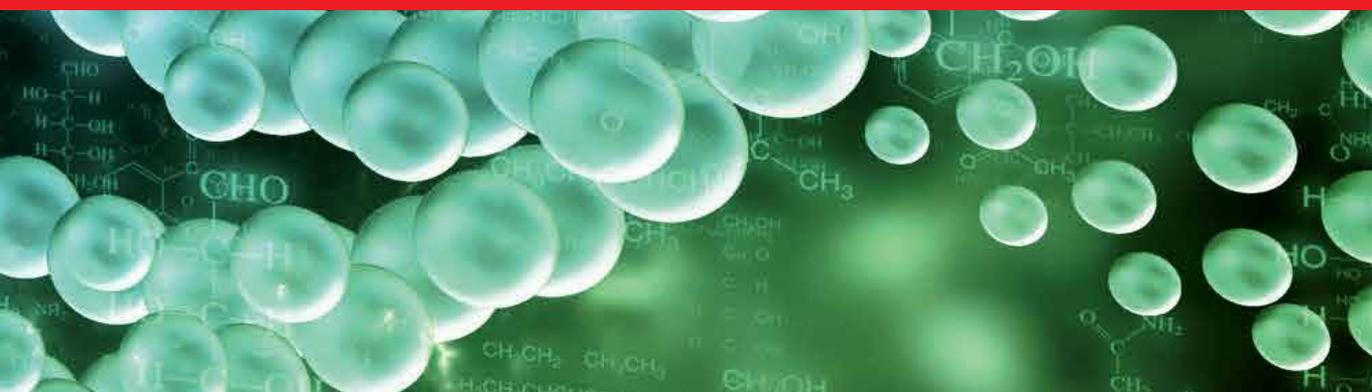




IntechOpen

# Fundamental Aspects of DNA Replication

*Edited by Jelena Kušić-Tišma*





---

# **FUNDAMENTAL ASPECTS OF DNA REPLICATION**

---

Edited by Jelena Kušić-Tišma

## Fundamental Aspects of DNA Replication

<http://dx.doi.org/10.5772/766>

Edited by Jelena Kušić-Tišma

### Contributors

Masaharu Takemura, Maria Fernandez, Adriana Fiorini, Fabrícia Gimenes, Quirino Alves De Lima Neto, Fábio Rogério Rosado, Ambrose Jong, Ann Kirchmaier, Jennifer Jacobi, Takeshi Mizuno, Masako Izumi, Christian Stefan Eichinger, Eisuke Gotoh, Takuro Nakagawa, Faria Zafar, Agustino Martínez-Antonio, Laura Espíndola-Serna, César Quiñones-Valles, Daniel Simmons, Jian-Hua Luo, Yan Yu, Chun Liang, Yuanliang Zhai, Philip Y.K. Yung, Jeanette Cook, Srikrupa Chandrasekaran, Karen Reidy, Nikolay Zenkin, Daniel Castro-Roa, Takeo Kubota, Kunio Miyake, Takae Hirasawa, Dianne C. Daniel, Edward Johnson, Malik Lutzmann

### © The Editor(s) and the Author(s) 2011

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](mailto: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 found at <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, 2011 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](mailto:orders@intechopen.com)

Fundamental Aspects of DNA Replication

Edited by Jelena Kušić-Tišma

p. cm.

ISBN 978-953-307-259-3

eBook (PDF) ISBN 978-953-51-5157-9

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,000+

Open access books available

116,000+

International authors and editors

120M+

Downloads

151

Countries delivered to

Our authors are among the  
Top 1%

most cited scientists

12.2%

Contributors from top 500 universities



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](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)





# Meet the editor

Dr. Jelena Kušić-Tišma is currently a research associate at the Institute of Molecular Genetics and Genetic Engineering, Belgrade. She received her PhD in Molecular Biology from Belgrade University in 2005. Her main research interest includes origin recognition mechanisms during initiation step of DNA replication in higher eukaryotes. She published several articles in peer review journals and is presently a visiting lecturer at the postgraduate studies of the Belgrade University.



---

# Contents

---

## **Preface XI**

- Chapter 1 **Assembly and Regulation of the Pre-Replication Complex: Increasing Complexity in Sight of Diversified Function and Regulation 3**  
Malik Lutzmann
- Chapter 2 **Regulation of DNA Replication Origin Licensing 13**  
Srikripa Chandrasekaran, Karen T. Reidy and Jeanette Gowen Cook
- Chapter 3 **Addressing the Enigma of MCM8 in DNA Replication 37**  
Dianne C. Daniel and Edward M. Johnson
- Chapter 4 **Regulation of MCM7 DNA Replication Licensing Activity 53**  
Jian-Hua Luo and Yan P. Yu
- Chapter 5 **Regulation of Minichromosome Maintenance (MCM) Helicase in Response to Replication Stress 65**  
Faria Zafar and Takuro Nakagawa
- Chapter 6 **Cell Cycle Control of DNA Replication by Phosphorylation and Dephosphorylation of Replication-Initiation Proteins in Budding Yeast 87**  
Yuanliang Zhai, Philip Y.K. Yung and Chun Liang
- Chapter 7 **Cdc6 Knockdown Renders p16<sup>INK4a</sup> Re-Activation, Leading to Senescence Human Breast Carcinoma Cells 107**  
Luo Feng, Jerry R. Barnhart, Lingtao Wu, Greg Shackelford, Sheng-he Huang and Ambrose Jong
- Chapter 8 **Visualize Dynamics of Chromosome Structure Formation and DNA Repair/Recombination Coupled With DNA Replication: Tight Coupled Role of DNA Replication in Chromosome Compaction and DNA Recombination 119**  
Eisuke Gotoh

- Chapter 9 **Sequence-Directed DNA Curvature in Replication Origins Segments 145**  
Adriana Fiorini, Fabrícia Gimenes, Quirino Alves de Lima Neto, Fábio Rogério Rosado and Maria Aparecida Fernandez
- Chapter 10 **Initiation of DNA Replication from Closed Circular DNA 161**  
Daniel Simmons
- Chapter 11 **Function of DNA Polymerase  $\alpha$  in a Replication Fork and its Putative Roles in Genomic Stability and Eukaryotic Evolution 187**  
Masaharu Takemura
- Chapter 12 **Quality Control of DNA Polymerase  $\alpha$  205**  
Takeshi Mizuno, Masako Izumi and Christian S. Eichinger
- Chapter 13 **Mechanisms and Controls of DNA Replication in Bacteria 219**  
César Quiñones-Valles, Laura Espíndola-Serna and Agustino Martínez-Antonio
- Chapter 14 **Propagating Epigenetic States During DNA Replication**  
Jennifer L. Jacobi and Ann L. Kirchmaier
- Chapter 15 **Epigenetic Modifications: Genetic Basis of Environmental Stress Response 271**  
Takeo Kubota, Kunio Miyake and Takae Hirasawa
- Chapter 16 **Relations Between Replication and Transcription 289**  
Daniel Castro-Roa and Nikolay Zenkin

---

## Preface

---

DNA molecule is the carrier of genetic information in the cell. Its faithful transmission through the generations is the basis of genome stability and integrity. A number of pathways ensure that the genome is replicated once and only once each time the cell divides. Several chapters in this book are dealing with regulation of DNA replication in eukaryotes. Targeted phosphorylation and dephosphorylation of key proteins that carry out initiation step of DNA replication is presented in chapter by Zhai et al., how initiator protein Cdc6 may promote cell proliferation by preventing senescence is demonstrated in chapter by Feng et al., while additional mechanisms, beside cyclin-CDK activity, adopted by cells to prevent re-replication during cell cycle arrest is addressed by Chandrasekaran and coworkers.

This book highlights different aspects of DNA replication process in eukaryotes, from assembly and regulation of prereplication complex (Lutzmann), mechanism of unwinding the double stranded DNA during initiation step (Simmons) and possible importance of intrinsically bent DNA regions in replication origins (Fiorini et al.), through structure and function of DNA polymerase  $\alpha$  (Takemura) and functional analysis of its individual subunits (Mizuno et al.), to formation and propagation of epigenetic patterns during the cell cycle (Jacobi et al.) and its changes by various environmental factors (Kubota).

On the other hand, minichromosome maintenance complex (MCM) is essential component of pre-replication complex and plays the central role in the progression of replication forks. Regulation of MCM complex under replication stress in order to preserve the replication machinery is reviewed in chapter by Zafar and Nakagawa. Possible role of MCM7 protein in human malignancies and its potential use in gene therapy is analyzed in chapter by Luo and Yan Yu, whereas possible function of another member of MCM protein family, MCM8 protein, in DNA replication and during the cell cycle is discussed in chapter by Daniel.

Interesting on its own right is the question of collision between replication fork progression and ongoing transcription discussed in chapter by Castro-Roa and Zenkin. Visualization of chromosome structure dynamics during S phase of the cell cycle, covered by Gotoh, has provided much of our understanding of relation between DNA replication and chromosome condensation.

Research of DNA replication, carried out in *E. coli*, revealed the basic mechanism and enzymology of replication fork and Espindola-Serna and coworkers in their chapter present an overview of DNA replication process and its regulation in prokaryotes.

The chapters in this book bring the current understanding of the mechanisms leading to genome stability and replication fidelity and we hope the reader will find it interesting, helpful and inspiring.

**Dr. Jelena Kušić-Tišma**

Laboratory for Molecular Biology  
Institute of Molecular Genetics and Genetic Engineering  
Belgrade  
Serbia

# Assembly and Regulation of the Pre-Replication Complex: Increasing Complexity in Sight of Diversified Function and Regulation

Malik Lutzmann  
*Institute of Human Genetics, IGH, CNRS  
France*

## 1. Introduction

Perhaps the two most fundamental rules in DNA replication say firstly to replicate the entire genome without leaving even small parts unreplicated and secondly, not to replicate any part, even a small one, more than just once during each cell cycle. In light of the size of an average eukaryotic genome and the time of S-phase, this is an enormous challenge. If in a gedankenexperiment a mammalian cell were scaled up to the size of a football and its DNA proportionally expanded in length, the DNA would span more than the distance from the earth to the sun - more than 1700 millions of kilometers in a football that has to be replicated in approximately eight hours.

In contrast to a (small) bacterial genome on a single chromosome, the replication of an eukaryotic genome (portioned in several chromosomes) has to start from multiple starting points, called origins of replication, in order to be accomplished in a reasonable time. Around 30.000 such origins are activated in an average mammalian cell and give rise to replication forks in each S-phase. In order to hold the fundamental rules of once per cell cycle replication, this complexity demands an extremely sophisticated regulation of replication initiation events. The formation of the pre-replication complex (pre-RC) on potential origins of replication at a time different from active replication is a key step in this regulation.

Through the formation of the pre-RC, the replicative helicase, the MCM2-7 complex, is loaded onto chromatin. The activity of this helicase separates the DNA double strand at the beginning and during S-phase and permits access of the polymerase machinery to the template strands. In addition, only the presence of the pre-RC proteins on chromatin permits the recruitment of other proteins that are necessary to load the polymerases. Therefore, in order to avoid that replication can start more than once per cell cycle at the same site, which would lead to re-replication, the formation of the pre-RC is restricted to a period with low S-phase kinase activity, that is, at the exit from mitosis and in G1. This system normally assures the once-per cell cycle replication: The license to start replication can only be given before replication itself has started. The cell only has to safeguard in addition that a pre-RC that once gave rise to a replication fork is by this action deactivated for the remaining cell cycle.

However, this very efficient system to avoid rereplication comprises on the other hand the danger that not sufficient replication forks will be created to replicate the entire genome if the cell encounters stress situations that block replication forks or lead to DNA breaks.

Therefore, far more pre-RCs are assembled as will normally be used later in S-phase to start a replication fork. This poses the questions how sites of pre-RC assembly are chosen and how the subpopulation of pre-RCs that will be activated and thus finally lead to the establishment of a replication fork are picked.

In this chapter, I will first review the basic protein machinery of the pre-RC and will then discuss additional factors that were recently characterized, which might influence or even decide the localization and fate of a given pre-RC. I will then give a short insight into situations, where re-replication is physiologically normal and give a summary of how origin licensing is regulated in these special cell cycles. At the end of this chapter, I will present examples that relate the increasing complexity of pre-RC factors in higher eukaryotes with their more and more complex organization of S-phase with respect to cell fate and differentiation.

## 2. Assembly and function of the pre-RC

The yeast *Sacharomyces cerevisiae* is special as it assembles its pre-RCs on sites on the chromatin that are defined by sequence motifs, so called *ARS* (**A**utonomous **R**eplication **S**equences) sites. Nevertheless, the yeast pre-RC serves as a paradigm, since its basic machinery is conserved throughout eukaryotes.

Pre-RC assembly begins in G1 with the binding of the hexameric ORC-complex (**O**rgan **R**ecognition **C**omplex, ORC1 - ORC6) to the *ARS* consensus sequences. The ORC-subunits 1,4 and 5 are ATPases. Once the ORC-complex is chromatin-bound, CDC6, another ATPase, joins and recruits the Cdt1 protein. Cdt1 has no enzymatic activity, but is nevertheless also an essential factor for pre-RC assembly. Importantly, Cdt1 seems to be the most regulated factor of the pre-RC, since it is tightly regulated by chromatin binding, several degradation pathways and finally by binding to geminin (see below).

Pre-RC assembly is then finally accomplished by the loading of the replicative helicase, the MCM2-7 complex, onto chromatin. In *S. cerevisiae*, Cdt1 is already in its soluble, non-chromatin bound form associated with the MCM2-7 complex, whereas in other organisms, Cdt1 and the MCM2-7 complex bind one after the other to the forming pre-RC. The concerted ATP consuming action of ORC, CDC6 and Cdt1 loads multiple hexameric MCM2-7 complexes, 20 to 40, for each pre-RC, on DNA. A huge amount of work and effort was invested to understand this ATP-driven MCM2-7 loading machinery functionally and (so far partially) also structurally (Mizushima et al, 2000), (Harvey & Newport, 2003), (Bowers et al, 2004), (Speck et al, 2005), (Randell et al, 2006).

ATP hydrolysis by ORC and CDC6 is necessary for iterative MCM2-7 loading and thus for viability. However, ORC bound to ATP, but defective of ATP hydrolysis, can support one single round of MCM2-7 loading, whereas the ATPase activity of CDC6 is strictly required for any MCM2-7 loading (Bowers et al, 2004). The "loaded" chromatin binding state of the MCM2-7 complex is defined as a salt-resistant binding that is, once loading is accomplished, even independent of the presence of the other pre-RC proteins.

Speck et al. added a milestone by presenting the first structural information of the "MCM2-7 loading machine" (Speck et al, 2005). The authors presented an image reconstruction from electron micrographs of the yeast ORC complex alone and in complex with CDC6. These structures with a resolution of about 25 Å show the ORC complex as an elongated entity composed of three domains. Association with CDC6 creates a ring-like structure that has quite similar dimensions to an archeal MCM helicase ring that could be visualized by protein crystallography.

Another fundamental step forward in understanding pre-RC formation functionally and structurally was achieved in 2009 when Evrien et al. and Remus et al. managed independently to reconstitute yeast pre-RC assembly *in vitro* from purified proteins onto immobilized DNA and to follow and dissect this assembly functionally and structurally (Remus et al, 2009), (Evrin et al, 2009). Importantly, they could present for the first time structural data showing how the MCM2-7 complex is topologically "loaded" onto double stranded DNA. The ORC-CDC6 assembly on chromatin recruits Cdt1-MCM2-7 heteroheptamers and loads MCM2-7 complexes stably onto DNA in a way that each time two of them form stable head-to head double hexamers that can freely slide along the DNA. They also presented data arguing that these double hexamers are loaded in a way that the DNA double strand passes through the central channel of the helicase, at least during its inactive state.

### **3. Regulation of the pre-RC in different organisms: Assembly and activation**

As mentioned, the yeast pre-RC is a simple, but without doubt the best understood "MCM2-7 loading machine" known, but as organisms became more and more complex in respect of their development and cellular differentiation, additional factors seemed to be necessary to provide either back-up mechanisms of regulation or the possibility of fine-tuning pre-RC assembly and activation in different cellular contexts.

In this section, I will present these additional factors. Some of them have been studied for more than ten years, others were only very recently discovered or described and their functions still await confirmation and a coherent embedding in the known context of origin licensing and cell identity.

#### **Geminin:**

Geminin was discovered 1998 in a screen for proteins being degraded by the anaphase promoting complex (APC) as a protein that blocks DNA replication in *Xenopus laevis* egg extracts (McGarry & Kirschner, 1998). Soon after the discovery of Cdt1 (Maiorano et al, 2000), it became clear that geminin binds to Cdt1 and inhibits Cdt1 in recruiting the MCM2-7 complex to the assembling pre-RC (Tada et al, 2001).

Whereas in yeast, Cdt1's activity is regulated solely by its degradation at S-phase entry, in higher eukaryotes geminin provides thus another layer of MCM2-7 recruitment control. Geminin is highly expressed at S-phase onset and blocks a second round of pre-RC assembly by binding and inhibiting Cdt1. Not before the end of mitosis, when the APC targets geminin for destruction, a new round of pre-RC assembly is thus permitted. Geminin is absent from the genomes of yeast and fungi and *C. elegans* might have a diverged form that however still awaits comprehensive functional analysis (Yanagi et al, 2005).

*Drosophila* has a geminin that shares the functions of controlling and blocking DNA replication by direct interaction with Cdt1 (Quinn et al, 2001). Plants also encode a protein that shares partial homology with geminin that functions in cell differentiation and that also interacts with plant Cdt1, however, a clear role in replication control by blocking Cdt1 needs to be shown (Caro & Gutierrez, 2007).

Importantly, geminin is also implicated in development and cell differentiation, as it interacts with key transcription factors, chromatin remodelers and seems further to be able to act as a transcription factor for developmental genes on its own (see down). An interesting model arises with these findings stating that geminin is placed at the intersection

between the "replication road" and the "differentiation road" and its presence and binding to different partners influences which way a cell goes: Less geminin (or no geminin at a particular localization at a particular moment in the cell cycle) leads the cell to assemble pre-RCs and to license origins and so to engage in another cell cycle. More geminin blocks replication licensing and thus proliferation, but at the same time activates the transcription of genes that lead to cell differentiation like neuronal transcription factors (see below). However, these regulating functions of geminin are surely more complex than known so far and we await in the future more data about how precisely in a given situation geminin influences cell fates.

In the earliest and extremely fast embryonic cell cycles, before transcription is activated, geminin is not (completely) degraded at mitosis, and is present also in G1 during pre-RC formation. Several studies investigated how pre-RC formation in the presence of geminin in these early cell cycles could be possible. Hodgson et al. (Hodgson et al, 2002) reported that the geminin that survived mitosis does not bind to Cdt1 and seems to be modified (as Li et al. suggested later, possibly by ubiquitination (Li & Blow, 2004)) in a way that its interaction with Cdt1 is blocked, thus Cdt1 is kept active in recruiting MCM2-7 complexes to chromatin. After licensing is achieved, the nuclear envelope reforms around chromatin, and only after this step is completed, S-phase promoting Cdks can be concentrated inside the nucleus in order to trigger S-phase. Then, once the nuclear envelope is resealed, also geminin becomes imported into the nucleus and seems to lose its modification by the import process itself so that it is again competent of binding and inhibiting Cdt1.

Another study reports the existence of two different stoichiometries of the Cdt1-geminin complex, one containing less geminin and a second one containing more geminin bound to each Cdt1 molecule (Lutzmann et al, 2006). The two different complexes can be assembled *in vitro*, and by using the *Xenopus* egg extract system it was shown that both complexes indeed behave differently: Cdt1 with less geminin bound is active in recruiting the MCM2-7 complex, but licenses origins in a strictly controlled "once per cell cycle" manner (under these conditions, Cdt1 alone leads to re-replication) whereas the complex of Cdt1 with more geminin associated is licensing inactive. This change in the stoichiometry seems to take place exactly at the onset of S-phase. These models of how licensing is permitted in the presence of geminin are not mutually exclusive. Several structural studies of Cdt1- and geminin-domains, alone and in complex, support the idea of different stoichiometries of the Cdt1-geminin complex (Lee et al, 2004), (Okorokov et al, 2004), (De Marco et al, 2009).

#### **MCM8 and MCM9:**

Besides the heterohexameric MCM2-7 complex, the family of MCM-domain containing DNA helicases also includes MCM8 and MCM9. In contrast to the MCM2-7 complex, MCM8 and MCM9 do not exist in all eukaryotic genomes, but with the only exception of *Drosophila melanogaster*, both are either together present or absent in a given genome, suggesting that the two proteins have interdependent roles (Liu et al, 2009). Yeast and *C.elegans* do not have MCM8 and MCM9, but several fungi possess open reading frames coding for shortened versions of MCM8 and MCM9. However, a long version of MCM9 with a C-terminal extension of about 500 amino acids is found only in vertebrates (Lutzmann et al, 2005).

Whereas one report on MCM8 employing overexpression of the protein places MCM8 on the pre-RC (Volkening & Hoffmann, 2005), another report depleting the endogenous protein using *Xenopus* egg extracts shows that pre-RC formation and the initiation of DNA

replication is not affected by the absence of MCM8, but that elongation of replication forks is hindered after depletion of MCM8 (Maiorano et al, 2005).

Recently, single nucleotide polymorphisms (SNPs) in human MCM8 exons were shown to affect the fertile life span in women. However, the biological mechanism of these results is at present unknown (He et al, 2009), (Stolk et al, 2009).

Using *Xenopus* egg extracts, which recapitulate the early embryonic cell cycles during embryonic development, it was shown that MCM9 is necessary for pre-RC assembly by binding to Cdt1 and so limiting the amount of geminin that can associate to Cdt1 during the period where Cdt1 has to be active in recruiting the MCM2-7 complex in the presence of geminin (Lutzmann & Mechali, 2008). However, whether the functions of MCM9 (and MCM8) are similar in these early embryonic cell cycles and in normal somatic cycles has still to be investigated.

