross-section, heart rate variability analysis of diabetic neuropathy, and EEG analysis in brain function assessment. Biotechnology covers the topics of hydrophobic interaction chromatography, protein scaffolds engineering, liposomes for construction of vaccines, induced pluripotent stem cells to fix genetic diseases by regenerative approaches, polymeric drug conjugates for improving the efficacy of anticancer drugs, and genetic modification of animals for agricultural use. Physiological engineering deals with mathematical modelling of physiological (cardiac, lung ventilation, glucose regulation) systems and formulation of indices for medical assessment (such as cardiac contractility, lung disease status, and diabetes risk). Finally, Hospital management science and technology involves the application of both biomedical engineering and industrial engineering for cost-effective operation of a hospital.

Photo by Olivier Le Queinec / Shutterstock

IntechOpen

ISBN 978-953-51-4376-5