### **HBO1:**

A highly active field in investigating the regulation of DNA replication origin analyses the chromatin environment at origins. The identification of the HBO1 protein, a MYST-family histone acetyltransferase that is necessary for normal origin licensing and MCM2-7 recruitment was an important step in deciphering chromatin remodeling during the licensing reaction. HBO1 (**H**uman acetylase **B**inding to **ORC1**) was originally identified as an ORC-interacting acetylase in a yeast two hybrid-screen (Iizuka & Stillman, 1999). By affinity purification, HBO1 is also found in a large chromatin-remodeling complex containing ING5 and this complex was shown to be essential for normal replication and S-phase progression (Doyon et al, 2006).

In such different systems as cultured cells from humans, *Drosophila* or *Xenopus* egg extracts, the depletion or inactivation of HBO1 affects the chromatin recruitment of the MCM2-7 complex and leads to replication defects. However, the effects on pre-RC formation vary. For example, whereas in *Xenopus* egg extract, depletion of the putative *Xenopus* HBO1 leads not only to a defect in MCM2-7 loading and replication but also to a recruitment defect of Cdt1 (Iizuka et al, 2006), knock-down of HBO1 in human cells does not affect the chromatin association of Cdt1 (Miotto & Struhl, 2008). HBO1 is found in its soluble, non-chromatin bound form in complex with Cdt1 and seems to enhance Cdt1's MCM2-7 recruiting activity once bound to chromatin.

Using ChIP, it was shown that HBO1 bound to Cdt1 is specifically recruited to origins, and acetylates at these sites histone H4 tails in order to facilitate MCM2-7 loading. Interestingly, it was reported further that geminin binding to Cdt1 does not physically affect the association of HBO1 with Cdt1 or its targeting to origins, but inhibits HBO1's acetylase activity specifically toward H4 at origins (Miotto & Struhl, 2010). Consistent with this, HBO1 is absent in organisms that also do not possess geminin, including yeast and possibly worms.

### **HOX proteins:**

Several studies reported that geminin binds not only to the promoter regions of several HOX genes, but also directly to a subset of HOX proteins. One study investigated the binding of HOX proteins to replication origins and reported binding of several HOX-members, especially HOXD13 to origin DNA (Salsi et al, 2009). HOXD13 was found in this study to interact further with CDC6, in order to enhance the recruitment of pre-RC proteins to origins. When overexpressed, HOXD13 seems to accelerate DNA replication. The authors of this study claim further that geminin binding to HOXD13 blocks the interaction with CDC6 and the positive effect of HOXD13 to origin licensing.

**ORCA:**

Up to now, the most upstream event of pre-RC formation known is the recruitment of the ORC complex to chromatin. Since besides yeast, no clear consensus site of ORC binding could be found on DNA, it remains unclear which factors specify ORC-binding sites. Chromatin context, influenced by multiple factors such as histone modifications, transcription and the presence of other chromatin binding proteins seem to play crucial roles.

Very recently, Vermeulen et al. (Vermeulen et al, 2010) and Shen et al. (Shen et al, 2010) identified in human cells a protein that was called ORCA (ORC Associated) that stabilizes ORC binding to chromatin.

ORCA was identified by mass spectrometric analysis of ORC2-immunoprecipitates from human cells and homologous proteins were identified in higher eukaryotes from human to flies, but the protein seems to be absent in yeast and worms. Interestingly, when the reciprocal ORCA-IP was analyzed, not only ORC subunits were found to interact with ORCA, but also Cdt1 and geminin (Shen et al, 2010). ORCA has about 650 amino acids and contains leucin-rich repeats and a WD-repeat containing domain. It is this WD-repeat containing domain that confers both chromatin association of ORCA and also binding to ORC. ORCA was shown to associate predominantly with chromatin during the G1 phase and to colocalize partially with HP1, mainly at heterochromatic regions.

Knockdown of ORCA leads in all cell lines analyzed to a reduction of ORC binding to chromatin. Interestingly, whereas cancer cell lines do not show a concomitant reduction in MCM2-7 recruitment or cell cycle alterations, primary cells indeed recruit less MCM2-7 and are enriched in G1 after knock-down of ORCA.

Also remarkable is the different colocalization of ORCA and ORC on centromeres and telomeres in several different immortalized cell lines. Telomerase-positive cells show ORCA and ORC localization at the centromeres whereas cells that use alternative lengthening of telomeres (ALT) through recombination recruit ORCA and ORC to telomeres. These findings indicate that ORCA and ORC might also have a function in recombination and / or DNA repair.

**GEMC1:**

GEMC1 was identified by database mining for ORFs containing sequence motifs of known replication proteins (Balestrini et al, 2010). The protein was named GEMC1 since its sequence contains a central coiled-coil region that shares similarities with geminin's coiled-coil region (**Geminin coiled-coil Containing protein 1**). The presence of GEMC1 was reported so far only in the genomes of vertebrates. However, despite its partial homology to geminin, GEMC1 does not exert a blocking function in DNA replication, but an activating one.

GEMC1 is necessary for the recruitment of CDC45, an essential factor that plays a key initiating role in turning the pre-RC into the pre-initiation complex (pre-IC) at S-phase onset. CDC45 binds to chromatin downstream of pre-RC formation, when S-phase kinase activity is already high, but upstream of the recruitment of replisome-factors such as MCM10, GINS and polymerase  $\alpha$ . GemC1 is highly phosphorylated by S-phase kinase activities and is recruited to chromatin through its association with the TopBP1 protein. This association is strengthened by GEMC1 phosphorylation. In *Xenopus* egg extracts, GemC1 and TopBP1 however bind to chromatin already during pre-RC formation. Whereas Balestrini et al. present data showing that binding of GEMC1 is indeed ORC-dependent, it is still debated whether the chromatin association of TopBP1 is equally ORC-dependent.

However, the question of whether GEMC1 is a true pre-RC protein is not clearly answered by this fact since its absence (by depletion from *Xenopus* egg extract) does not inhibit the final step of pre-RC assembly, the recruitment of the MCM2-7 helicase. Thus, pre-RC formation and origin licensing are not dependent on GEMC1.

Maybe, somehow similar to the ORCA protein, GEMC1 might be a factor that does not fit entirely into the textbook model of pre-RC assembly, were the chromatin-binding of one factor is strictly and entirely dependent on the presence of its upstream predecessor on chromatin. Once ORC and / or ORCA are bound to chromatin at a potential origin of replication, several, and from there on independent, branches of protein-machinery assembly might take place. One is surely the loading of the replicative helicase, the MCM2-7 complex, and another might be the recruitment of factors that permit the activation of this helicase later at S-phase onset. If helicase loading is blocked, at least a part of the machinery that would be needed further downstream of MCM2-7 loading to activate an origin might still be assembled: GemC1 binding to chromatin is dependent on ORC, but GemC1 is not necessary to load the MCM2-7 helicase, and if MCM2-7 chromatin loading is blocked (by addition of exogenous geminin), GEMC1 binds to chromatin nevertheless.

GEMC1 might be an interesting factor that forces us to question and maybe re-think the formerly-presumed strict dependences and hierarchies of licensing and replisome assembly.

#### 4. Controlled re-replication

Before I finish the chapter by trying to interpret the increasing number of factors that are required for either pre-RC assembly or its regulation in multicellular organisms that have a more and more complex embryonic development and cell differentiation program, I want to mention situations in which controlled re-replication is required and physiological normal.

Whereas during aberrant re-replication cells re-replicate parts of the genome in an uncontrolled number, in most physiological situations of re-replication, either the whole genome (in so called *endocycles*) or only very particular loci (so called *gene amplifications*) are precisely re-replicated multiple times. Controlled re-replication is normally programmed in cells that have a very high metabolic turnover or have to produce a certain (sub-) set of proteins in extremely large amounts. Not surprisingly, in most of the known cases of gene amplification, ribosomal genes are amplified. Another prominent example is the amplification of chorion genes in *Drosophila* ovary cells that encode the proteins forming the eggshell.

Endocycles, in which cells re-replicate their entire genome (albeit often less accurately than in normal cycles) are executed when cells pass directly from a G2 phase into a new G1 phase (without passing through mitosis) or if only a partial mitosis that lacks cytokinesis is interposed. Endocycles are common in plants and flies and also in some vertebrates (most often in fish and reptiles) but are rare in mammals. Very generally, endocycles require that Cdk1 activity is inhibited in order to block mitosis and that the APC(Cdh1) activity is high to degrade mitotic cyclins and - in higher eukaryotes - also geminin at moments when Cdk activity is low.

In addition, another important change in cell cycle control mechanisms has to take place: Normally, in order to block the propagation of damaged DNA or of extra copies of DNA to daughter cells, the DNA damage checkpoint becomes activated when re-replication occurs and this response activates either cell senescence or cellular death via apoptosis. Therefore, another distinctive mark of cells undergoing scheduled endoreplication is the inactivation of

the DNA damage checkpoint during re-replication of the genome: Even if artificially (by overexpression of the *Drosophila* Cdt1 homologue Dup1) in endocycling *Drosophila* cells DNA damage is caused, the DNA damage checkpoint is not activated in these cell cycles (Mehrotra et al, 2008).

The best described situation of mammalian cells that undergo endocycles is found during the differentiation of trophoblast stem cells into trophoblast giant cells that are produced to facilitate the implantation of the early embryo into the epithelium of the uterus. Importantly, the licensing machinery is essential and active also in endocycling cells. Therefore, like in normal mitotic cycles, gap phases are required where Cdk2 activity is low and pre-RC assembly is allowed to take place.

## 5. Origin choice and pre-RC complexity

In this final part of the chapter, I will resume the ideas in the field why multicellular eukaryotes with a complex developmental biology and cell differentiation program obviously need more factors to regulate replication origins as do single-cell eukaryotes. As mentioned in the beginning, the yeast *S. cerevisiae* possesses a simple licensing apparatus, maybe the most *sized-down* of all eukaryotes. It is surely not a coincidence that this organism specifies all possible replication origins through clearly defined DNA sequence motifs, the autonomous replication sequences (ARS), which work when placed on plasmids outside the chromosomal context as well as in their natural genomic environment (even though if not all ARS sequences in the yeast genome lead to active replication origins in each cell cycle). Furthermore, besides a meiotic cell cycle, yeast does not have a developmental- or cell differentiation program.

For multicellular organisms (and for most unicellular ones as well), no consensus sequence or sequences could be found that specify pre-RC assembly sites or origins of replication. In contrast, the positions of assembled and finally also used replication origins can change due to gene expression, growth conditions and chromatin structure (for review see (Mechali, 2010)).

It seems that during evolution, epigenetic factors took at least a part of origin determination from the pure genetic influence (the sequence) in order to cope with the replication of an always similar sequence that is present in very different chromatin states depending on cell identity and transcriptional programs.

As pointed out, in higher eukaryotes a number of pre-RC factors interact with chromatin-shaping proteins and chromatin remodelers. The ORC complex interacts with HP1 and seems to be dependent on ORCA for its normal localization, especially at heterochromatic sites. ORC and MCM2 further interact with the histone acetylase HBO1. Cdt1 is dependent on HBO1 and its histone acetylation activity to exert its function in recruiting the MCM2-7 complex.

Furthermore, geminin, likely not an essential pre-RC factor itself, but essential for its proper regulation, has complex roles in origin regulation. The protein interacts not only with crucial developmental transcription factors but also with the chromatin remodeler Brg1 (Seo et al, 2005). Thus, several chromatin-remodeling activities seem to be required to assemble and regulate pre-RCs.

It was shown that origin positioning changes with the transcriptional status of a chromatin domain (Gregoire et al, 2006) and further that binding of a transcription factor to a plasmid specifies the site where replication of the plasmid starts when incubated in *Xenopus* egg extract that otherwise does not show sequence specificity of DNA templates (Danis et al, 2004). Interestingly, the moment during G1 when origins are chosen that will be activated later in S-phase is precisely defined and called the **Origin Decision Point (ODP)** (Wu & Gilbert,

1996), (Li et al, 2003). Once this program is established, cells have to go through mitosis in order to "reset" this programming. Incubation of differentiated nuclei (having large origin distances) in mitotic egg extract (resembling the undifferentiated state with short inter-origin distances) is sufficient to switch from large to small replicons (Lemaitre et al, 2005).

Mostly, cell differentiation has to go along with a block in proliferation. In organisms composed of billions of cells, which are found at all possible positions between the extremes of completely undifferentiated, fast proliferating stem cells and on the other side fully differentiated cells that dropped completely out of the cell cycle, the correct tuning between differentiation and proliferation is a complex and vital task. Several data suggest that also here, the pre-RC and its regulation play an important role.

Since its discovery, two roles of geminin have been described: Its regulating (blocking) function on the licensing system (McGarry & Kirschner, 1998) and its involvement in neuronal differentiation (Kroll et al, 1998). Several situations were described where geminin is placed on the crossroads between proliferation and differentiation.

Geminin binds the homeobox-containing transcription factor Six3 that is necessary for proliferation of retinal precursor cells. Six3 can displace Cdt1 from geminin. By this means, even without affecting transcription, Six3 can specifically activate the proliferation of retinal precursor cells by sequestering geminin from Cdt1 (Del Bene et al, 2004).

Geminin also interacts with HOX proteins and blocks their binding to their promoter targets. In addition, it also binds to polycomb group proteins that repress HOX transcription and function. Like in the case of Six3, HOX proteins seem to be able to displace geminin from Cdt1. Therefore, a model arises in that geminin can block differentiation by binding and blocking HOX proteins and at the same time releasing Cdt1 to favor proliferation. Or, in contrast, geminin can release HOX proteins and block Cdt1, which leads the cell to take the road away from proliferation towards differentiation (Luo et al, 2004).

In yet another study, geminin was shown to bind the catalytic subunit of the SWI/SNF chromatin-remodeling complex, Brg1 (Seo et al, 2005). This complex is necessary for the transcription of neuronal genes that lead to neuronal differentiation in neuronal progenitor tissue. Thus, in this case, more geminin means a block of Brg1 activity, no transcription of neuronal genes and ongoing proliferation of neuronal progenitor tissue. Less geminin keeps Brg1 active, neuronal genes are transcribed and differentiation wins over proliferation, even if less geminin (which means more free Cdt1) is present.

Finally, in *Xenopus* egg extracts recapitulating the very early embryonic cell cycles in which geminin is not degraded but present in G1, the MCM9 protein limits the binding of geminin to Cdt1, keeping Cdt1 active in recruiting the MCM2-7 complex to chromatin (Lutzmann & Mechali, 2008).

Clearly, most of the principal, and obviously all of the fine-tuning mechanisms that use the licensing reaction to balance proliferation and differentiation have still to be discovered and characterized. Likely, in the coming years we will learn much more about how replication and its regulation are connected to and interdependent with cell differentiation and development. Pre-RC assembly and origin choice might play crucial roles in these processes.

## 6. References

- Balestrini A, Cosentino C, Errico A, Garner E, Costanzo V (2010) GEMC1 is a TopBP1-interacting protein required for chromosomal DNA replication. *Nat Cell Biol* 12: 484-491

- Bowers JL, Randell JC, Chen S, Bell SP (2004) ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. *Mol Cell* 16: 967-978
- Caro E, Gutierrez C (2007) A green GEM: intriguing analogies with animal geminin. *Trends Cell Biol* 17: 580-585
- Danis E, Brodolin K, Menut S, Maiorano D, Girard-Reydet C, Mechali M (2004) Specification of a DNA replication origin by a transcription complex. *Nat Cell Biol* 6: 721-730
- De Marco V, Gillespie PJ, Li A, Karantzelis N, Christodoulou E, Klompmaker R, van Gerwen S, Fish A, Petoukhov MV, Iliou MS, Lygerou Z, Medema RH, Blow JJ, Svergun DI, Taraviras S, Perrakis A (2009) Quaternary structure of the human Cdt1-Geminin complex regulates DNA replication licensing. *Proc Natl Acad Sci U S A* 106: 19807-19812
- Del Bene F, Tessmar-Raible K, Wittbrodt J (2004) Direct interaction of geminin and Six3 in eye development. *Nature* 427: 745-749
- Doyon Y, Cayrou C, Ullah M, Landry AJ, Cote V, Selleck W, Lane WS, Tan S, Yang XJ, Cote J (2006) ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Mol Cell* 21: 51-64
- Evrin C, Clarke P, Zech J, Lurz R, Sun J, Uhle S, Li H, Stillman B, Speck C (2009) A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc Natl Acad Sci U S A* 106: 20240-20245
- Gregoire D, Brodolin K, Mechali M (2006) HoxB domain induction silences DNA replication origins in the locus and specifies a single origin at its boundary. *EMBO Rep* 7: 812-816
- Harvey KJ, Newport J (2003) Metazoan origin selection: origin recognition complex chromatin binding is regulated by CDC6 recruitment and ATP hydrolysis. *J Biol Chem* 278: 48524-48528
- He C, Kraft P, Chen C, Buring JE, Pare G, Hankinson SE, Chanock SJ, Ridker PM, Hunter DJ, Chasman DI (2009) Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. *Nat Genet* 41: 724-728
- Hodgson B, Li A, Tada S, Blow JJ (2002) Geminin becomes activated as an inhibitor of Cdt1/RLF-B following nuclear import. *Curr Biol* 12: 678-683
- Iizuka M, Matsui T, Takisawa H, Smith MM (2006) Regulation of replication licensing by acetyltransferase Hbo1. *Mol Cell Biol* 26: 1098-1108
- Iizuka M, Stillman B (1999) Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein. *J Biol Chem* 274: 23027-23034
- Kroll KL, Salic AN, Evans LM, Kirschner MW (1998) Geminin, a neuralizing molecule that demarcates the future neural plate at the onset of gastrulation. *Development* 125: 3247-3258
- Lee C, Hong B, Choi JM, Kim Y, Watanabe S, Ishimi Y, Enomoto T, Tada S, Cho Y (2004) Structural basis for inhibition of the replication licensing factor Cdt1 by geminin. *Nature* 430: 913-917
- Lemaitre JM, Danis E, Pasero P, Vassetzky Y, Mechali M (2005) Mitotic remodeling of the replicon and chromosome structure. *Cell* 123: 787-801
- Li A, Blow JJ (2004) Non-proteolytic inactivation of geminin requires CDK-dependent ubiquitination. *Nat Cell Biol* 6: 260-267
- Li F, Chen J, Solessio E, Gilbert DM (2003) Spatial distribution and specification of mammalian replication origins during G1 phase. *J Cell Biol* 161: 257-266

- Liu Y, Richards TA, Aves SJ (2009) Ancient diversification of eukaryotic MCM DNA replication proteins. *BMC Evol Biol* 9: 60
- Luo L, Yang X, Takihara Y, Knoetgen H, Kessel M (2004) The cell-cycle regulator geminin inhibits Hox function through direct and polycomb-mediated interactions. *Nature* 427: 749-753
- Lutzmann M, Maiorano D, Mechali M (2005) Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2-8 protein family. *Gene* 362: 51-56
- Lutzmann M, Maiorano D, Mechali M (2006) A Cdt1-geminin complex licenses chromatin for DNA replication and prevents rereplication during S phase in *Xenopus*. *EMBO J* 25: 5764-5774
- Lutzmann M, Mechali M (2008) MCM9 binds Cdt1 and is required for the assembly of prereplication complexes. *Mol Cell* 31: 190-200
- Maiorano D, Cuvier O, Danis E, Mechali M (2005) MCM8 is an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. *Cell* 120: 315-328
- Maiorano D, Moreau J, Mechali M (2000) XCDT1 is required for the assembly of pre-replicative complexes in *Xenopus laevis*. *Nature* 404: 622-625
- McGarry TJ, Kirschner MW (1998) Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell* 93: 1043-1053
- Mechali M (2010) Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat Rev Mol Cell Biol* 11: 728-738
- Mehrotra S, Maqbool SB, Kolpakas A, Murnen K, Calvi BR (2008) Endocycling cells do not apoptose in response to DNA rereplication genotoxic stress. *Genes Dev* 22: 3158-3171
- Miotto B, Struhl K (2008) HBO1 histone acetylase is a coactivator of the replication licensing factor Cdt1. *Genes Dev* 22: 2633-2638
- Miotto B, Struhl K (2010) HBO1 histone acetylase activity is essential for DNA replication licensing and inhibited by Geminin. *Mol Cell* 37: 57-66
- Mizushima T, Takahashi N, Stillman B (2000) Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro. *Genes Dev* 14: 1631-1641
- Okorokov AL, Orlova EV, Kingsbury SR, Bagneris C, Gohlke U, Williams GH, Stoerber K (2004) Molecular structure of human geminin. *Nat Struct Mol Biol* 11: 1021-1022
- Quinn LM, Herr A, McGarry TJ, Richardson H (2001) The *Drosophila* Geminin homolog: roles for Geminin in limiting DNA replication, in anaphase and in neurogenesis. *Genes Dev* 15: 2741-2754
- Randell JC, Bowers JL, Rodriguez HK, Bell SP (2006) Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Mol Cell* 21: 29-39
- Remus D, Beuron F, Tolun G, Griffith JD, Morris EP, Diffley JF (2009) Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* 139: 719-730
- Salsi V, Ferrari S, Ferraresi R, Cossarizza A, Grande A, Zappavigna V (2009) HOXD13 binds DNA replication origins to promote origin licensing and is inhibited by geminin. *Mol Cell Biol* 29: 5775-5788
- Seo S, Herr A, Lim JW, Richardson GA, Richardson H, Kroll KL (2005) Geminin regulates neuronal differentiation by antagonizing Brg1 activity. *Genes Dev* 19: 1723-1734

- Shen Z, Sathyan KM, Geng Y, Zheng R, Chakraborty A, Freeman B, Wang F, Prasanth KV, Prasanth SG (2010) A WD-repeat protein stabilizes ORC binding to chromatin. *Mol Cell* 40: 99-111
- Speck C, Chen Z, Li H, Stillman B (2005) ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nat Struct Mol Biol* 12: 965-971
- Stolk L, Zhai G, van Meurs JB, Verbiest MM, Visser JA, Estrada K, Rivadeneira F, Williams FM, Cherkas L, Deloukas P, Soranzo N, de Keyzer JJ, Pop VJ, Lips P, Lebrun CE, van der Schouw YT, Grobbee DE, Witteman J, Hofman A, Pols HA, Laven JS, Spector TD, Uitterlinden AG (2009) Loci at chromosomes 13, 19 and 20 influence age at natural menopause. *Nat Genet* 41: 645-647
- Tada S, Li A, Maiorano D, Mechali M, Blow JJ (2001) Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. *Nat Cell Biol* 3: 107-113
- Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, Mann M (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142: 967-980
- Volkening M, Hoffmann I (2005) Involvement of human MCM8 in prereplication complex assembly by recruiting hcdc6 to chromatin. *Mol Cell Biol* 25: 1560-1568
- Wu JR, Gilbert DM (1996) A distinct G1 step required to specify the Chinese hamster DHFR replication origin. *Science* 271: 1270-1272
- Yanagi K, Mizuno T, Tsuyama T, Tada S, Iida Y, Sugimoto A, Eki T, Enomoto T, Hanaoka F (2005) *Caenorhabditis elegans* geminin homologue participates in cell cycle regulation and germ line development. *J Biol Chem* 280: 19689-19694

# Regulation of DNA Replication Origin Licensing

Srikripa Chandrasekaran, Karen T. Reidy and Jeanette Gowen Cook  
*University of North Carolina at Chapel Hill  
United States of America*

## 1. Introduction

DNA replication is a fundamental biological process that serves to create two copies of the genetic material during each cell division. Complete and precise replication enables identical sets of genes to be faithfully delivered to daughter cells during each cell division. To achieve rapid duplication of the entire genome, eukaryotic cells initiate DNA replication at multiple locations on each chromosome termed *origins of DNA replication*. Origin DNA is unwound and complementary DNA is then synthesized from bi-directionally moving replication forks. The replication forks eventually merge to form two identical chromosomes.

The cell expends tremendous energy ensuring that a single origin of replication does not initiate replication twice within the same cell cycle. One of the most highly regulated steps in DNA replication is assembly of pre-replication complexes (pre-RCs). Pre-RC assembly begins as cells exit mitosis and continues through G1 phase, culminating in chromosomes poised for replication by the end of G1. At the onset of S phase, origins fire and replication begins. During this time, several overlapping mechanisms prevent pre-RC assembly on origins that have already fired to avoid utilizing any origins twice.

An abnormal situation in which replication is triggered multiple times from the same origin during a single cell cycle is termed *re-replication* (Figure 1). Re-replication is detrimental to genome stability, because it generates multiple replication forks on the same DNA strand. Ultimately such structures result in double strand breaks, genome instability, and in some cases, tumorigenesis (Arentson et al., 2002; Karakaidos et al., 2004; Xouri et al., 2004; Liontos et al., 2007). This chapter focuses on mechanisms to prevent re-replication during normal and perturbed cell cycles.

## 2. Pre-replication complex (pre-RC) assembly

To faithfully replicate its genomic information in a timely manner, a cell must initiate replication at thousands of sites across the genome. These origins of replication are prepared for replication through assembly of pre-RC complexes, beginning in late mitosis and continuing through G1 phase of the cell cycle. Origins with a fully assembled pre-RC are said to be "licensed" for replication. It is essential that origins assemble pre-RCs *only* in G1 because assembly of pre-RCs in S or G2 can lead to re-replication.

Pre-RC assembly begins when the six-subunit origin recognition complex (ORC) binds to an origin of replication (Figure 2). ORC is composed of the constitutively-expressed subunits Orc2-6, as well as the cell cycle-regulated Orc1 protein, and acts as an ATPase (Dhar et al.,

2001; Vashee et al., 2001; Bowers et al., 2004; Mendez et al., 2002). Once bound to origins, ORC recruits the remaining licensing factors Cdc6 and Cdt1 to origins (Cocker et al., 1996; Nishitani et al., 2000).

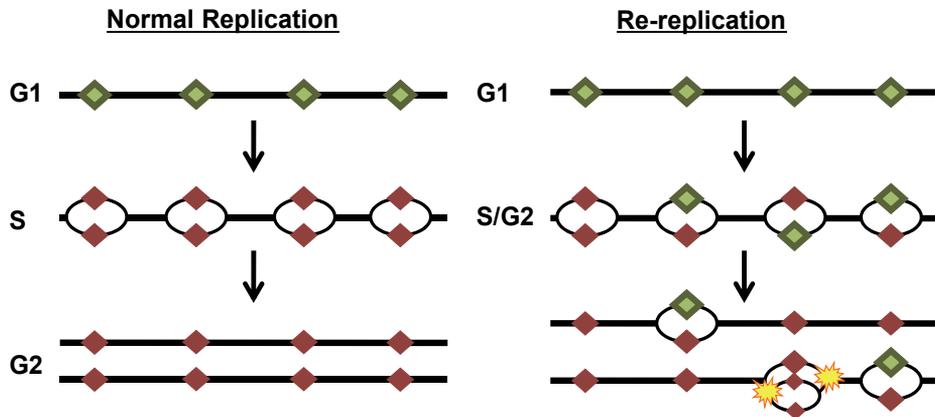


Fig. 1. Re-replication leads to double strand breaks. In a normal replicating cell (left), origins are licensed only in G1 phase (indicated by the green diamonds). Replication begins in S phase, and origins are inhibited from firing again (indicated by the red diamonds); replication is completed by G2 phase. In re-replicating cells (right), origins are licensed again in either S or G2 phase, leading to origin re-firing and re-replicated stretches of DNA. The consequences of re-replication include DNA damage, genome instability, and tumorigenesis.

Cdc6 was discovered in *Saccharomyces cerevisiae* and is essential for DNA replication; if Cdc6 is absent, yeast cells not only fail to replicate but also undergo reductional anaphase in which mitosis initiates without genome replication (Hartwell, 1976; Zwierschke et al., 1994; Piatti et al., 1995). Cdc6 is a member of the AAA+ ATPase family, and is closely related to Orc1 (Gaudier et al., 2007; Liu et al., 2000). ATP hydrolysis by Cdc6 and ORC is needed to load the helicase complex onto DNA (Randell et al., 2006; Bowers et al., 2004; Weinreich et al., 1999; Herbig et al., 1999; Donovan et al., 1997; Cook et al., 2002). Due to its tight association with ORC and its partially conserved DNA binding domain, it has been suggested that Cdc6 may also play a role in defining ORC binding sites (Mizushima et al., 2000).

Cdt1 was first discovered in *Schizosaccharomyces pombe* and, while possessing no enzymatic activity, is essential for the licensing reaction (Hofmann & Beach, 1994; Nishitani et al., 2000). Cdt1 binds the core replicative helicase Mini-Chromosome Maintenance (MCM) complex and recruits MCM to origins through direct interactions with ORC and Cdc6 (Tanaka & Diffley, 2002; Cook et al., 2004; Bruschi et al., 1995; Chen & Bell, 2011). While both Cdc6 and Cdt1 are needed to load the MCM complex, they bind in a sequential manner; Cdt1 can only bind to chromatin-bound Cdc6 and ORC (Tsuyama et al., 2005). Both Cdc6 and ORC hydrolyze ATP to load MCM complexes onto DNA (Randell et al., 2006). ATP hydrolysis by Cdc6 also releases Cdt1 to recruit additional MCM complexes (Randell et al., 2006). Once MCM complexes are loaded, the origin is licensed and can initiate replication once the MCM helicase is activated in S phase. After MCM complexes have been loaded, ORC, Cdc6, and Cdt1 are no longer needed, and replication can continue in their absence

(Donovan et al., 1997; Rowles et al., 1999; Maiorano et al., 2000). This property of the loaded MCM complex is key to preventing re-replication because, as discussed below, ORC, Cdc6, and Cdt1 are inactivated beginning in S phase.

At each origin, at least two MCM hexamer complexes are loaded at a time, with multiple rounds of loading at each origin (Evrin et al., 2009; Remus et al., 2009; Edwards et al., 2002; Lei et al., 1996). The exact mechanism of MCM loading is not currently understood, but electron microscopy images suggest ORC and Cdc6 form a structure similar to known clamp loaders such as RFC (Chen et al., 2008; Speck et al., 2005). While multiple MCM complexes can be loaded at each origin, perhaps as many ten copies per origin, the majority of the MCM complexes that associate with chromatin do not travel with the replication fork suggesting that they are not normally activated (Edwards et al., 2002; Krude et al., 1996; Dimitrova et al., 1999). These additional MCM complexes may be loaded as a backup mechanism to ensure that a sufficient number of origins fire in S phase (Ge & Blow, 2010).

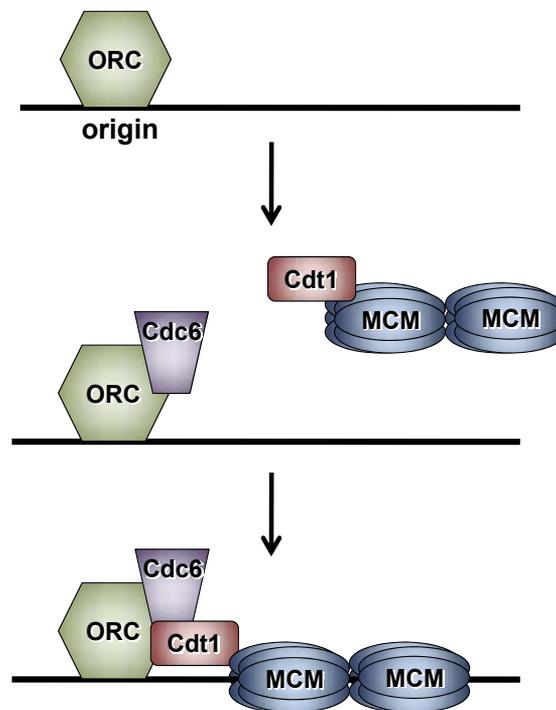


Fig. 2. Pre-RC assembly in G1 phase. Pre-RC assembly begins when the Origin Recognition Complex (ORC) binds to origin DNA. ORC recruits Cdc6, which in turn recruits Cdt1 bound to the Mini-Chromosome Maintenance (MCM) core helicase complex. Through the ATPase activity of ORC and Cdc6, the MCM complex is loaded onto DNA and the origin is licensed for replication.

MCM loading is highly regulated by multiple overlapping mechanisms (summarized in Table 1). Cdc6 and Cdt1 protein levels peak at different stages of the cell cycle; Cdt1 levels peak in G1 phase whereas Cdc6 peaks in S/G2 phase in mammalian cells (Nishitani et al., 2000; Petersen et al., 2000). Additionally, a member of the ORC complex, Orc1, is degraded or inactivated at the onset of S phase (Mendez et al., 2002; Li & DePamphilis, 2002; Li et al., 2004).

Due in part to these alternating protein levels, there are only two short windows in the cell cycle when pre-RC formation can occur. Pre-RC assembly begins at the end of mitosis, before the Anaphase Promoting Complex/Cyclosome (APC/C) becomes active in G1, and targets Cdc6 for degradation. The second round of pre-RC assembly occurs in late G1 phase when activated Cdk2 stabilizes Cdc6 but before Cdt1 is degraded at the onset of S phase (Diffley, 2004). Furthermore, the MCM subunits undergo post-translational modifications that facilitate MCM complex formation as well as their ability to be loaded onto DNA (Lin et al., 2008; Chuang et al., 2009). These mechanisms will be discussed in depth in the subsequent sections with specific emphasis on the regulation of metazoan pre-RC assembly.

Regulator	Substrate	Function
<b>Cyclin/Cdk</b>		
cyclin A/ Cdk2	Cdt1	ubiquitination by SCF <sup>Skp2</sup>
	Cdc6 (Ser106)	nuclear export of Cdc6
	Orc1	nuclear export and/or degradation depending on cell type
cyclin E/ Cdk2	Cdc6 (Ser54)	stabilization by protection from APC/C <sup>Cdh1</sup> ubiquitination
	Mcm3	facilitates Mcm (2-7) complex formation
<b>Ubiquitin ligases</b>		
SCF <sup>Skp2</sup>	Cdt1	ubiquitination in S phase
APC/ C <sup>Cdc20</sup>	cyclin A	ubiquitination in G2 phase
APC/ C <sup>Cdh1</sup>	geminin	ubiquitination in M and G1 phases
	Cdc6	
CRL4 <sup>Cdt2</sup>	Cdt1	ubiquitination in S phase and after DNA damage
	Set8	
<b>Acetyltransferases</b>		
Gcn5	Cdc6	Promotes phosphorylation of Ser6 by cyclin A/Cdk2
Hbo1	histone H4	promotes pre-RC formation (Cdt1 stimulation?)
	pre-RC proteins?	
<b>Methyltransferases</b>		
Set8	histone H4 (K20)	Monomethylation promotes chromatin condensation and pre-RC assembly

Table 1. Regulation of pre-RC components and related factors.

### 3. Pre-RC regulation

Pre-RC formation and origin licensing occurs during late mitosis and early G1 phase of the cell cycle (Bell & Dutta, 2002). It is important to limit the assembly of pre-RCs to only the G1 phase since failure to restrict pre-RC formation results in re-replication. Re-replication can



### 3.1.1 Cyclin-Cdks as negative regulators of pre-RC assembly

Several experimental findings support a role for Cdks in preventing re-replication. For example, transient over-expression of the Cdk inhibitor, p21, in G2 caused re-replication (Bates et al., 1998). In addition, ORC and MCM complexes were recruited to human chromatin in G2 when Cdk activity was inhibited genetically or pharmacologically (Coverley et al., 1996; Fujita et al., 1998; Ballabeni et al., 2004; Li et al., 2004; Sugimoto et al., 2004). Finally, chemical inactivation of Cdk1, combined with genetic ablation of Cdk2, allowed pre-RCs to assemble inappropriately during mitosis (Ballabeni et al., 2004). These observations highlight a role for Cdks in preventing inappropriate origin licensing.

Re-licensing of origins during S phase is prevented, in part, by Cdk2, in association with cyclin A (Wheeler et al., 2008). Cdt1 interacts with the S phase cyclin A/Cdk2 complex, which results in Cdt1 phosphorylation at threonine 29 (Li et al., 2004; Liu et al., 2004). Phosphorylated Cdt1 binds to the F-box protein, Skp2, the substrate receptor for the ubiquitin ligase SCF<sup>Skp2</sup>. Cdt1 is polyubiquitinated by SCF<sup>Skp2</sup> and targeted for degradation by the 26S proteasome, thus reducing the pool of Cdt1 protein available to participate in origin licensing (Takeda et al., 2005; Sugimoto et al., 2004; Kim & Kipreos, 2007).

Coincident with Cdt1 destruction in S phase, Cdc6 is acetylated by Gcn5 on lysines 92, 105, and 109 which promotes cyclin A/Cdk2 phosphorylation on Cdc6 at serine 106 (Paolinelli et al., 2009; Mailand & Diffley, 2005). Serine 106 phosphorylation results in exclusion of Cdc6 protein from the nucleus, preventing re-replication (Paolinelli et al., 2009; Saha et al., 1998; Coverley et al., 2000; Kim et al., 2007; Fujita et al., 1999; Jiang et al., 1999; Petersen et al., 1999). The small amount of Cdc6 that remains nuclear throughout S phase is chromatin-bound and likely participates in the ATR-dependent intra-S phase checkpoint by mechanisms that are not yet understood (Mendez & Stillman, 2000; Lau et al., 2006).

Additionally, the Orc1 subunit of ORC is phosphorylated by cyclin A/Cdk1 during S phase, and this phosphorylation promotes Orc1 degradation in HeLa cells (Mendez et al., 2002). The same phosphorylation on Orc1 in Chinese Hamster Ovary cells (CHO) does not affect Orc1 stability, but lowers the affinity of Orc1 for chromatin (Li et al., 2004). In both HeLa and CHO cells, Orc1 phosphorylation allows the export of Orc1 to the cytoplasm (Saha et al., 2006). Over-expression of cyclin A from Kaposi's Sarcoma-associated herpes virus also facilitates re-localization of Orc1 to the cytoplasm. These results show that Orc1 is subject to phosphorylation by cyclin A/Cdk1, and this event modulates the stability and/or localization of Orc1, thereby contributing to the prevention of re-replication.

Recent evidence from *S. cerevisiae* suggests that Orc2 and Orc6 may also be targets of cyclin/Cdk inhibition. Phosphorylation of these subunits leads to a marked decrease in MCM loading (Green et al., 2006; Tanny et al., 2006; Nguyen et al., 2001). Interaction between Orc6 and the S phase Cdk, Clb5, is needed to prevent MCM loading outside of G1 phase; this interaction occludes the Cdt1 binding site on the ORC complex (Wilmes et al., 2004; Tanny et al., 2006; Chen & Bell, 2011). In addition to steric hindrance, Clb5 phosphorylates Orc6; this modification also partially blocks the Cdt1 binding site and prevents MCM loading (Chen & Bell, 2011). It remains to be determined if similar mechanisms also apply to Cdk regulation of mammalian ORC (DePamphilis, 2005).

### 3.1.2 Cyclin-Cdks as positive regulators of pre-RC assembly

Cdc6 protein levels are very low in both quiescent cells and in early/mid G1 phase cells due to ubiquitin-mediated proteolysis, but Cdc6 protein accumulates in late G1 just prior to a

burst of MCM loading. In late G1, cyclin E/Cdk2 phosphorylates human Cdc6 on serine 54, which protects Cdc6 from the ubiquitin E3 ligase APC/C<sup>Cdh1</sup> (Mailand & Diffley, 2005). Many cancer cells have both high cyclin E/Cdk2 activity and high Cdc6 protein levels which may reflect not only the transcriptional up-regulation of *cyclin E* and *cdc6* genes in tumors but also the stabilizing effect of cyclin E/Cdk2 on Cdc6 protein (Nielsen et al., 1996). Cyclin E may also have an as yet poorly understood direct role in MCM loading through association with Cdt1 and MCM subunits independently of its association with Cdk2. When cyclin E was depleted from mammalian cells, MCM loading did not occur as cells re-entered G1 from quiescence (Geng et al., 2007; Geng et al., 2003).

In addition to regulating Cdc6, Cdk complexes also facilitate the formation and loading of the MCM helicase complex. Mcm3 is phosphorylated on serine 112 by Cdk1, which triggers MCM complex assembly (Lin et al., 2008). Cdk activity affects MCM loading in an indirect fashion as well. The activity of cyclin E/Cdk2 is required for the accumulation of Cdc7 mRNA, which in turn functions in origin firing (Chuang et al., 2009; Francis et al., 2009; Jiang et al., 1999; Masai et al., 2000; Sheu & Stillman, 2006). In quiescent cells, phosphorylation of Mcm2 at serine 5 is necessary to promote MCM loading (Chuang et al., 2009; Geng et al., 2007).

### 3.2 Cdk-independent regulation of pre-RCs

#### 3.2.1 Regulation of Cdt1 by CRL4<sup>Cdt2</sup>

Cdk-independent mechanisms also prevent re-replication by targeting Cdt1. Non-phosphorylatable (Cdk-resistant) Cdt1 mutants are degraded during S phase despite being unable to bind to Skp2 (Takeda et al., 2005; Nishitani et al., 2006; Senga et al., 2006). An alternate, DNA-dependent, mechanism for Cdt1 degradation was subsequently uncovered (Jin et al., 2006; Arias & Walter, 2006; Senga et al., 2006; Higa et al., 2006; Nishitani et al., 2006; Hu & Xiong, 2006; Ralph et al., 2006). Proliferating Cell Nuclear Antigen (PCNA) is a cofactor required to increase the processivity of DNA pol $\delta$  during leading strand synthesis (Maga & Hubscher, 2003). Thus, PCNA travels with active replication forks. Cdt1 interacts with PCNA through a highly conserved region called the PIP box during S phase. Cdt1 binding to DNA-loaded PCNA is essential for Cdt1 recognition by the CRL4<sup>Cdt2</sup> ubiquitin E3 ligase. CRL4 associates with Cdt1 via direct binding to the Cdt2 substrate adapter which links to the Cul4 scaffold. Since Cdt1 only binds PCNA on chromatin, it is only ubiquitinated by chromatin-associated CRL4<sup>Cdt2</sup>. In this manner Cdt1 degradation is directly coupled to DNA synthesis. A Cdt1 mutant that cannot bind either PCNA or cyclin/Cdk is stable during S phase and causes re-replication (Nishitani et al., 2006). Likewise, Cdt2 depletion stabilizes Cdt1 in S phase, causing re-licensing of fired origins, and extensive re-replication (Jin et al., 2006; Higa et al., 2006; Ralph et al., 2006).

#### 3.2.2 Regulation of Cdt1 by geminin accumulation

As another layer of regulation, metazoan S phase cells accumulate the protein geminin, which binds to Cdt1 and blocks the Cdt1-MCM interaction (Lee et al., 2004; Wohlschlegel et al., 2000; Tada et al., 2001). Geminin is expressed throughout the S, G2 and M phases of the cell cycle when origin licensing is inhibited (McGarry & Kirschner, 1998; Wohlschlegel et al., 2000). Geminin is targeted for degradation in G1, but begins to accumulate at the end of G1 when the ubiquitin ligase responsible for its degradation (APC/C<sup>Cdh1</sup>) is inactivated (McGarry & Kirschner, 1998; Rape et al., 2006). In S phase, geminin binds to the residual Cdt1 that escaped degradation and renders it unavailable to re-license origins.

Recent biochemical evidence has suggested that geminin-Cdt1 complexes exist in several forms (De Marco et al., 2009). These forms include a licensing-inhibitory heterohexamer that consists of two Cdt1 molecules and four geminin molecules, and a licensing-permissive heterotrimer, comprised of one Cdt1 molecule and two geminin molecules (Lutzmann et al., 2006). Binding of geminin to Cdt1 in a heterohexamer can tether several Cdt1 molecules together, creating chromatin-bound foci that may cooperatively inhibit licensing (Ode et al., 2011). Depending on the amount of geminin in the cell, geminin may switch from being an inhibitor of origin licensing when geminin levels are high to an activator when levels of geminin are low.

### 3.2.3 APC/C as a regulator of pre-RC assembly

The cell spends a significant amount of energy to ensure that the correct proteins are expressed at the appropriate time. Before one cell cycle phase begins, cells ensure that the previous step has been properly completed and, in many cases, inactivated by controlling protein activity abundance. One mechanism for enforcing the proper order of events is through regulated protein degradation. The Anaphase Promoting Complex/Cyclosome (APC/C) is uniquely tied to cell cycle progression and control of DNA replication as evidenced by the fact its regulation and activity are modulated in every phase of the cell cycle. APC/C is a RING-type E3 ubiquitin ligase originally discovered through its association with its substrates, the mitotic cyclins (Sudakin et al., 1995; King et al., 1995). Two activator subunits, Cdc20 and Cdh1, interact dynamically with the APC/C holoenzyme to influence substrate recognition (Figure 4). APC/C<sup>Cdh1</sup> targets in G1 include Skp2, a member of the SCF ubiquitin ligase complex, the licensing factor Cdc6, and the inhibitor protein geminin (Wei et al., 2004; Petersen et al., 2000; McGarry & Kirschner, 1998). Degradation of Skp2 results in accumulation of the Cdk2 inhibitors p21 and p27, and prevents premature S phase entry due to low Cdk2 activity (Wei et al., 2004). APC/C<sup>Cdh1</sup> also acts to limit the amount of Cdc6 that is available in the cell (discussed in 3.1.2).

As cells progress through S phase, Cdh1 is phosphorylated by cyclin A/Cdk2 complexes; since hyper-phosphorylated Cdh1 cannot interact with APC/C, the ubiquitin ligase complex is inactive (Zachariae et al., 1998). This inactivation allows geminin to accumulate and bind any remaining Cdt1 (McGarry & Kirschner, 1998). During S phase, APC/C<sup>Cdh1</sup> is also bound by its inhibitor protein, Emi1 (Hsu et al., 2002). Interestingly, Emi1 accumulation is not needed to begin S phase but is needed to signal the stop of replication and mitotic entry, even though APC/C<sup>Cdc20</sup> can still ubiquitinate its targets if Emi1 is present in mitosis (Di Fiore & Pines, 2007). Emi1 remains bound until prophase, when it is phosphorylated by Plk1 (Moshe et al., 2004). While Emi1 accumulation is not needed for S phase entry, it is essential to inhibit re-replication. Depletion of Emi1 leads to re-replication in human cells, due to the untimely activation of APC/C<sup>Cdh1</sup> (Sivaprasad et al., 2007; Machida & Dutta, 2007). This stabilization allows geminin levels to drop when Cdt1 levels are high; at the same time, increased activity of cyclin A/Cdk2 allows Cdc6 to become stabilized. With both licensing factors present, origins are licensed outside of G1 and re-replication occurs.

## 4. Re-replication causes DNA damage and cell cycle arrest

Manipulation of key components of the replication machinery can induce re-replication. For instance, over-expression of replication proteins Cdt1 and Cdc6 causes re-replication in human cancer cells (Nishitani & Nurse, 1995; Vaziri et al., 2003; Ekholm-Reed et al., 2004;

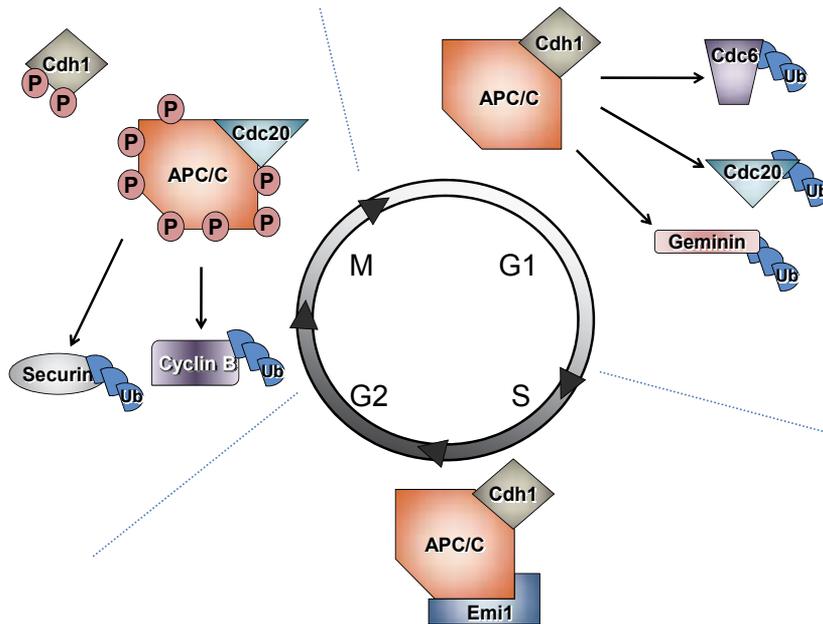


Fig. 4. APC/C is cell cycle regulated and controls several important pre-RC proteins. In G1 phase, APC/C is bound to the adaptor protein Cdh1 and APC/C<sup>Cdh1</sup> ubiquitinates the licensing factor Cdc6 and the Cdt1 inhibitor geminin. During this time, it also targets another APC/C adaptor protein, Cdc20, for degradation. As cells progress into S phase, APC/C<sup>Cdh1</sup> is bound by its inhibitor protein, Emi1. At the beginning of mitosis, APC/C and Cdh1 become hyper-phosphorylated, leading to dissociation of Cdh1 from APC/C. Phosphorylated APC/C can then bind the adaptor protein, Cdc20. The APC/C<sup>Cdc20</sup> complex is responsible for degrading cyclin B and securin, thereby promoting sister chromatid separation and mitotic exit.

Tanaka & Diffley, 2002; Lontos et al., 2007). Depletion of the Cdt1 inhibitor geminin also induces re-replication in certain cancer cell types (Melixetian et al., 2004; Zhu et al., 2004). Depletion or inhibition of the CRL4<sup>Cdt2</sup> ubiquitin E3 ligase, the machinery responsible for the ubiquitination of Cdt1, stabilizes Cdt1 and stimulates re-replication (Jin et al., 2006; Zhong et al., 2003; Lin et al., 2010). Thus, misregulation of pre-RC components or proteins involved in regulating the pre-RC complex can promote re-replication.

Re-replication activates the DNA damage response (Blow & Dutta, 2005). Mammalian cells depleted of geminin, Cdt2, or Ddb1 (another subunit of CRL4<sup>Cdt2</sup>), fail to properly down-regulate Cdt1 during S phase. These cells, or cells that overproduce Cdt1, also acquire phosphorylated  $\gamma$ H2AX, a histone variant incorporated into chromatin at sites of DSBs (Vaziri et al., 2003; Melixetian et al., 2004; Zhu et al., 2004; Jin et al., 2006; Lovejoy et al., 2006; Zhu & Dutta, 2006). The appearance of double stranded break markers (DSBs) in re-replicating cells suggests that multiple forks on the same DNA result in replication fork collision and/or collapse.

Checkpoint sensors recognize impaired fork activity, generated by either fork collapse or stalled replication forks, and shut down cell cycle progression. This DNA damage response involves activation of the serine-threonine kinases, ATM (Ataxia Telangiectasia Mutated)

and ATR (ATM-related) (Shiloh, 2003). Single-stranded regions are recognized by ATR, which preferentially activates and phosphorylates the Chk1 protein kinase (Guo et al., 2000; Zhao & Piwnica-Worms, 2001; Kramer et al., 2004; Niida et al., 2007; Zou & Elledge, 2003). Chk1 in turn inactivates the Cdc25 phosphatase causing an intra-S or G2/S phase arrest (Donzelli & Draetta, 2003; Jin et al., 2003; Ferguson et al., 2005). Inactivation of Cdc25 is also accomplished by induction of MAPKAP-K2 (MK2), a downstream effector kinase of the p38 MAP kinase pathway, which is stimulated by genotoxic stress (Manke et al., 2005; Reinhardt et al., 2007; Lemaire et al., 2006; Huard et al., 2008). Without active Cdc25, the mitotic cyclin B/Cdk1 complex remains phosphorylated and inactive. Thus, the DNA damage response pathway leads to a G2/M cell cycle arrest (Tang et al., 2006).

Damage caused by DSBs is recognized by ATM (Shechter et al., 2004; Costanzo et al., 2001). ATM activates and phosphorylates the checkpoint protein kinase Chk2 (Melchionna et al., 2000; Zhou & Elledge, 2000). Activated Chk2 then phosphorylates and activates the transcriptional regulator, p53 (Meek, 1994; Milczarek et al., 1997). Active p53 induces expression of many genes needed for DNA repair and cell cycle arrest, including the Cdk inhibitor, p21 (Smith et al., 1995). The accumulation of p21 inhibits Cdk2 function and blocks cells from entering S phase (Taylor & Stark, 2001; Levine, 1997; Vogelstein et al., 2000). DNA damage in G2 phase also leads to p53-dependent p21 accumulation, inhibition of Cdk1, and arrest in G2 (Agarwal et al., 1995; Deng et al., 1995; Brugarolas et al., 1995).

Cells can recover from arrest by activating repair pathways and continuing cell division, but if the damage is too extensive, they die by apoptosis (Hartwell & Weinert, 1989; Weinert & Hartwell, 1990; Nasmyth, 1996; Abraham, 2001). Re-replication induced by Cdt1 stabilization or geminin depletion results in many of the checkpoint events outlined above, including phosphorylation of Chk2, Chk1, and p53, failure to enter mitosis or initiation of apoptosis depending on the degree of re-replication-associated damage (Zhu & Dutta, 2006; Zhu et al., 2004; McGarry, 2002). G2/M arrest caused by loss of geminin is primarily mediated by ATR recognition of single-stranded regions generated early in re-replication. As cells eventually accumulate more damage from more extensive re-replication and generate DSBs, the ATM pathway is induced (Lin & Dutta, 2007). Such sequential activation may permit low-level damage to be repaired during an ATR-mediated G2/M arrest prior to ATM-mediated p53 activation and permanent arrest or cell death.

Interestingly, ATR-deficient cells show a higher propensity to re-replicate than their ATR-proficient counterparts do (Lin & Dutta, 2007). These observations suggest that the ATR pathway can restrict re-replication while re-replication simultaneously activates ATR. Another potential interplay between replication licensing and the ATR pathway is suggested by the binding of Cdc6 to ATR, which promotes ATR chromatin binding (Yoshida et al., 2010). Cdc6 can also activate p21-bound Cdk2 and override the DNA damage checkpoint (Kan et al., 2008).

## 5. Pre-RC regulation during a DNA damage response

As described above, cyclin/Cdk activity prevents inappropriate formation of pre-RC complexes in S and G2 phases of the cell cycle. Exposure to DNA-damaging agents or generation of DSBs during replication elicits the DNA-damage response, which can reduce Cdk activity, particularly in G2 phase. Therefore, DNA damage can perturb the normal Cdk-dependent regulation of pre-RC components such that there is potential danger of

initiating re-replication events. To counteract this risk, the DNA damage induces Cdk-independent mechanisms to inhibit pre-RC assembly.

### 5.1 Degradation of Cdc6

When cells are exposed to various genotoxins, such as ultraviolet (UV) irradiation, ionizing radiation (IR), base alkylating agents, etc., Cdc6 is actively degraded by two mechanisms. As outlined above, the DNA damage response results in p53-dependent induction of the Cdk inhibitor p21<sup>Cip1</sup>. Reduced Cdk activity prevents phosphorylation of Cdc6 at serine 54 (Duursma & Agami, 2005). Without this protective phosphorylation, Cdc6 can be ubiquitinated by APC/C<sup>Cdh1</sup> and degraded, particularly in G1 phase due to the high activity of APC/C<sup>Cdh1</sup> (Mailand & Diffley, 2005). Cdc6 ubiquitination by APC/C<sup>Cdh1</sup> is further facilitated by the activation of APC/C<sup>Cdh1</sup> by the DNA damage checkpoint and cyclin destruction (which also reduces Cdc6 serine 54 phosphorylation) (Sudo et al., 2001).

Cdc6 is also targeted for degradation by a different E3 ubiquitin ligase, Huwe1 (Figure 5). Recognition of Cdc6 by Huwe1 is independent of both p53 induction of p21 and changes in Cdk-mediated phosphorylation (Hall et al., 2007). Huwe1 is a monomeric HECT family E3 ligase that has been linked to the regulation of many proteins central to cell proliferation control, including maintaining low p53 levels in the absence of DNA damage (Chen et al., 2005). Huwe1-mediated Cdc6 degradation is also activated in response to endogenous re-replication-associated DNA damage ultimately leading to loss of Cdc6 from re-replicating cells (Hall et al., 2008).

### 5.2 Degradation of Cdt1

When cells are exposed to exogenous DNA damaging agents, Cdt1 is rapidly degraded (Higa et al., 2003; Hu et al., 2004). DNA damage-induced Cdt1 degradation is unaffected by Cdk activity, and is therefore unrelated to ubiquitination by the SCF<sup>Skp2</sup> ubiquitin ligase. Instead, Cdt1 is degraded through association with the CRL4<sup>Cdt2</sup> ubiquitin ligase (Higa et al., 2003; Hu & Xiong, 2006; Jin et al., 2006; Nishitani et al., 2006; Ralph et al., 2006; Sansam et al., 2006; Senga et al., 2006). PCNA is loaded onto damaged DNA not only during normal replication, but also in response to DNA damage both for DNA repair synthesis and in a checkpoint role (Lee & Myung, 2008). As a result, Cdt1 interacts with loaded PCNA after DNA damage in the same way that it interacts with PCNA during S phase. The interaction of Cdt1 with PCNA at sites of DNA damage recruits CRL4<sup>Cdt2</sup> via direct binding of Cdt2 to the Cdt1-PCNA complex, leading to Cdt1 ubiquitination and proteasomal degradation.

Like Cdc6, Cdt1 is targeted for destruction in response to both exogenous and endogenous DNA damage. Human or *Drosophila* cells depleted of geminin also rapidly lose Cdt1 and Cdc6 (Mihaylov et al., 2002; Ballabeni et al., 2004). Since the lack of geminin induces re-replication-related DNA damage, Cdt1 and Cdc6 are simultaneously targeted for degradation by CRL4<sup>Cdt2</sup> and Huwe1, respectively (Hall et al., 2008). In cells depleted of either CRL4<sup>Cdt2</sup> components or Huwe1, geminin depletion allowed more extensive re-replication, suggesting that the destruction of pre-RC proteins in re-replicating cells protects them from even further re-replication.

### 5.3 Phosphorylation of pre-RC components by DNA damage-stimulated kinases

Cdt1 and Cdc6 destruction after DNA damage is largely independent of the DNA damage checkpoint, but other pre-RC members are specifically targeted by the ATM and ATR pathways. Mcm2 is phosphorylated in both *Xenopus* and human cancer cells (on serine 92

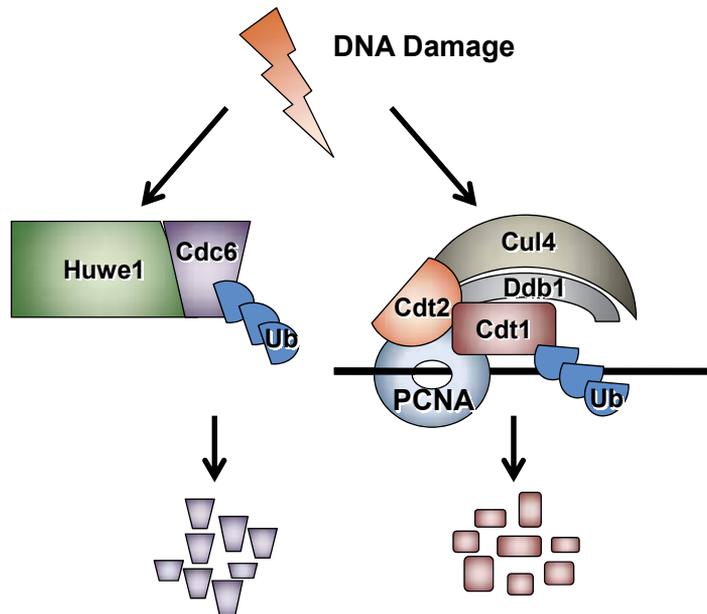


Fig. 5. DNA damage causes degradation of Cdc6 and Cdt1. Exogenous genotoxic damage and endogenous re-replication-associated damage induce the degradation of the licensing factors Cdc6 and Cdt1. Cdc6 is targeted by the ubiquitin ligase Huwe1. Cdt1 binds to PCNA, which is loaded at sites of DNA damage. This association allows the CRL4<sup>Cdt2</sup> ubiquitin ligase to bind Cdt1 and target it for degradation.

and serine 108, respectively) in response to many DNA damaging agents (Yoo et al., 2004; Cortez et al., 2004). In *Xenopus* egg extracts, both ATM and ATR phosphorylate the Mcm2 subunit of the MCM complex, whereas in human cells ATR is the primary kinase involved (Yoo et al., 2004; Cortez et al., 2004). Mcm3 and Mcm4 are also phosphorylated in response to DNA damage. Mcm4 becomes phosphorylated after DNA damage in a caffeine-sensitive manner. Caffeine is a known inhibitor of the ATM and ATR kinases, so these kinases or their downstream targets are presumed to be responsible. Mcm3 is phosphorylated on serine 535 by ATM in response to ionizing radiation (Cortez et al., 2004). Additionally, Orc3 is phosphorylated in response to ionizing radiation (Matsuoka et al., 2007). The functional consequences of these phosphorylations are not yet known, but it is reasonable to predict that these modifications prevent pre-RC assembly.

## 6. Novel interfaces between pre-RC regulation and cell cycle control

### 6.1 Additional control of replication licensing

In addition to the core pre-RC members described thus far, other proteins have also been implicated in regulating DNA replication licensing. A protein related to MCM subunits, Mcm9, is not itself incorporated into the core helicase complex, but has been suggested to be essential for licensing in eukaryotes. Mcm9 contains an ATPase and a helicase domain, and may directly bind and influence Cdt1 function (Lutzmann & Mechali, 2008). Geminin and Mcm9 compete for binding to Cdt1; as such, Mcm9 could protect Cdt1 from geminin during origin licensing (Lutzmann & Mechali, 2009).

Another protein that plays a role in influencing DNA replication is the acetyltransferase, Hbo1. Hbo1 binds both Orc1 and Mcm2, and is thereby recruited to origins (Miotto & Struhl, 2008; Iizuka et al., 2006; Burke et al., 2001). Hbo1 depletion causes licensing defects in human cells and in *Xenopus* egg extracts by preventing the loading of MCM complexes but not the chromatin association of ORC, Cdc6, or Cdt1 (Iizuka et al., 2006; Miotto & Struhl, 2008). These defects can be reversed by over-expressing Cdt1, suggesting that Hbo1 is an activator of MCM loading at the Cdt1 step in pre-RC assembly. Hbo1 acetylates histone H4 which may reflect a role for Hbo1 in modulating chromatin at origins (Miotto & Struhl, 2010). Hbo1 also associates with the tumor suppressor, p53, and this association negatively regulates Hbo1 enzymatic activity. Cellular stress agents other than DNA damage, such as hyperosmotic stress, activate p53 (Vogelstein et al., 2000), and this activation was suggested to inhibit the HAT activity of Hbo1 and consequently, origin licensing (Iizuka et al., 2008). Hbo1 can also acetylate multiple pre-RC components *in vitro* however, so its role in pre-RC assembly may also be through direct modification of licensing proteins (Iizuka et al., 2006). In addition to histone acetylation, monomethylation of histone H4 facilitates pre-RC formation. Tethering the H4 Lys20 methyltransferase, Set8 (PR-Set7), to an ectopic sequence induced pre-RC formation at that site, suggesting that H4K20 methylation may also promote pre-RC assembly at normal chromosomal origins (Tardat et al., 2010). Set8 is ubiquitinated by the CRL4<sup>Cdt2</sup> ubiquitin ligase in PCNA-dependent manner during S phase and in response to DNA damage (Jorgensen et al., 2011; Wu & Rice, 2010; Oda et al., 2010). Failure to degrade Set8 in S phase leads to re-replication and lack of chromatin condensation during mitosis (Wu & Rice, 2010; Jorgensen et al., 2011). These data further suggest an involvement of chromatin structure in the regulation of origin licensing.

Each of the genes encoding pre-RC components is transcriptionally regulated by the Rb-E2F pathway. Given that tumor cells frequently exhibit high-level expression of E2F target genes, (Rb or p16 loss, cyclin overproduction, etc.) it is not surprising that Cdt1 and Cdc6 are overproduced in many cancers (Ohta et al., 2001; Karakaidos et al., 2004; Pinyol et al., 2006; Di Micco et al., 2006). Overproduction of Cdt1 or Cdc6 in cultured human cells induces re-replication, raising the possibility that tumor cells also re-replicate *in vivo*. Recently it has been suggested that cancer cells “hyper-replicate” and that this form of replication stress is a driving force in oncogenesis. It has also been suggested that excessive pre-RC assembly may even downregulate expression of the INK4/ARF tumor suppressor locus due to interference between a nearby origin and the INK4 promoter (Gonzalez et al., 2006). Recently, mutations in genes for several components of the pre-RC, including Orc1, Orc4, Orc6, Cdt1, and Cdc6 have been linked to the autosomal recessive primordial dwarfism syndrome, Meier Gorlin syndrome (Guernsey et al., 2011; Bicknell et al., 2011). This is the first report implicating impaired licensing in a developmental disorder. Taken together there are now clear links between pre-RC formation, normal human development, and tumorigenesis.

## 7. References

- Abraham, R.T. (2001). Cell cycle checkpoint signaling through the ATM and ATR kinases. *Genes Dev*, Vol. 15, No. 17, 2177-2196.
- Agarwal, M.L., Agarwal, A., Taylor, W.R. and Stark, G.R. (1995). p53 controls both the G2/M and the G1 cell cycle checkpoints and mediates reversible growth arrest in human fibroblasts. *Proc Natl Acad Sci U S A*, Vol. 92, No. 18, 8493-8497.

- Arentson, E., Faloon, P., Seo, J., *et al.* (2002). Oncogenic potential of the DNA replication licensing protein CDT1. *Oncogene*, Vol. 21, No. 8, 1150-1158.
- Arias, E.E. and Walter, J.C. (2006). PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. *Nat Cell Biol*, Vol. 8, No. 1, 84-90.
- Ballabeni, A., Melixetian, M., Zamponi, R., Masiero, L., Marinoni, F. and Helin, K. (2004). Human geminin promotes pre-RC formation and DNA replication by stabilizing CDT1 in mitosis. *Embo J*, Vol. 23, No. 15, 3122-3132.
- Bates, S., Ryan, K.M., Phillips, A.C. and Vousden, K.H. (1998). Cell cycle arrest and DNA endoreduplication following p21Waf1/Cip1 expression. *Oncogene*, Vol. 17, No. 13, 1691-1703.
- Bell, S.P. and Dutta, A. (2002). DNA replication in eukaryotic cells. *Annu Rev Biochem*, Vol. 71, No. 333-374.
- Bicknell, L.S., Bongers, E.M., Leitch, A., *et al.* (2011). Mutations in the pre-replication complex cause Meier-Gorlin syndrome. *Nat Genet*, Vol. 43, No. 4, 356-359.
- Blow, J.J. and Dutta, A. (2005). Preventing re-replication of chromosomal DNA. *Nat Rev Mol Cell Biol*, Vol. 6, No. 6, 476-486.
- Bochman, M.L. and Schwacha, A. (2009). The Mcm complex: unwinding the mechanism of a replicative helicase. *Microbiol Mol Biol Rev*, Vol. 73, No. 4, 652-683.
- Bowers, J.L., Randell, J.C., Chen, S. and Bell, S.P. (2004). ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. *Mol Cell*, Vol. 16, No. 6, 967-978.
- Braden, W.A., McClendon, A.K. and Knudsen, E.S. (2008). Cyclin-dependent kinase 4/6 activity is a critical determinant of pre-replication complex assembly. *Oncogene*, Vol. 27, No. 56, 7083-7093.
- Brugarolas, J., Chandrasekaran, C., Gordon, J.I., Beach, D., Jacks, T. and Hannon, G.J. (1995). Radiation-induced cell cycle arrest compromised by p21 deficiency. *Nature*, Vol. 377, No. 6549, 552-557.
- Bruschi, C.V., Mcmillan, J.N., Coglievina, M. and Esposito, M.S. (1995). The genomic instability of yeast *cdc6-1/cdc6-1* mutants involves chromosome structure and recombination. *Mol Gen Genet*, Vol. 249, No. 1, 8-18.
- Burke, T.W., Cook, J.G., Asano, M. and Nevins, J.R. (2001). Replication factors MCM2 and ORC1 interact with the histone acetyltransferase HBO1. *J Biol Chem*, Vol. 276, No. 18, 15397-15408.
- Chen, D., Kon, N., Li, M., Zhang, W., Qin, J. and Gu, W. (2005). ARF-BP1/Mule is a critical mediator of the ARF tumor suppressor. *Cell*, Vol. 121, No. 7, 1071-1083.
- Chen, S. and Bell, S.P. (2011). CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes Dev*, Vol. 25, No. 4, 363-372.
- Chen, Z., Speck, C., Wendel, P., Tang, C., Stillman, B. and Li, H. (2008). The architecture of the DNA replication origin recognition complex in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA*, Vol. 105, No. 30, 10326-10331.
- Chuang, L.C., Teixeira, L.K., Wohlschlegel, J.A., *et al.* (2009). Phosphorylation of Mcm2 by Cdc7 promotes pre-replication complex assembly during cell-cycle re-entry. *Mol Cell*, Vol. 35, No. 2, 206-216.
- Cocker, J.H., Piatti, S., Santocanale, C., Nasmyth, K. and Diffley, J.F. (1996). An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. *Nature*, Vol. 379, No. 6561, 180-182.

- Cook, J.G., Chasse, D.A. and Nevins, J.R. (2004). The regulated association of Cdt1 with minichromosome maintenance proteins and Cdc6 in mammalian cells. *J Biol Chem*, Vol. 279, No. 10, 9625-9633.
- Cook, J.G., Park, C.H., Burke, T.W., *et al.* (2002). Analysis of Cdc6 function in the assembly of mammalian prereplication complexes. *Proc Natl Acad Sci USA*, Vol. 99, No. 3, 1347-1352.
- Cortez, D., Glick, G. and Elledge, S.J. (2004). Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proc Natl Acad Sci U S A*, Vol. 101, No. 27, 10078-10083.
- Costanzo, V., Robertson, K., Bibikova, M., *et al.* (2001). Mre11 protein complex prevents double-strand break accumulation during chromosomal DNA replication. *Mol Cell*, Vol. 8, No. 1, 137-147.
- Coverley, D., Pelizon, C., Trewick, S. and Laskey, R.A. (2000). Chromatin-bound Cdc6 persists in S and G2 phases in human cells, while soluble Cdc6 is destroyed in a cyclin A-cdk2 dependent process. *J Cell Sci*, Vol. 113 ( Pt 11), No. 1929-1938.
- Coverley, D., Wilkinson, H.R. and Downes, C.S. (1996). A protein kinase-dependent block to reinitiation of DNA replication in G2 phase in mammalian cells. *Exp Cell Res*, Vol. 225, No. 2, 294-300.
- De Marco, V., Gillespie, P.J., Li, A., *et al.* (2009). Quaternary structure of the human Cdt1-Geminin complex regulates DNA replication licensing. *Proc Natl Acad Sci U S A*, Vol. 106, No. 47, 19807-19812.
- Deng, C., Zhang, P., Harper, J.W., Elledge, S.J. and Leder, P. (1995). Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell*, Vol. 82, No. 4, 675-684.
- Depamphilis, M.L. (2005). Cell cycle dependent regulation of the origin recognition complex. *Cell Cycle*, Vol. 4, No. 1, 70-79.
- Dhar, S.K., Yoshida, K., Machida, Y., *et al.* (2001). Replication from oriP of Epstein-Barr virus requires human ORC and is inhibited by geminin. *Cell*, Vol. 106, No. 3, 287-296.
- Di Fiore, B. and Pines, J. (2007). Emi1 is needed to couple DNA replication with mitosis but does not regulate activation of the mitotic APC/C. *J Cell Biol*, Vol. 177, No. 3, 425-437.
- Di Micco, R., Fumagalli, M., Cicalese, A., *et al.* (2006). Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature*, Vol. 444, No. 7119, 638-642.
- Diffley, J.F. (2004). Regulation of early events in chromosome replication. *Curr Biol*, Vol. 14, No. 18, 778-786.
- Dimitrova, D.S., Todorov, I.T., Melendy, T. and Gilbert, D.M. (1999). Mcm2, but not RPA, is a component of the mammalian early G1-phase prereplication complex. *J Cell Biol*, Vol. 146, No. 4, 709-722.
- Donovan, S., Harwood, J., Drury, L.S. and Diffley, J.F. (1997). Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *Proc Natl Acad Sci U S A*, Vol. 94, No. 11, 5611-5616.
- Donzelli, M. and Draetta, G.F. (2003). Regulating mammalian checkpoints through Cdc25 inactivation. *EMBO Rep*, Vol. 4, No. 7, 671-677.
- Duursma, A. and Agami, R. (2005). p53-Dependent regulation of Cdc6 protein stability controls cellular proliferation. *Mol Cell Biol*, Vol. 25, No. 16, 6937-6947.

- Edwards, M.C., Tutter, A.V., Cvetic, C., Gilbert, C.H., Prokhorova, T.A. and Walter, J.C. (2002). MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in *Xenopus* egg extracts. *J Biol Chem*, Vol. 277, No. 36, 33049-33057.
- Ekholm-Reed, S., Mendez, J., Tedesco, D., Zetterberg, A., Stillman, B. and Reed, S.I. (2004). Deregulation of cyclin E in human cells interferes with prereplication complex assembly. *J Cell Biol*, Vol. 165, No. 6, 789-800.
- Evrin, C., Clarke, P., Zech, J., *et al.* (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc Natl Acad Sci U S A*, Vol. 106, No. 48, 20240-20245.
- Ferguson, A.M., White, L.S., Donovan, P.J. and Piwnicka-Worms, H. (2005). Normal cell cycle and checkpoint responses in mice and cells lacking Cdc25B and Cdc25C protein phosphatases. *Mol Cell Biol*, Vol. 25, No. 7, 2853-2860.
- Francis, L.I., Randell, J.C., Takara, T.J., Uchima, L. and Bell, S.P. (2009). Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation. *Genes Dev*, Vol. 23, No. 5, 643-654.
- Fujita, M., Yamada, C., Goto, H., *et al.* (1999). Cell cycle regulation of human CDC6 protein. Intracellular localization, interaction with the human mcm complex, and CDC2 kinase-mediated hyperphosphorylation. *J Biol Chem*, Vol. 274, No. 36, 25927-25932.
- Fujita, M., Yamada, C., Tsurumi, T., Hanaoka, F., Matsuzawa, K. and Inagaki, M. (1998). Cell cycle- and chromatin binding state-dependent phosphorylation of human MCM heterohexameric complexes. A role for cdc2 kinase. *J Biol Chem*, Vol. 273, No. 27, 17095-17101.
- Gaudier, M., Schuwirth, B.S., Westcott, S.L. and Wigley, D.B. (2007). Structural basis of DNA replication origin recognition by an ORC protein. *Science*, Vol. 317, No. 5842, 1213-1216.
- Ge, X.Q. and Blow, J.J. (2010). Chk1 inhibits replication factory activation but allows dormant origin firing in existing factories. *J Cell Biol*, Vol. 191, No. 7, 1285-1297.
- Geng, Y., Lee, Y.M., Welcker, M., *et al.* (2007). Kinase-independent function of cyclin E. *Mol Cell*, Vol. 25, No. 1, 127-139.
- Geng, Y., Yu, Q., Sicinska, E., *et al.* (2003). Cyclin E ablation in the mouse. *Cell*, Vol. 114, No. 4, 431-443.
- Gonzalez, S., Klatt, P., Delgado, S., *et al.* (2006). Oncogenic activity of Cdc6 through repression of the INK4/ARF locus. *Nature*, Vol. 440, No. 7084, 702-706.
- Green, B.M., Morreale, R.J., Ozaydin, B., Derisi, J.L. and Li, J.J. (2006). Genome-wide mapping of DNA synthesis in *Saccharomyces cerevisiae* reveals that mechanisms preventing reinitiation of DNA replication are not redundant. *Mol Biol Cell*, Vol. 17, No. 5, 2401-2414.
- Guernsey, D.L., Matsuoka, M., Jiang, H., *et al.* (2011). Mutations in origin recognition complex gene ORC4 cause Meier-Gorlin syndrome. *Nat Genet*, Vol. 43, No. 4, 360-364.
- Guo, Z., Kumagai, A., Wang, S.X. and Dunphy, W.G. (2000). Requirement for Atr in phosphorylation of Chk1 and cell cycle regulation in response to DNA replication blocks and UV-damaged DNA in *Xenopus* egg extracts. *Genes Dev*, Vol. 14, No. 21, 2745-2756.

- Hall, J.R., Kow, E., Nevis, K.R., *et al.* (2007). Cdc6 stability is regulated by the Huwe1 ubiquitin ligase after DNA damage. *Mol Biol Cell*, Vol. 18, No. 9, 3340-3350.
- Hall, J.R., Lee, H.O., Bunker, B.D., *et al.* (2008). CDT1 and CDC6 are destabilized by rereplication-induced DNA damage. *J Biol Chem*, Vol. No. 25356-25363.
- Hartwell, L.H. (1976). Sequential function of gene products relative to DNA synthesis in the yeast cell cycle. *J Mol Biol*, Vol. 104, No. 4, 803-817.
- Hartwell, L.H. and Weinert, T.A. (1989). Checkpoints: controls that ensure the order of cell cycle events. *Science*, Vol. 246, No. 4930, 629-634.
- Herbig, U., Marlar, C.A. and Fanning, E. (1999). The Cdc6 nucleotide-binding site regulates its activity in DNA replication in human cells. *Mol Biol Cell*, Vol. 10, No. 8, 2631-2645.
- Higa, L.A., Banks, D., Wu, M., Kobayashi, R., Sun, H. and Zhang, H. (2006). L2DTL/CDT2 interacts with the CUL4/DDB1 complex and PCNA and regulates CDT1 proteolysis in response to DNA damage. *Cell Cycle*, Vol. 5, No. 15, 1675-1680.
- Higa, L.A., Mihaylov, I.S., Banks, D.P., Zheng, J. and Zhang, H. (2003). Radiation-mediated proteolysis of CDT1 by CUL4-ROC1 and CSN complexes constitutes a new checkpoint. *Nat Cell Biol*, Vol. 5, No. 11, 1008-1015.
- Hofmann, J.F. and Beach, D. (1994). *cdt1* is an essential target of the Cdc10/Sct1 transcription factor: requirement for DNA replication and inhibition of mitosis. *Embo J*, Vol. 13, No. 2, 425-434.
- Hsu, J.Y., Reimann, J.D., Sorensen, C.S., Lukas, J. and Jackson, P.K. (2002). E2F-dependent accumulation of hEmi1 regulates S phase entry by inhibiting APC(Cdh1). *Nat Cell Biol*, Vol. 4, No. 5, 358-366.
- Hu, J., McCall, C.M., Ohta, T. and Xiong, Y. (2004). Targeted ubiquitination of CDT1 by the DDB1-CUL4A-ROC1 ligase in response to DNA damage. *Nat Cell Biol*, Vol. 6, No. 10, 1003-1009.
- Hu, J. and Xiong, Y. (2006). An evolutionarily conserved function of proliferating cell nuclear antigen for Cdt1 degradation by the Cul4-Ddb1 ubiquitin ligase in response to DNA damage. *J Biol Chem*, Vol. 281, No. 7, 3753-3756.
- Huard, S., Elder, R.T., Liang, D., Li, G. and Zhao, R.Y. (2008). Human immunodeficiency virus type 1 Vpr induces cell cycle G2 arrest through Srk1/MK2-mediated phosphorylation of Cdc25. *J Virol*, Vol. 82, No. 6, 2904-2917.
- Iizuka, M., Matsui, T., Takisawa, H. and Smith, M.M. (2006). Regulation of replication licensing by acetyltransferase Hbo1. *Mol Cell Biol*, Vol. 26, No. 3, 1098-1108.
- Iizuka, M., Sarmiento, O.F., Sekiya, T., Scrabble, H., Allis, C.D. and Smith, M.M. (2008). Hbo1 Links p53-dependent stress signaling to DNA replication licensing. *Mol Cell Biol*, Vol. 28, No. 1, 140-153.
- Jiang, W., Wells, N.J. and Hunter, T. (1999). Multistep regulation of DNA replication by Cdk phosphorylation of HsCdc6. *Proc Natl Acad Sci U S A*, Vol. 96, No. 11, 6193-6198.
- Jin, J., Arias, E.E., Chen, J., Harper, J.W. and Walter, J.C. (2006). A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. *Mol Cell*, Vol. 23, No. 5, 709-721.
- Jin, J., Shirogane, T., Xu, L., *et al.* (2003). SCFbeta-TRCP links Chk1 signaling to degradation of the Cdc25A protein phosphatase. *Genes Dev*, Vol. 17, No. 24, 3062-3074.

- Jorgensen, S., Eskildsen, M., Fugger, K., *et al.* (2011). SET8 is degraded via PCNA-coupled CRL4(CDT2) ubiquitylation in S phase and after UV irradiation. *J Cell Biol*, Vol. 192, No. 1, 43-54.
- Kan, Q., Jinno, S., Yamamoto, H., Kobayashi, K. and Okayama, H. (2008). ATP-dependent activation of p21WAF1/CIP1-associated Cdk2 by Cdc6. *Proc Natl Acad Sci USA*, Vol. 105, No. 12, 4757-4762.
- Karakaidos, P., Taraviras, S., Vassiliou, L.V., *et al.* (2004). Overexpression of the replication licensing regulators hCdt1 and hCdc6 characterizes a subset of non-small-cell lung carcinomas: synergistic effect with mutant p53 on tumor growth and chromosomal instability--evidence of E2F-1 transcriptional control over hCdt1. *Am J Pathol*, Vol. 165, No. 4, 1351-1365.
- Kim, J., Feng, H. and Kipreos, E.T. (2007). *C. elegans* CUL-4 prevents rereplication by promoting the nuclear export of CDC-6 via a CKI-1-dependent pathway. *Curr Biol*, Vol. 17, No. 11, 966-972.
- Kim, Y. and Kipreos, E.T. (2007). Cdt1 degradation to prevent DNA re-replication: conserved and non-conserved pathways. *Cell Div*, Vol. 2, No. 18.
- King, R.W., Peters, J.M., Tugendreich, S., Rolfe, M., Hieter, P. and Kirschner, M.W. (1995). A 20S complex containing CDC27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. *Cell*, Vol. 81, No. 2, 279-288.
- Kramer, A., Mailand, N., Lukas, C., *et al.* (2004). Centrosome-associated Chk1 prevents premature activation of cyclin-B-Cdk1 kinase. *Nat Cell Biol*, Vol. 6, No. 9, 884-891.
- Krude, T., Musahl, C., Laskey, R.A. and Knippers, R. (1996). Human replication proteins hCdc21, hCdc46 and P1Mcm3 bind chromatin uniformly before S-phase and are displaced locally during DNA replication. *J Cell Sci*, Vol. 109 ( Pt 2), No. 309-318.
- Lau, E., Zhu, C., Abraham, R.T. and Jiang, W. (2006). The functional role of Cdc6 in S-G2/M in mammalian cells. *EMBO Rep*, Vol. 7, No. 4, 425-430.
- Lee, C., Hong, B., Choi, J.M., *et al.* (2004). Structural basis for inhibition of the replication licensing factor Cdt1 by geminin. *Nature*, Vol. 430, No. 7002, 913-917.
- Lee, K.Y. and Myung, K. (2008). PCNA modifications for regulation of post-replication repair pathways. *Mol Cells*, Vol. 26, No. 1, 5-11.
- Lei, M., Kawasaki, Y. and Tye, B.K. (1996). Physical interactions among Mcm proteins and effects of Mcm dosage on DNA replication in *Saccharomyces cerevisiae*. *Mol Cell Biol*, Vol. 16, No. 9, 5081-5090.
- Lemaire, M., Froment, C., Boutros, R., *et al.* (2006). CDC25B phosphorylation by p38 and MK-2. *Cell Cycle*, Vol. 5, No. 15, 1649-1653.
- Levine, A.J. (1997). p53, the cellular gatekeeper for growth and division. *Cell*, Vol. 88, No. 3, 323-331.
- Li, C.J. and Depamphilis, M.L. (2002). Mammalian Orc1 protein is selectively released from chromatin and ubiquitinated during the S-to-M transition in the cell division cycle. *Mol Cell Biol*, Vol. 22, No. 1, 105-116.
- Li, C.J., Vassilev, A. and Depamphilis, M.L. (2004). Role for Cdk1 (Cdc2)/cyclin A in preventing the mammalian origin recognition complex's largest subunit (Orc1) from binding to chromatin during mitosis. *Mol Cell Biol*, Vol. 24, No. 13, 5875-5886.
- Lin, D.I., Aggarwal, P. and Diehl, J.A. (2008). Phosphorylation of MCM3 on Ser-112 regulates its incorporation into the MCM2-7 complex. *Proc Natl Acad Sci USA*, Vol. 105, No. 23, 8079-8084.

- Lin, J.J. and Dutta, A. (2007). ATR pathway is the primary pathway for activating G2/M checkpoint induction after re-replication. *J Biol Chem*, Vol. 282, No. 42, 30357-30362.
- Lin, J.J., Milhollen, M.A., Smith, P.G., Narayanan, U. and Dutta, A. (2010). NEDD8-targeting drug MLN4924 elicits DNA rereplication by stabilizing Cdt1 in S phase, triggering checkpoint activation, apoptosis, and senescence in cancer cells. *Cancer Res*, Vol. 70, No. 24, 10310-10320.
- Liontos, M., Koutsami, M., Sideridou, M., *et al.* (2007). Deregulated overexpression of hCdt1 and hCdc6 promotes malignant behavior. *Cancer Res*, Vol. 67, No. 22, 10899-10909.
- Liu, E., Li, X., Yan, F., Zhao, Q. and Wu, X. (2004). Cyclin-dependent kinases phosphorylate human Cdt1 and induce its degradation. *J Biol Chem*, Vol. 279, No. 17, 17283-17288.
- Liu, J., Smith, C.L., Deryckere, D., Deangelis, K., Martin, G.S. and Berger, J.M. (2000). Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Mol Cell*, Vol. 6, No. 3, 637-648.
- Lovejoy, C.A., Lock, K., Yenamandra, A. and Cortez, D. (2006). DDB1 maintains genome integrity through regulation of Cdt1. *Mol Cell Biol*, Vol. 26, No. 21, 7977-7990.
- Lutzmann, M., Maiorano, D. and Mechali, M. (2006). A Cdt1-geminin complex licenses chromatin for DNA replication and prevents rereplication during S phase in *Xenopus*. *Embo J*, Vol. 25, No. 24, 5764-5774.
- Lutzmann, M. and Mechali, M. (2008). MCM9 binds Cdt1 and is required for the assembly of prereplication complexes. *Mol Cell*, Vol. 31, No. 2, 190-200.
- Lutzmann, M. and Mechali, M. (2009). How to load a replicative helicase onto chromatin: a more and more complex matter during evolution. *Cell Cycle*, Vol. 8, No. 9, 1309-1313.
- Machida, Y.J. and Dutta, A. (2007). The APC/C inhibitor, Emi1, is essential for prevention of rereplication. *Genes Dev*, Vol. 21, No. 2, 184-194.
- Maga, G. and Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *J Cell Sci*, Vol. 116, No. Pt 15, 3051-3060.
- Mailand, N. and Diffley, J.F. (2005). CDKs promote DNA replication origin licensing in human cells by protecting Cdc6 from APC/C-dependent proteolysis. *Cell*, Vol. 122, No. 6, 915-926.
- Maiorano, D., Moreau, J. and Mechali, M. (2000). XCDT1 is required for the assembly of pre-replicative complexes in *Xenopus laevis*. *Nature*, Vol. 404, No. 6778, 622-625.
- Malumbres, M. and Barbacid, M. (2009). Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer*, Vol. 9, No. 3, 153-166.
- Manke, I.A., Nguyen, A., Lim, D., Stewart, M.Q., Elia, A.E. and Yaffe, M.B. (2005). MAPKAP kinase-2 is a cell cycle checkpoint kinase that regulates the G2/M transition and S phase progression in response to UV irradiation. *Mol Cell*, Vol. 17, No. 1, 37-48.
- Masai, H., Matsui, E., You, Z., Ishimi, Y., Tamai, K. and Arai, K. (2000). Human Cdc7-related kinase complex. In vitro phosphorylation of MCM by concerted actions of Cdk5 and Cdc7 and that of a critical threonine residue of Cdc7 by Cdk5. *J Biol Chem*, Vol. 275, No. 37, 29042-29052.
- Matsuoka, S., Ballif, B.A., Smogorzewska, A., *et al.* (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*, Vol. 316, No. 5828, 1160-1166.
- Mcgarry, T.J. (2002). Geminin deficiency causes a Chk1-dependent G2 arrest in *Xenopus*. *Mol Biol Cell*, Vol. 13, No. 10, 3662-3671.

- McGarry, T.J. and Kirschner, M.W. (1998). Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell*, Vol. 93, No. 6, 1043-1053.
- Meek, D.W. (1994). Post-translational modification of p53. *Semin Cancer Biol*, Vol. 5, No. 3, 203-210.
- Melchionna, R., Chen, X.B., Blasina, A. and McGowan, C.H. (2000). Threonine 68 is required for radiation-induced phosphorylation and activation of Cds1. *Nat Cell Biol*, Vol. 2, No. 10, 762-765.
- Melixetian, M., Ballabeni, A., Masiero, L., et al. (2004). Loss of Geminin induces rereplication in the presence of functional p53. *J Cell Biol*, Vol. 165, No. 4, 473-482.
- Mendez, J. and Stillman, B. (2000). Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol Cell Biol*, Vol. 20, No. 22, 8602-8612.
- Mendez, J., Zou-Yang, X.H., Kim, S.Y., Hidaka, M., Tansey, W.P. and Stillman, B. (2002). Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. *Mol Cell*, Vol. 9, No. 3, 481-491.
- Mihaylov, I.S., Kondo, T., Jones, L., et al. (2002). Control of DNA replication and chromosome ploidy by geminin and cyclin A. *Mol Cell Biol*, Vol. 22, No. 6, 1868-1880.
- Milczarek, G.J., Martinez, J. and Bowden, G.T. (1997). p53 Phosphorylation: biochemical and functional consequences. *Life Sci*, Vol. 60, No. 1, 1-11.
- Miotto, B. and Struhl, K. (2008). HBO1 histone acetylase is a coactivator of the replication licensing factor Cdt1. *Genes Dev*, Vol. 22, No. 19, 2633-2638.
- Miotto, B. and Struhl, K. (2010). HBO1 histone acetylase activity is essential for DNA replication licensing and inhibited by Geminin. *Mol Cell*, Vol. 37, No. 1, 57-66.
- Mizushima, T., Takahashi, N. and Stillman, B. (2000). Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro. *Genes Dev*, Vol. 14, No. 13, 1631-1641.
- Moshe, Y., Boulaire, J., Pagano, M. and Hershko, A. (2004). Role of Polo-like kinase in the degradation of early mitotic inhibitor 1, a regulator of the anaphase promoting complex/cyclosome. *Proc Natl Acad Sci U S A*, Vol. 101, No. 21, 7937-7942.
- Nasmyth, K. (1996). Viewpoint: putting the cell cycle in order. *Science*, Vol. 274, No. 5293, 1643-1645.
- Nguyen, V.Q., Co, C. and Li, J.J. (2001). Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature*, Vol. 411, No. 6841, 1068-1073.
- Nielsen, N.H., Arnerlov, C., Emdin, S.O. and Landberg, G. (1996). Cyclin E overexpression, a negative prognostic factor in breast cancer with strong correlation to oestrogen receptor status. *Br J Cancer*, Vol. 74, No. 6, 874-880.
- Niida, H., Katsuno, Y., Banerjee, B., Hande, M.P. and Nakanishi, M. (2007). Specific role of Chk1 phosphorylations in cell survival and checkpoint activation. *Mol Cell Biol*, Vol. 27, No. 7, 2572-2581.
- Nishitani, H., Lygerou, Z., Nishimoto, T. and Nurse, P. (2000). The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature*, Vol. 404, No. 6778, 625-628.
- Nishitani, H. and Nurse, P. (1995). p65cdc18 plays a major role controlling the initiation of DNA replication in fission yeast. *Cell*, Vol. 83, No. 3, 397-405.

- Nishitani, H., Sugimoto, N., Roukos, V., *et al.* (2006). Two E3 ubiquitin ligases, SCF-Skp2 and DDB1-Cul4, target human Cdt1 for proteolysis. *Embo J*, Vol. 25, No. 5, 1126-1136.
- Oda, H., Hubner, M.R., Beck, D.B., *et al.* (2010). Regulation of the histone H4 monomethylase PR-Set7 by CRL4(Cdt2)-mediated PCNA-dependent degradation during DNA damage. *Mol Cell*, Vol. 40, No. 3, 364-376.
- Ode, K.L., Fujimoto, K., Kubota, Y. and Takisawa, H. (2011). Inter-origin cooperativity of geminin action establishes an all-or-none switch for replication origin licensing. *Genes Cells*, Vol. 16, No. 4, 380-396.
- Ohta, S., Koide, M., Tokuyama, T., Yokota, N., Nishizawa, S. and Namba, H. (2001). Cdc6 expression as a marker of proliferative activity in brain tumors. *Oncol Rep*, Vol. 8, No. 5, 1063-1066.
- Paolinelli, R., Mendoza-Maldonado, R., Cereseto, A. and Giacca, M. (2009). Acetylation by GCN5 regulates CDC6 phosphorylation in the S phase of the cell cycle. *Nat Struct Mol Biol*, Vol. 16, No. 4, 412-420.
- Petersen, B.O., Lukas, J., Sorensen, C.S., Bartek, J. and Helin, K. (1999). Phosphorylation of mammalian CDC6 by cyclin A/CDK2 regulates its subcellular localization. *Embo J*, Vol. 18, No. 2, 396-410.
- Petersen, B.O., Wagener, C., Marinoni, F., *et al.* (2000). Cell cycle- and cell growth-regulated proteolysis of mammalian CDC6 is dependent on APC-CDH1. *Genes Dev*, Vol. 14, No. 18, 2330-2343.
- Piatti, S., Lengauer, C. and Nasmyth, K. (1995). Cdc6 is an unstable protein whose de novo synthesis in G1 is important for the onset of S phase and for preventing a 'reductional' anaphase in the budding yeast *Saccharomyces cerevisiae*. *Embo J*, Vol. 14, No. 15, 3788-3799.
- Pinyol, M., Salaverria, I., Bea, S., *et al.* (2006). Unbalanced expression of licensing DNA replication factors occurs in a subset of mantle cell lymphomas with genomic instability. *Int J Cancer*, Vol. 119, No. 12, 2768-2774.
- Ralph, E., Boye, E. and Kearsley, S.E. (2006). DNA damage induces Cdt1 proteolysis in fission yeast through a pathway dependent on Cdt2 and Ddb1. *EMBO Rep*, Vol. 7, No. 11, 1134-1113.
- Randell, J.C., Bowers, J.L., Rodriguez, H.K. and Bell, S.P. (2006). Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Mol Cell*, Vol. 21, No. 1, 29-39.
- Rape, M., Reddy, S.K. and Kirschner, M.W. (2006). The processivity of multiubiquitination by the APC determines the order of substrate degradation. *Cell*, Vol. 124, No. 1, 89-103.
- Reinhardt, H.C., Aslanian, A.S., Lees, J.A. and Yaffe, M.B. (2007). p53-deficient cells rely on ATM- and ATR-mediated checkpoint signaling through the p38MAPK/MK2 pathway for survival after DNA damage. *Cancer Cell*, Vol. 11, No. 2, 175-189.
- Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P. and Diffley, J.F. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell*, Vol. 139, No. 4, 719-730.
- Rowles, A., Tada, S. and Blow, J.J. (1999). Changes in association of the *Xenopus* origin recognition complex with chromatin on licensing of replication origins. *J Cell Sci*, Vol. 112 ( Pt 12), No. 2011-2018.

- Saha, P., Chen, J., Thome, K.C., *et al.* (1998). Human CDC6/Cdc18 associates with Orc1 and cyclin-cdk and is selectively eliminated from the nucleus at the onset of S phase. *Mol Cell Biol*, Vol. 18, No. 5, 2758-2767.
- Saha, T., Ghosh, S., Vassilev, A. and Depamphilis, M.L. (2006). Ubiquitylation, phosphorylation and Orc2 modulate the subcellular location of Orc1 and prevent it from inducing apoptosis. *J Cell Sci*, Vol. 119, No. Pt 7, 1371-1382.
- Sansam, C.L., Shepard, J.L., Lai, K., *et al.* (2006). DTL/CDT2 is essential for both CDT1 regulation and the early G2/M checkpoint. *Genes Dev*, Vol. 20, No. 22, 3117-3129.
- Senga, T., Sivaprasad, U., Zhu, W., *et al.* (2006). PCNA is a cofactor for Cdt1 degradation by CUL4/DDB1-mediated N-terminal ubiquitination. *J Biol Chem*, Vol. 281, No. 10, 6246-6252.
- Shechter, D., Costanzo, V. and Gautier, J. (2004). ATR and ATM regulate the timing of DNA replication origin firing. *Nat Cell Biol*, Vol. 6, No. 7, 648-655.
- Sheu, Y.J. and Stillman, B. (2006). Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Mol Cell*, Vol. 24, No. 1, 101-113.
- Shiloh, Y. (2003). ATM and related protein kinases: safeguarding genome integrity. *Nat Rev Cancer*, Vol. 3, No. 3, 155-168.
- Sivaprasad, U., Machida, Y.J. and Dutta, A. (2007). APC/C--the master controller of origin licensing? *Cell Div*, Vol. 2, No. 8.
- Smith, M.L., Chen, I.T., Zhan, Q., O'connor, P.M. and Fornace, A.J., Jr. (1995). Involvement of the p53 tumor suppressor in repair of u.v.-type DNA damage. *Oncogene*, Vol. 10, No. 6, 1053-1059.
- Speck, C., Chen, Z., Li, H. and Stillman, B. (2005). ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nat Struct Mol Biol*, Vol. 12, No. 11, 965-971.
- Sudakin, V., Ganoth, D., Dahan, A., *et al.* (1995). The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis. *Mol Biol Cell*, Vol. 6, No. 2, 185-197.
- Sudo, T., Ota, Y., Kotani, S., *et al.* (2001). Activation of Cdh1-dependent APC is required for G1 cell cycle arrest and DNA damage-induced G2 checkpoint in vertebrate cells. *Embo J*, Vol. 20, No. 22, 6499-6508.
- Sugimoto, N., Tatsumi, Y., Tsurumi, T., *et al.* (2004). Cdt1 phosphorylation by cyclin A-dependent kinases negatively regulates its function without affecting geminin binding. *J Biol Chem*, Vol. 279, No. 19, 19691-19697.
- Tada, S., Li, A., Maiorano, D., Mechali, M. and Blow, J.J. (2001). Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. *Nat Cell Biol*, Vol. 3, No. 2, 107-113.
- Takeda, D.Y., Parvin, J.D. and Dutta, A. (2005). Degradation of Cdt1 during S phase is Skp2-independent and is required for efficient progression of mammalian cells through S phase. *J Biol Chem*, Vol. 280, No. 24, 23416-23423.
- Tanaka, S. and Diffley, J.F. (2002). Dereglated G1-cyclin expression induces genomic instability by preventing efficient pre-RC formation. *Genes Dev*, Vol. 16, No. 20, 2639-2649.
- Tang, J., Erikson, R.L. and Liu, X. (2006). Checkpoint kinase 1 (Chk1) is required for mitotic progression through negative regulation of polo-like kinase 1 (Plk1). *Proc Natl Acad Sci U S A*, Vol. 103, No. 32, 11964-11969.

- Tanny, R.E., Macalpine, D.M., Blitzblau, H.G. and Bell, S.P. (2006). Genome-wide analysis of re-replication reveals inhibitory controls that target multiple stages of replication initiation. *Mol Biol Cell*, Vol. 17, No. 5, 2415-2423.
- Tardat, M., Brustel, J., Kirsh, O., *et al.* (2010). The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells. *Nat Cell Biol*, Vol. No.
- Taylor, W.R. and Stark, G.R. (2001). Regulation of the G2/M transition by p53. *Oncogene*, Vol. 20, No. 15, 1803-1815.
- Tsuyama, T., Tada, S., Watanabe, S., Seki, M. and Enomoto, T. (2005). Licensing for DNA replication requires a strict sequential assembly of Cdc6 and Cdt1 onto chromatin in *Xenopus* egg extracts. *Nucleic Acids Res*, Vol. 33, No. 2, 765-775.
- Van Den Heuvel, S. and Harlow, E. (1993). Distinct roles for cyclin-dependent kinases in cell cycle control. *Science*, Vol. 262, No. 5142, 2050-2054.
- Vashee, S., Simancek, P., Challberg, M.D. and Kelly, T.J. (2001). Assembly of the human origin recognition complex. *J Biol Chem*, Vol. 276, No. 28, 26666-26673.
- Vaziri, C., Saxena, S., Jeon, Y., *et al.* (2003). A p53-dependent checkpoint pathway prevents rereplication. *Mol Cell*, Vol. 11, No. 4, 997-1008.
- Vogelstein, B., Lane, D. and Levine, A.J. (2000). Surfing the p53 network. *Nature*, Vol. 408, No. 6810, 307-310.
- Wei, W., Ayad, N.G., Wan, Y., Zhang, G.J., Kirschner, M.W. and Kaelin, W.G., Jr. (2004). Degradation of the SCF component Skp2 in cell-cycle phase G1 by the anaphase-promoting complex. *Nature*, Vol. 428, No. 6979, 194-198.
- Weinert, T.A. and Hartwell, L.H. (1990). Characterization of RAD9 of *Saccharomyces cerevisiae* and evidence that its function acts posttranslationally in cell cycle arrest after DNA damage. *Mol Cell Biol*, Vol. 10, No. 12, 6554-6564.
- Weinreich, M., Liang, C. and Stillman, B. (1999). The Cdc6p nucleotide-binding motif is required for loading mcm proteins onto chromatin. *Proc Natl Acad Sci USA*, Vol. 96, No. 2, 441-446.
- Wheeler, L.W., Lents, N.H. and Baldassare, J.J. (2008). Cyclin A-CDK activity during G1 phase impairs MCM chromatin loading and inhibits DNA synthesis in mammalian cells. *Cell Cycle*, Vol. 7, No. 14, 2179-2188.
- Wilmes, G.M., Archambault, V., Austin, R.J., Jacobson, M.D., Bell, S.P. and Cross, F.R. (2004). Interaction of the S-phase cyclin Clb5 with an "RXL" docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. *Genes Dev*, Vol. 18, No. 9, 981-991.
- Wohlschlegel, J.A., Dwyer, B.T., Dhar, S.K., Cvetic, C., Walter, J.C. and Dutta, A. (2000). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science*, Vol. 290, No. 5500, 2309-23012.
- Wu, S. and Rice, J.C. (2010). A new regulator of the cell cycle: the PR-Set7 histone methyltransferase. *Cell Cycle*, Vol. 10, No. 1, 68-72.
- Xouri, G., Lygerou, Z., Nishitani, H., Pachnis, V., Nurse, P. and Taraviras, S. (2004). Cdt1 and geminin are down-regulated upon cell cycle exit and are over-expressed in cancer-derived cell lines. *Eur J Biochem*, Vol. 271, No. 16, 3368-3378.
- Yoo, H.Y., Shevchenko, A. and Dunphy, W.G. (2004). Mcm2 is a direct substrate of ATM and ATR during DNA damage and DNA replication checkpoint responses. *J Biol Chem*, Vol. 279, No. 51, 53353-53364.

- Yoshida, K., Sugimoto, N., Iwahori, S., *et al.* (2010). CDC6 interaction with ATR regulates activation of a replication checkpoint in higher eukaryotic cells. *J Cell Sci*, Vol. 123, No. Pt 2, 225-235.
- Zachariae, W., Schwab, M., Nasmyth, K. and Seufert, W. (1998). Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. *Science*, Vol. 282, No. 5394, 1721-1724.
- Zhao, H. and Piwnicka-Worms, H. (2001). ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1. *Mol Cell Biol*, Vol. 21, No. 13, 4129-4139.
- Zhong, W., Feng, H., Santiago, F.E. and Kipreos, E.T. (2003). CUL-4 ubiquitin ligase maintains genome stability by restraining DNA-replication licensing. *Nature*, Vol. 423, No. 6942, 885-889.
- Zhou, B.B. and Elledge, S.J. (2000). The DNA damage response: putting checkpoints in perspective. *Nature*, Vol. 408, No. 6811, 433-439.
- Zhu, W., Chen, Y. and Dutta, A. (2004). Rereplication by depletion of geminin is seen regardless of p53 status and activates a G2/M checkpoint. *Mol Cell Biol*, Vol. 24, No. 16, 7140-7150.
- Zhu, W. and Dutta, A. (2006). An ATR- and BRCA1-mediated Fanconi anemia pathway is required for activating the G2/M checkpoint and DNA damage repair upon rereplication. *Mol Cell Biol*, Vol. 26, No. 12, 4601-4611.
- Zou, L. and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science*, Vol. 300, No. 5625, 1542-1548.
- Zwerschke, W., Rottjakob, H.W. and Kuntzel, H. (1994). The *Saccharomyces cerevisiae* CDC6 gene is transcribed at late mitosis and encodes a ATP/GTPase controlling S phase initiation. *J Biol Chem*, Vol. 269, No. 37, 23351-23356.

# Addressing the Enigma of MCM8 in DNA Replication

Dianne C. Daniel and Edward M. Johnson  
*Eastern Virginia Medical School  
United States of America*

## 1. Introduction

MCM8 is a relatively new protein about which little is currently known regarding its function within the cell. Even so, there is controversy surrounding what its role might be in DNA replication. In this chapter, information regarding the role of MCM8 in DNA replication gleaned from studies carried out in different species will be discussed. In addition, sequence differences in the protein, itself, among different species will be presented. This review will focus on MCM8 in three species, *Homo sapiens*, *Xenopus laevis* and *Drosophila melanogaster*, since the bulk of published experimental data was obtained using these organisms.

## 2. Structure of MCM8

The discovery of the MCM8 gene was the direct result of sequencing of the human genome. The MCM8 gene was discovered based on direct search for homologies to the genes of other family members (Johnson et al., 2003) and on direct database comparison of sequences in which a hepatitis virus had integrated in DNA from human hepatocellular carcinoma tissue (Gozuacik et al., 2003). MCM8 belongs to the Minichromosome Maintenance (MCM) family of proteins so named based on the role of founding members in the maintenance of centromeric plasmids containing an origin of replication in *Saccharomyces cerevisiae* (Maine et al., 1984). MCM family members other than MCM8 will be discussed in another chapter and will not be further discussed here except for comparative purposes. Two groups simultaneously named the newly discovered protein and gene MCM8, essentially because the names MCM1, MCM2-7 and MCM10 were already taken (Gozuacik et al., 2003; Johnson et al., 2003). MCM 1 and MCM10 are not homologous in sequence to MCM2-7 family members (Tye, 1999). Whereas MCM2-7 family members have a secured position within the mechanistic framework of DNA replication based on studies in both yeast and *Xenopus* (Blow & Laskey, 1988; Chong et al., 1995; Forsburg, 2004; Gomez et al., 2002; Hennessy et al., 1991; Kubota et al., 1995; Labib et al., 2000; Thommes et al., 1997; Tye, 1999; Yan et al., 1993), the role of MCM8 is enigmatic.

In *Homo sapiens*, the MCM8 gene is located at chromosome band 20p12.3-13 and is located contrapodal to a gene encoding a homolog of the yeast GCD10 gene. It is composed of 19 exons (Johnson et al., 2003). MCM8 will be italicized hereafter when specifically referring to nucleic acid. *MCM8* also exists in a variant form (isoform 2) in a choriocarcinoma found

among Expressed Sequence Tags, which could be the result of aberrant splicing (Johnson et al., 2003). Both MCM8 and MCM8 isoform 2 have canonical helicase domains, Walker A and Walker B (Walker et al., 1982). Although other family members have an unusual A-box motif, either AKS or SKS, MCM8 has the canonical A-box motif GKS, indicating the possibility of intrinsic helicase activity (Johnson et al., 2003). Sequences deleted in the MCM8 isoform 2 are near a zinc finger-like motif located N-terminal to the Walker boxes (Fig. 1). The reported *Xenopus laevis* MCM8 homolog contains similar structural motifs, including a Zn finger-type motif, highly homologous to the human one, and Walker A and Walker B motifs. It also contains the canonical Walker A GKS sequence described above for human MCM8 (see Fig. 2). *Xenopus* MCM8 is highly homologous to human MCM8 (74%), with most variability in the N terminus. This variability is in a region of 60 amino acids that are arginine- and glycine-rich in both species (Maiorano et al., 2005).

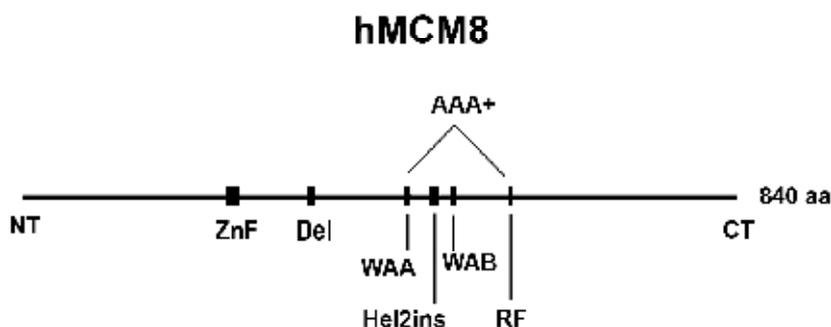


Fig. 1. Diagram showing the conserved structural features for human MCM8. Lengths and positions of the noted motifs are approximately to scale. WAA and WAB are the Walker A and Walker B boxes, respectively (Walker et al., 1982). Hel2ins is the helix-2 insert (Jenkinson & Chong, 2006) and ZnF and RF refer to the zinc finger-like motif and arginine finger motif, respectively (Forsburg, 2004; Tye, 1999). Del refers to the sixteen amino acids deleted in human MCM8 isoform 2 (Johnson et al., 2003).

The *rec* gene in *Drosophila melanogaster* was later reported to be the MCM8 ortholog, and it is located on *Drosophila* Chromosome 3 between *c(3)G* and *spn-E*. The *rec* gene consists of two exons and an intron. Note that *rec* will indicate the gene and REC will indicate the protein. It may have diverged from other MCM8 orthologs through accumulation of more changes as indicated by its longer branch length in the phylogenetic tree compared with other MCMs (Blanton et al., 2005). Based on ancient diversification, MCM8 is concordantly either present or absent along with another family member, MCM9, in various taxa. Exceptions are the *Drosophila* species, which lack MCM9. The sequence divergence of REC is suggested to relate to the lack of MCM9 in this species or to a major change in protein function (Blanton et al., 2005; Liu et al., 2009). MCM9 is not included in all compilations of MCM proteins because MCM9 lacks the signature Walker B box sequence, IDEFDKM, present in all other MCM members. Human MCM9 contains instead the somewhat conserved IDEFNSL (Shultz et al., 2007). Compared to other MCM family members, MCM9 proteins have a longer and poorly conserved C-terminus (Lutzmann et al., 2005; Shultz et al., 2007). From an evolutionary standpoint, the highly divergent MCM8 ortholog in all species of *Drosophila* may not make it a good model for other eukaryotes (Liu et al., 2009).

MCM8 is not present in yeast where family members MCM2-7 have been extensively studied using genetic approaches (Tye, 1999). The archaeal MCM proteins and MCM2-7 proteins are AAA+ (ATPases associated with a variety of cellular activities) enzymes (Bae et al., 2009; Bochman & Schwacha, 2009, review; Koonin, 1993; Neuwald et al., 1999). The ATPase active site is known as the AAA+ domain (Neuwald et al., 1999). Subdomains of AAA+ proteins contain ATPase active site motifs found in a P loop domain and a lid domain C-terminal to the P loop. The P loop domain consists of Walker A and Walker B boxes and a Sensor I motif. The lid contains the arginine finger motif and a Sensor II motif. In MCM proteins, the AAA+ active sites are formed from the P loop *cis* motif of one subunit and the lid domain *trans* motif of another subunit. Nucleotide binding and hydrolysis can lead to conformational changes that drive mechanical work (Bochman & Schwacha, 2009; Erzberger & Berger, 2006). Conformational changes among subunits of toroidal hexamers, such as the MCM2-7 complex, may be coupled and propagated by the combinatorial arrangement of active sites (Davey et al., 2003; Bochman & Schwacha, 2009).

A large presensor-I (PS-I) AAA+ superclade is comprised of clade 4 viral superfamily III helicases (e.g., SV40 large T-antigen), the clade 5 HCLR (HslU, ClpX, Lon and Ruv) and the clade 6 H2 insert family in which MCM family members were initially placed (Erzberger & Berger, 2006). The clade 6 H2 insert family has a beta-alpha-beta insertion in helix 2 (Iyer et al., 2004; Jenkinson & Chong, 2006). As in MCM2-7, Fig. 2 sequence alignment shows that the helix-2 insert (Hel2ins) is also present between the Walker A and B boxes in MCM8 in *Homo sapiens*, *Xenopus laevis* and *Drosophila melanogaster*. In addition to the presence of an insertion in helix 2 of clade 6 members, an additional insertion has been reported to define Clade 7 (Erzberger & Berger, 2006). Based on structural studies, this group reported that MCM2-7 members contain a Sensor-II motif within the lid domain that is uniquely repositioned to act in *trans* through a helical insertion positioned N-terminal to it (presensor-II). They used this insert to define a presensor-II (PS-II) insertion clade 7, which includes MCM2-7. The divergent members reported for clade 7 were chosen based on the presence of the helical insertion when the H2 insert family members were examined. Whereas clade 6 H2 insert mutations disrupted interaction of clade 6 member NtrC with its remodeling target, the clade 7 helical insertion is thought to coordinate the stability of adjacent subunits or protomers, perhaps by changing contacts or accessible surface area (Erzberger & Berger, 2006). Our sequence analyses reveal that *Homo sapiens*, *Xenopus laevis* and *Drosophila melanogaster* MCM8 contain a potential helical insertion in PS-II. Further structural information about this MCM8 insert is not available, but it could potentially place MCM8 in AAA+ clade 7. Notably, this region appears much longer in *Drosophila* than in the other two species based on these analyses (not shown). This region in all three species requires further analysis, since it has been described as potentially important in subunit interactions (Erzberger & Berger, 2006) and could help account for active-site differences within MCM2-7 subunits (Bochman & Schwacha, 2009).

### 3. Sequence variations of MCM8 proteins among and within different species

The following figures show important sequence alignments for the three species to be discussed in this review: *Hs1* is *Homo sapiens* isoform 1, *Hs2* is *Homo sapiens* isoform 2 (Johnson et al., 2003), *Xl* is *Xenopus laevis*, and *Dm* is *Drosophila melanogaster*. All sequence alignments in the figures in Section 3 were performed using the Multalin algorithm (Corpet, 1988).

**Fig. 2** is an alignment of the conserved Walker A and Walker B ATPase domains, the AAA+ clade 6 helix-2 insert that is present in MCM2-7 and the canonical MCM arginine finger

domain (SRFD). The conserved motifs are part of the AAA+ domain of many ATP-dependent molecular motors (Koonin, 1993; Neuwald et al., 1999). Note that there is significant homology in three of these four motifs among the three species. Only the helix-2 insert of *Drosophila* MCM8 is completely divergent.

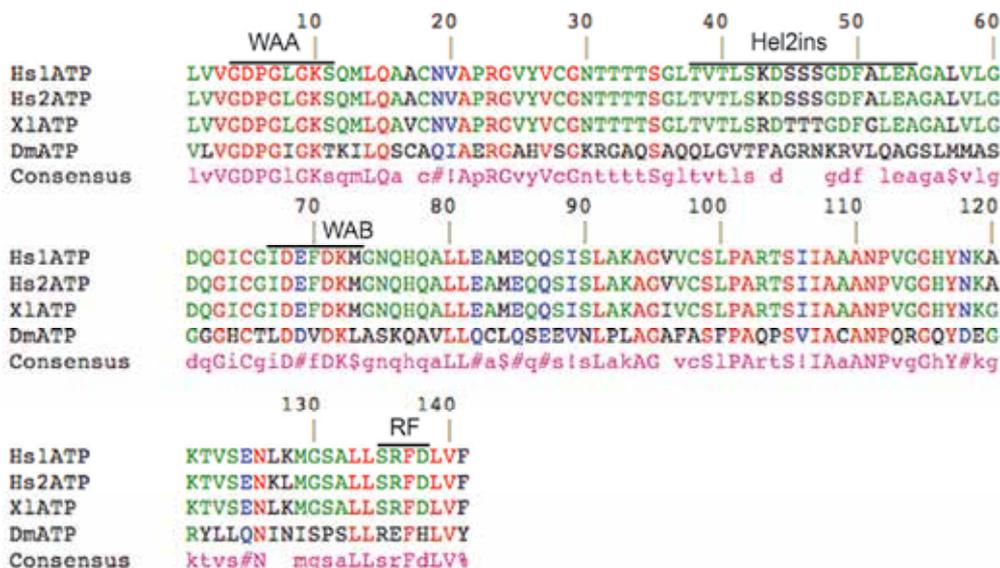


Fig. 2. The Mg-ATP binding domain of MCM8 from three species. Comparative alignment of MCM8 conserved motifs in *Hs* (forms 1 and 2), *Xl* and *Dm*. Abbreviations: WAA (Walker box A), WAB (Walker box B), RF (arginine finger), Hel2ins (helix-2 insert).

The Walker box A is important for P loop conformation and is implicated in ATP binding by interacting with the phosphate moiety (Saraste et al., 1990; Walker et al., 1982). The Walker B motif is thought to contribute to ATP hydrolysis by mediating interaction with ATP through Mg<sup>2+</sup>. (Koonin, 1993; Walker et al., 1982). The arginine finger is the SRFD amino acid motif found in other MCM2-7 family members (Forsburg, 2004). At the AAA+ active site, the arginine finger is in a different subunit from the one that binds ATP and helps to complete the active site interface between two subunits. It may help coordinate the order of hydrolysis in the MCM2-7 oligomer (Bae et al., 2009; Davey et al., 2003; Davey et al., 2002). The helix-2 insert is implicated in helicase activity (Brewster et al., 2008; Jenkinson & Chong, 2006). The helix-2 insert domain plays a key role in transducing or coupling energy of hydrolysis to unwinding of target. Removal of the helix-2 insert in archaeon *Methanothermobacter thermautotrophicus* (*MthMCM*) resulted in the loss of DNA unwinding and increased dsDNA-dependent ATP hydrolysis and the affinity for single-stranded and double-stranded DNA. Since this motif is not as conserved as other neighboring AAA+ motifs, its role in MCM activity was predicted to be mechanical (Jenkinson & Chong, 2006). In the Fig. 3, the zinc finger-like motif, a common feature of other MCM family members (Forsburg, 2004; Tye, 1999) is presented for the three species under discussion. Note that this domain in *Drosophila melanogaster* is almost completely divergent from the domain in *Homo sapiens* isoform 1, *Homo sapiens* isoform 2 and *Xenopus laevis*. Only a P residue and the Zn-chelating four C residues are conserved in *Drosophila* MCM8.

	10	20
Hs1Zn	CAACGEIQSFPLPDGKYS	LP TKCPVPVC
Hs2Zn	CAACGEIQSFPLPDGKYS	LP TKCPVPVC
XlZn	CNMCGLIQCFPLPDGKY	TVP TKCPVPEC
DmZn	CSRCQMEIAMRQRGTFQ	PRPYQCKRSEC
Consensus	C Cg iq fplpdgky	PtkCpvpeC

Fig. 3. Homologies among zinc finger domains in three species.

The Zn<sup>2+</sup> finger motif is not involved directly in DNA binding. It stabilizes folding of the N-terminal domain, and this function is predicted to aid in the formation of the double hexamer in the *Mth*MCM archaeon (Fletcher et al., 2003). For discussion of this motif in eukaryotic MCMs, see the review by Bochman and Schwacha, 2009.

In Fig. 4, the region of *Homo sapiens* isoform 1 MCM8 found to be deleted in choriocarcinoma is aligned to compare the two *Homo sapiens* isoforms with *Xenopus laevis* and *Drosophila melanogaster*. Note that *Xl* and *Hs1* are highly homologous in the region of the deletion. The deletion occurs in the *Hs2* variant, which has thus far only been reported in a case of choriocarcinoma (Johnson et al., 2003). A similar region is missing in *Dm* MCM8. This deletion most likely does not eliminate DNA binding because *Dm* MCM8 is reportedly involved in DNA repair synthesis in meiosis (Matsubayashi & Yamamoto, 2003; Blanton et al., 2005).

	10	20	30	40	50	60
XlMCM8	VDSCVPCDMI	TVTGIVKVS	NTRDGGFKNKNNKCMFLLYIEAN	-VNSKQGQK	KSTEDSGN	
HsMCM81	VDSCVPGDTV	ITIGIVKVS	NAEEGS-RNKNDKCMFLLYIEANS	ISNSKQGQK	TKSSE	DGCK
HsMCM82	VDSCVPGDTV	ITIGIVKVS	NAE-----	-----EANS	ISNSKQGQK	TKSSE
DmMCM8	VDAVRVCQEVV	VTGILKLQELGDDT	-----	-----TTGDT	SNQMQLKAVS	IRDA
Consensus	VDscvpG#	It!TGIVkvs#	d		ean	SNskgqk Ks ed

Fig. 4. A deletion in *Hs* MCM8 isoform 2 is also present in *Dm* MCM8.

Based on archaeon *Mth*MCM mutational analysis, N- and C-terminal sequences were found to play a regulatory role in ATP hydrolysis with effects on substrate binding and on processivity (Jenkinson & Chong, 2006). Whether this particular region of deletion in *Hs* isoform 2 has similar regulatory effects remains to be determined.

#### 4. Research regarding MCM8 protein function in DNA replication in *Homo sapiens* and *Xenopus laevis*

A report on MCM8 from human studies resulted from the integration of hepatitis B virus at a site that was subsequently identified as being the sequence coding for MCM8 (Gozuacik et al., 2003). The group developed a rabbit polyclonal antibody against the N-terminus of the protein and cloned the MCM8 gene using RT-PCR. Further investigation was carried out using cultured cells for *in vitro* studies. Following the release of density arrested Hs27 newborn foreskin fibroblasts, MCM8 mRNA accumulated from G1 through S phase. The MCM8 protein was detectable in these same cells throughout the cell cycle. Although no nuclear localization sequence could be found, MCM8 was detected in the nuclei of Hs27 and HeLa cells. Both nucleosolic and structure-bound MCM8 were observed. Detergent-permeabilized nuclei released nucleosolic MCM8 into the lysate, whereas the remaining structure-bound MCM8 portion in the pellet could be released by increasing salt

concentrations. Unlike the MCM3 control, MCM8 was structure bound in S, but not G1, based on HeLa cell synchronization at G2/M, late G1 and early S phases using nocodazole, mimosine and aphidicolin, respectively. Using these same cells, this group did not pull down MCM8 during immunoprecipitation procedure using antibodies against MCM3 or MCM4. MCM8 was expressed in Hs27, HeLa, HEK-293 and HuH7 cell lines, and was more highly expressed in the hepatocellular carcinoma-derived HuH7 cells than in normal (nonproliferating) liver tissue. The authors proposed a specific role for MCM8 in DNA replication based on its being structure bound in S phase and their lack of detection of MCM8 with MCM3 and MCM4, members of the G1 pre-replication complex (pre-RC).

Another group reported on a new human *MCM8* gene following a comparison of *MCM* family sequences against sequence tags expressed mRNAs (Johnson et al., 2003). By arranging these expressed cDNA segments contiguously, they identified an open reading frame (ORF) that was not identical to any known *MCM*. DNA sequencing of IMAGE cDNA clones led to confirmation of the ORF. They also searched the HTGS database of human genomic sequences to locate a BAC clone containing a unique gene encoding the new ORF with chromosomal location on 20p12.3-13. *MCM8*, comprised of 19 exons, was found to be located contrapodal to another gene comprised of 11 exons that encodes a homolog of yeast gene product GCD10. The sequences between the two transcription units were found to be TATA-less and highly GC-rich with multiple CpG units and to contain E2F, Sp1 and Pur binding elements. Notably, MCM8 was reported to have a canonical Walker A helicase domain as distinct from MCM2-7. This group also identified an MCM8 variant in a choriocarcinoma that is devoid of sixteen amino acids that are located N-terminal to the conserved helicase domain. They prepared a monoclonal antibody that specifically detects MCM8. Using this antibody, a fraction of MCM8 was found by this group to coisolate through several steps with MCM6 and MCM7 from HeLa cells. There was also an MCM8 gradient fraction not coincident with other MCMs that could be free MCM8. In addition, antibodies against MCM4, MCM6, and MCM7 coimmunoprecipitated MCM8 from HeLa cells as detected with anti-MCM8 antibody using immunoblotting. Using RT-PCR with commercially prepared cDNA, this group reported expression of *MCM8* mRNA in placenta, lung, liver, pancreas and heart. The same procedure using matched normal and tumor cDNAs involving cases of colon adenocarcinoma showed that *MCM8* mRNA expression was reduced relative to noncancerous tissue from the same patient. Results of their studies led these authors to propose a role for MCM8 in DNA replication or repair processes. They also proposed that MCM8 might substitute for another MCM at certain times in development or during the cell cycle. Due to a unique helicase motif among MCM proteins, they proposed that such a substitution might add a regulatory dimension to the function of the MCM complex.

Investigators performing studies in human HeLa and 293T cells reported that MCM8 has an important function during G1 in pre-RC assembly (Volkening & Hoffmann, 2005). This group prepared N- and C-terminal rabbit polyclonal antibodies against MCM8. Based on biochemical fractionation of HeLa cells following cell synchronization with nocodazole, MCM8 was found to accumulate on chromatin in G1 prior to binding of the MCM2-7 complex. MCM8 was found to be chromatin-bound throughout the cell cycle paralleling the binding of Cdc6 and Orc2. Following transfection of 293T cells with expression constructs for Ha-MCM8, HA-Cdc6 and HA-ORC2 (hemagglutinin-derived tag), MCM8 was found to interact with proteins Cdc6 and Orc2, both of which are components of the pre-RC. HA-MCM8 was not, however, found to interact with MCM2 or MCM6, also components of the pre-RC. Interactions were verified for these endogenous proteins in HeLa extracts

using anti-MCM8 antibody for immunoprecipitation followed by Western blots with antibody against Cdc6 or Orc2. Small hairpin RNAs were used to down-regulate endogenous MCM8 in HeLa cells, and this down regulation led to a delay in the entry of these cells into S phase as verified by flow cytometry. The authors proposed that this delay has to do with a role for MCM8 in G1 progression. The down-regulation of MCM8 was found to lead to a reduction of Cdc6 and MCM2-7 complex loaded onto chromatin. All these findings led these authors to report that interaction of MCM8 and Cdc6 is required for assembly of the pre-RC.

Using the *Xenopus* model, a group reported that MCM8 is required for efficient replication of chromosomal DNA in the *Xenopus* cell free replication assay (Maiorano et al., 2005). This group identified the *Xenopus* MCM8 homolog and studied its function using *Xenopus* egg extracts and demembrated sperm nuclei. They prepared a rabbit polyclonal antibody against a *Xenopus* MCM8 N-terminal peptide. Immunoprecipitation procedure using an anti-MCM3 antibody was performed and followed by Western detection using an anti-MCM8 antibody. MCM8 was not found to associate with soluble MCM3 or to complex with other components of the MCM2-7 complex present in S phase egg cytosol. The investigators isolated detergent-resistant chromatin fractions over a time course from S phase egg extracts to which demembrated sperm nuclei had been added. MCM8 was found to bind this chromatin at initiation of DNA replication after chromatin binding by Cdt1 and MCM2. In addition, they also isolated chromatin from membrane-depleted egg extracts, competent only for formation of the pre-RC, but not for initiation of DNA replication. MCM8 did not bind this chromatin. This group used immunofluorescence microscopy and Western blotting of chromatin fractions to determine the effect on MCM8 of adding aphidicolin at the time of initiation of DNA replication (which is not inhibited by aphidicolin) or during elongation (which is inhibited due to inhibition of DNA polymerase alpha). Although MCM3 was present in the labeled nuclei after both aphidicolin treatments, MCM8 was only present during elongation. Although MCM3 was present when S-CDK inhibitor, p21, was added at initiation (where p21 proteins blocks initiation of DNA synthesis but not formation of the pre-RC) and at the elongation time points, MCM8 was only present during elongation. Based on these experiments, investigators proposed that MCM8 binds chromatin after the pre-RC licensing step and at initiation of DNA synthesis. Immunodepleting MCM8 did not affect the chromatin loading of MCM3, a member of the pre-RC. The authors interpreted these data to mean that MCM8 is not required for replication licensing. This group also showed that recombinant MCM8 prepared from a baculovirus expression system possessed *in vitro* DNA helicase and ATPase activities. The helicase activity was lost by creating a mutation in the ATP binding site. Knockdown of MCM8 resulted in a 40% reduction in DNA synthesis compared to controls and in a slow replication phenotype. Nuclear assembly, however, was not affected. After depletion of MCM8, replication products were found to be short DNA chains similar to those produced in the presence of low concentrations of aphidicolin, which slows down replication. They found that depletion of MCM8 reduced chromatin bound Replication Protein A (RPA) 34 subunit and DNA polymerase alpha. In extracts depleted of MCM8, DNA synthesis could be rescued by reconstitution with ATP bound MCM8. Using immunofluorescence microscopy, MCM8 was found to colocalize with the RPA34 subunit and with replication foci. The authors proposed that MCM8 is a helicase that facilitates RPA recruitment as well as the processivity of DNA polymerases. The investigators proposed a function for MCM8 in the elongation step of DNA replication in regulating

fork movement that is distinct from the exclusive role of MCM8 in pre-initiation studies in HeLa and 293T cells noted by Volkening & Hoffmann, 2005.

These apparently contradictory roles for MCM8 in initiation versus elongation may be at least partially resolved by more recent findings. In human studies using HeLa cells, MCM8 was reported to colocalize with certain proteins involved in different aspects of DNA replication (Kinoshita et al., 2008). Using HeLa cells synchronized by a double-thymidine block, this group performed immunoprecipitation procedure with an anti-MCM7 antibody followed by Western detection using an anti-MCM8 monoclonal antibody, which they had previously prepared and characterized to be MCM8-specific. The resulting data revealed that the association between MCM8 and MCM7 peaked in mid G1, at the time of assembly of the pre-replication complex. Double chromatin immunoprecipitation (ChIP), developed by this group to determine the presence of two proteins on a specific segment of DNA at an origin of replication, was then used to show association of MCM8 with proteins involved in DNA replication. Using HeLa cells synchronized by double thymidine block, cell cycle studies were combined with double ChIP procedure. Genomic sequences upstream of the *c-MYC* gene were targeted through double ChIP using an antibody against Cdc6 for the first ChIP and followed by an anti-MCM8 antibody for the second ChIP. This procedure localized MCM8 with Cdc6, a protein reported to be involved in the subsequent loading of MCM proteins into the pre-RC (Kinoshita & Johnson, 2004; Kneissl et al., 2003; Lei & Tye, 2001). This localization was on specific DNA segments flanking the approximate center of the *c-MYC* replication initiation zone (Vassilev & Johnson, 1990), and occurred during both G1 and S phases, but not continuously. The investigators found a role for MCM8 in elongation was likely to be discontinuous from any role in initiation due to an MCM8 off signal in regard to its interaction with Cdc6 at the G1/S border and an MCM8 on again signal at the beginning of S phase. They also used double ChIP to show that MCM8 is present simultaneously at the *c-MYC* initiation zone with chromatin-bound Cdk2, a G1-S phase kinase essential for G1 to S phase transition. Immunogold electron microscopy was performed using mid-S phase HeLa cells, and MCM8 was strongly localized to heterochromatin, which replicates during this time. In these same samples, MCM8 was also shown to be in close localization with RPA70, a protein involved in elongation, not initiation. Based on these findings, the authors proposed distinct roles for MCM8 in DNA replication during G1 and S phases of the cell cycle. In addition, they suggest that MCM8 may also participate in processes distinct from replication initiation and elongation.

## 5. MCM8 protein function in DNA synthesis in *Drosophila*

Following the first two reports identifying human MCM8, the *Drosophila rec* gene was identified as a new member of the MCM family and was reported to be required for meiotic recombination in this species (Matsubayashi & Yamamoto, 2003). This group recognized the MCM domain and the putative Zn-finger motif through comparison with MCM2-7. They reported that *rec* mutations result in a very low level of meiotic recombination with primary non-disjunction at high frequency. These defects could be reversed in transformants carrying a wild type transgene. They identified molecular lesions consistent with induced mutant phenotypes. The investigators used DNA damage via methylmethane sulfonate or X-rays to determine that in somatic cells, the *rec* gene plays a limited or no role in DNA recombination and repair. They report that a role in pre-meiotic DNA replication is unlikely because electron microscopy revealed normal

synaptonemal complexes in *rec* mutants. In addition, there was normal oogenesis and oocyte development except for the lack of recombination. The authors suggested a role for *Dm* MCM8 in meiotic recombination.

Studies conducted in the *Drosophila* model by a second group revealed that the MCM8 orthologue REC is required for sufficient DNA synthesis to permit formation of a necessary meiotic crossover intermediate to drive this process (Blanton et al., 2005). This group found that in the absence of MCM8, recombination occurred through synthesis-dependent strand annealing to generate noncrossover products only. They reported that in meiotic recombination, REC functions at an intermediate step, but is not essential for pre-meiotic S phase. They found evidence for this in genomic DNA, where formation of normal synaptonemal complexes between homologous chromosomes in female *rec* mutants provided the supporting data. *Rec* mutant females displayed twice the rate of noncrossover gene conversions, but there was little crossover repair of double strand breaks. There was an increase in noncrossover events, but crossovers did not follow the normal distribution. There was a significant reduction in tract length in these mutants, which the authors suggest is due to diminished DNA repair synthesis. A role for *rec* in nonmeiotic cells was not reported. These authors proposed a model in which REC functions in meiotic crossover at the repair synthesis step.

Another group using *Drosophila* S2 cells reported a role for REC in DNA replication (Crevel et al., 2007). This group used dsRNA interference to deplete MCM2-8 family members. The depletion of MCM8 resulted in a reduction in fork number by 30%. They did not observe a significant effect on cell cycle or viability when analyzing flow cytometry profiles. They co-depleted MCM8 and MCM5, but did not observe a synergistic effect on cell cycle distribution. Based on an immunoblot, depletion of MCM8 did not have an effect on Orc5, MCM2, MCM5 or Cdc45 loading in chromatin. In a similar experiment, the investigators found a decrease in the amount of chromatin-bound PCNA detected, which they reported as evidence for a role for MCM8 in DNA replication in *Drosophila* S2 cells.

## 6. Role of MCM8 during the cell cycle

Reports differ as to where in the cell cycle MCM8 exerts its activity. Two different groups find that human MCM8 associates with Cdc6 in cultured cells, using two different approaches and three unique antibodies (Kinoshita et al., 2008; Volkening & Hoffmann, 2005). The two procedures included immunoprecipitation with anti-HA for exogenously expressed proteins or with anti-MCM8 polyclonal antibody for pull down of endogenous MCM8 (Volkening & Hoffmann, 2005) and by double ChIP of cell cycle fractions using monoclonal antibodies against Cdc6 and against MCM8 (Kinoshita et al., 2008). Cdc6, along with Cdt1, is involved in the subsequent loading of MCM2-7 onto the ORC complex loading deck during formation of the pre-RC (Forsburg, 2004, review). MCM8 was found to bind to chromatin throughout the cell cycle paralleling the binding of Cdc6 (Volkening & Hoffmann, 2005). Although MCM8 was found to interact with Orc2 based on the same immunoprecipitation experiments, Orc2 binding to chromatin was not affected by MCM8 silencing. Down regulation of MCM8 led to a concurrent decrease in Cdc6 on chromatin and delay of entry into S phase based on flow cytometry profiles leading the authors to suggest that MCM8 is required to load Cdc6 to chromatin (Volkening & Hoffmann, 2005). Cell cycle analysis of chromatin binding of Cdc6 during MCM8 down regulation is, however, not available. These additional data are needed to determine whether Cdc6 chromatin binding

was inhibited in G1 or S or both. In addition, MCM8 was found to be concurrently present with Cdc6 at sequences in the *c-MYC* replication initiation zone in HeLa cells when double chromatin immunoprecipitation (ChIP) with anti-Cdc6 antibody was followed by ChIP using anti-MCM8 antibody. Cell cycle studies using double ChIP revealed that these proteins are both present at the *c-MYC* initiation zone sequence in G1 as well as in S phase, but not continuously due to an *on to off* signal at the G1/S border followed by an *off to on* signal (Kinoshita et al., 2008). Using the same double ChIP procedure, MCM8 was found to be present with Cdk2 on specific sequences at the initiation zone of the *c-MYC* gene. Cdk2 is required for transition into S phase, and this association of MCM8 with Cdk2 supports a role for MCM8 in the transition. During this same time range (G1/S), there was little to no association of MCM8 with Cdc6 on these same sequences. This transition period requires further study.

MCM8 was also found to colocalize with RPA subunits during S phase by two groups (Kinoshita et al., 2008; Maiorano et al., 2005). RPA binds to single-stranded DNA and therefore serves as a marker for unwound strands of replicating DNA. MCM8 was found to colocalize with the RPA34 subunit during S phase in the *Xenopus* replication model (Maiorano et al., 2005). In HeLa mid-S phase cells, dual immunogold electron microscopy studies showed that MCM8 colocalized with RPA70, RPA large subunit (Kinoshita et al., 2008). When MCM7 and MCM8 were compared in this study, there was a difference in the amount of MCM7 and MCM8 that colocalized with RPA70 where colocalization is defined as two beads within 10 nm distance of each other. MCM7 was most often present within 10 nm distance of RPA70. Whereas MCM8 was sometimes within 10 nm of RPA70, it was most often present within 100 nm distance of RPA70. Thus, both in the *Xenopus* replication assay and in HeLa cells, MCM8 colocalizes with or near a protein involved in DNA replication during S phase. Although MCM2-7 proteins have not generally been shown to do so, MCM8 was shown to colocalize with replication foci in *Xenopus* nuclei labeled with bromodeoxyuridine (Maiorano et al., 2005).

The major disagreement among studies regards whether MCM8 interacts with the other MCM2-7 family members. This has been addressed in the *Xenopus* model by Maiorano et al., 2005, who investigated whether MCM8 is present along with other members of the MCM complex when the pre-RC forms in demembranated sperm nuclei in egg extracts or when either initiation or elongation stages of replication are specifically inhibited in these nuclei. No association of MCM8 with MCM2 or MCM3 could be found in these studies. The data were in agreement with that of Gozuacik et al., 2003, where chromatin-bound MCM8 could not be detected in pull down lysates from the pellet 1 fraction containing the MCM2-7 complex using anti-MCM3 and anti-MCM4 antibodies. This same group specifically synchronized HeLa cells at G2/M, late G1 and early S phases using nocodazole, mimosine and aphidicolin, respectively. Following this specific synchronization, they did not detect MCM8 in the late G1 pellet fraction remaining after removal of the Triton X-100-extractable supernatant when using anti-MCM8 polyclonal antibody. Volkening et al., 2005, using HA-tagged MCM8, Cdc6 and Orc2, carried out transfection studies to analyze MCM8 association with Cdc6 and Orc2. Although association of MCM8 with Cdc6 and Orc2 was found, association of MCM8 with MCM2 and MCM6 was not found even though these proteins are also members of the pre-RC. They found that MCM8 loaded onto chromatin in G1 prior to the other MCMS in a profile similar to Cdc6 and Orc2. Using double thymidine block of HeLa cells and immunoprecipitation with anti-MCM7 antibody followed by Western detection using anti-MCM8 monoclonal antibody,

Kinoshita et al., 2008, found an association of MCM8 with MCM7 during both G1 and S phases of the cell cycle.

Perhaps differences in experimental protocols or synchronization procedures may have led to different end results when investigating the association of MCM8 with members of the MCM2-7 complex. Three of the studies were completed in human cells (Gozuacik et al., 2003, Kinoshita, 2008 #4; Volkening & Hoffmann, 2005). The third study was done using the *Xenopus* replication model (Maiorano et al., 2005). There is the possibility of species-specific differences or of questions regarding the ability of one model to recapitulate the same process in the other. There may be weak binding that is lost under certain conditions. More experiments where proteins are cross-linked to DNA, such as in the ChIP procedure, may be informative. It is possible that antibodies raised against different epitopes may not yield the same result due to differential accessibility of the epitopes when the MCM8 protein is folded or part of an oligomer or when it is part of a multi-protein complex. More work is needed to resolve the differences, which may be due in part to the multi-dimensional and dynamic aspects of this complex machinery.

## 7. Summary of MCM8 and interacting partners involved in DNA replication

Evidence described in Section 6 indicates a role for MCM8 in DNA replication during G1 and S phases of the cell cycle. This evidence is based on the apparent interaction of MCM8 with proteins forming the pre-RC during G1 or with proteins involved in elongation during S phase. Whereas all six MCM proteins of the MCM2-7 complex are essential for DNA replication fork progression (Labib et al., 2000), the exact role of these proteins during the cell cycle is not known. The role of MCM8 is also unknown as is its active assembly and configuration. In contrast to members of the MCM2-7 complex, MCM8 has its own intrinsic helicase activity based on studies using the *Xenopus* replication assay (Maiorano et al., 2005). Whether MCM8 could replace one or more members of the MCM2-7 complex at some time during development or during the cell cycle is unknown. The canonical Walker A helicase motif of MCM8, unique among MCMs, could indicate a regulatory role if replacing another family member (Johnson et al., 2003).

There is strong evidence for an interaction between MCM8 and Cdc6. Two independent groups confirmed the interaction of MCM8 with Cdc6 in human cells (Johnson et al., 2003; Volkening & Hoffmann, 2005). There are questions regarding whether the interaction occurs only during G1 phase prior to loading of other MCMs onto chromatin (Volkening & Hoffmann, 2005) or, if indeed, MCM8 is assembled onto chromatin during G1 (Maiorano et al., 2005). An interruption in MCM8 interaction with Cdc6 at the G1/S border could indicate that MCM8 may be involved in a switch mechanism or perform a different function during G1 than in S phase (Kinoshita et al., 2008). There is also strong evidence for the interaction of MCM8 with RPA during S phase based on immunofluorescence microscopy (Maiorano et al., 2005) and immunogold electron microscopy (Kinoshita et al., 2008). There is evidence that much of MCM8 acts with MCM6 and MCM7, while a significant fraction of MCM8 is independent (Johnson et al., 2003). Which MCM partners MCM8 might act with functionally is presently not known.

Studies in *Drosophila* point to a role in meiotic recombination (Matsubayashi & Yamamoto, 2003) and at the repair synthesis step in meiosis crossover (Blanton et al., 2005). Future studies should address whether there is a similar role for MCM8 in other species or whether this is a specific function of *Drosophila* MCM8 brought about by sequence divergence. In

*Drosophila* S2 cells, MCM8 was found to have a role in replication (Crevel et al., 2007), but direct interactions with proteins involved in DNA replication were not studied.

Based on transfection studies carried out in human HCT116 cells along with semiquantitative RT-PCR, *MCM8* mRNA was shown to be upregulated by exogenous E2F1 (Hayashi et al., 2006). This group also used chromatin immunoprecipitation to demonstrate that E2F1 and NF-Y each directly associate with the human *MCM8* promoter in HCT116 cells. The investigators carried out transfection of HeLa cells using *MCM8* promoter-luciferase constructs and an expression vector for E2F1 in a dual luciferase reporter assay to demonstrate that transcriptional activation required an E2F-binding motif near the site for transcription initiation. Activation of the human *MCM8* promoter was achieved by E2F1-4 transcription factors in these assays.

Fig. 5 presents MCM8 interacting partners detected by various experimental procedures based predominately on cell cycle studies as described in this review. MCM8 interaction with Cdk2 is also included since Cdk2 is a G1- S phase kinase that is essential for transition into S phase.

<b>MCM8 Interacting Partners</b>		
<b>G1</b>	<b>G1/S</b>	<b>S</b>
<b>Orc2</b>	<b>MCM8 switch?*</b>	<b>Cdc6</b>
<b>Cdc6</b>	<b>G1 on to off</b>	<b>RPA34</b>
	<b>S off to on</b>	<b>RPA70</b>
		<b>Cdk2</b>

\* E2F interaction with MCM8 promoter

Fig. 5. Proteins reported to interact with MCM8 during the cell cycle. References and experimental approaches used: *Orc2*, Volkening and Hoffmann, 2005, immunoprecipitation (IP); *Cdc6*, Volkening and Hoffmann, 2005, IP; *Cdc6*, Kinoshita et al., 2008, double chromatin immunoprecipitation, (ChIP); *RPA34*, Maiorano et al., 2005, immunofluorescence microscopy; *RPA70*, Kinoshita et al., 2008, immunogold electron microscopy; *Cdk2*, Kinoshita et al., 2008, double ChIP; *Cdc6/MCM8 switch G1 on to off, S off to on*, Kinoshita et al., 2008, double ChIP; *MCM8 promotor (E2F, NF-Y)*, Hayashi et al., 2006, ChIP.

Determination of whether MCM8 acts as an AAA+ molecular motor in the three model species discussed will provide some insight into the enigma of MCM8 function during the cell cycle and in DNA replication. The evolutionary advantage for adding a new member to the MCM family and what sets it apart from the MCM2-7 complex are important questions to answer as researchers investigate the MCM8 active three-dimensional configuration and interactive targets in *Hs*, *Xl* and *Dm*.

## 8. Acknowledgment

This project was supported by grant funding from Virginia's Commonwealth Health Research Board to DCD and by NIH (NINDS and NCI) funding to EMJ.

## 9. References

- Bae, B., Chen, Y.H., Costa, A., Onesti, S., Brunzelle, J.S., Lin, Y., Cann, I.K., & Nair, S.K. (2009). Insights into the architecture of the replicative helicase from the structure of an archaeal MCM homolog. *Structure* 17, 211-222, 0969-2126 (Print) 0969-2126 (Linking).
- Blanton, H.L., Radford, S.J., McMahan, S., Kearney, H.M., Ibrahim, J.G., & Sekelsky, J. (2005). REC, *Drosophila* MCM8, drives formation of meiotic crossovers. *PLoS Genet* 1, e40, 1553-7404 (Electronic).
- Blow, J.J. & Laskey, R.A. (1988). A role for the nuclear envelope in controlling DNA replication within the cell cycle. *Nature* 332, 546-548, 0028-0836 (Print) 0028-0836 (Linking).
- Bochman, M.L. & Schwacha, A. (2009). The Mcm complex: unwinding the mechanism of a replicative helicase. *Microbiol Mol Biol Rev* 73, 652-683, 1098-5557 (Electronic) 1092-2172 (Linking).
- Brewster, A.S., Wang, G., Yu, X., Greenleaf, W.B., Carazo, J.M., Tjajadia, M., Klein, M.G., & Chen, X.S. (2008). Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase. *Proc Natl Acad Sci U S A* 105, 20191-20196, 1091-6490 (Electronic) 0027-8424 (Linking).
- Chong, J.P., Mahbubani, H.M., Khoo, C.Y., & Blow, J.J. (1995). Purification of an MCM-containing complex as a component of the DNA replication licensing system. *Nature* 375, 418-421, 0028-0836 (Print) 0028-0836 (Linking).
- Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 16, 10881-10890, 0305-1048 (Print) 0305-1048 (Linking).
- Crevel, G., Hashimoto, R., Vass, S., Sherkow, J., Yamaguchi, M., Heck, M.M., & Cotterill, S. (2007). Differential requirements for MCM proteins in DNA replication in *Drosophila* S2 cells. *PLoS One* 2, e833, 1932-6203 (Electronic) 1932-6203 (Linking).
- Davey, M.J., Indiani, C., & O'Donnell, M. (2003). Reconstitution of the Mcm2-7p heterohexameric subunit arrangement, and ATP site architecture. *J Biol Chem* 278, 4491-4499, 0021-9258 (Print) 0021-9258 (Linking).
- Davey, M.J., Jeruzalmi, D., Kuriyan, J., & O'Donnell, M. (2002). Motors and switches: AAA+ machines within the replisome. *Nat Rev Mol Cell Biol* 3, 826-835, 1471-0072 (Print) 1471-0072 (Linking).
- Erzberger, J.P. & Berger, J.M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu Rev Biophys Biomol Struct* 35, 93-114, 1056-8700 (Print) 1056-8700 (Linking).
- Fletcher, R.J., Bishop, B.E., Leon, R.P., Sclafani, R.A., Ogata, C.M., & Chen, X.S. (2003). The structure and function of MCM from archaeal *M. thermoautotrophicum*. *Nat Struct Biol* 10, 160-167, 1072-8368 (Print) 1072-8368 (Linking).
- Forsburg, S.L. (2004). Eukaryotic MCM proteins: beyond replication initiation. *Microbiol Mol Biol Rev* 68, 109-131, 1092-2172 (Print) 1092-2172 (Linking).
- Gomez, E.B., Catlett, M.G., & Forsburg, S.L. (2002). Different phenotypes in vivo are associated with ATPase motif mutations in *Schizosaccharomyces pombe* minichromosome maintenance proteins. *Genetics* 160, 1305-1318, 0016-6731 (Print) 0016-6731 (Linking).

- Gozuacik, D., Chami, M., Lagorce, D., Faivre, J., Murakami, Y., Poch, O., Biermann, E., Knippers, R., Brechot, C., & Paterlini-Brechot, P. (2003). Identification and functional characterization of a new member of the human Mcm protein family: hMcm8. *Nucleic Acids Res* 31, 570-579, 1362-4962 (Electronic) 0305-1048 (Linking).
- Hayashi, R., Goto, Y., Haga, A., Kobayashi, D., Ikeda, R., & Yoshida, K. (2006). Comparative genomics on MCM8 orthologous genes reveals the transcriptional regulation by transcription factor E2F. *Gene* 367, 126-134, 0378-1119 (Print) 0378-1119 (Linking).
- Hennessy, K.M., Lee, A., Chen, E., & Botstein, D. (1991). A group of interacting yeast DNA replication genes. *Genes Dev* 5, 958-969, 0890-9369 (Print) 0890-9369 (Linking).
- Iyer, L.M., Leipe, D.D., Koonin, E.V., & Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *J Struct Biol* 146, 11-31, 1047-8477 (Print) 1047-8477 (Linking).
- Jenkinson, E.R. & Chong, J.P. (2006). Minichromosome maintenance helicase activity is controlled by N- and C-terminal motifs and requires the ATPase domain helix-2 insert. *Proc Natl Acad Sci U S A* 103, 7613-7618, 0027-8424 (Print) 0027-8424 (Linking).
- Johnson, E.M., Kinoshita, Y., & Daniel, D.C. (2003). A new member of the MCM protein family encoded by the human MCM8 gene, located contrapodal to GCD10 at chromosome band 20p12.3-13. *Nucleic Acids Res* 31, 2915-2925, 1362-4962 (Electronic) 0305-1048 (Linking).
- Kinoshita, Y. & Johnson, E.M. (2004). Site-specific loading of an MCM protein complex in a DNA replication initiation zone upstream of the c-MYC gene in the HeLa cell cycle. *J Biol Chem* 279, 35879-35889, 0021-9258 (Print) 0021-9258 (Linking).
- Kinoshita, Y., Johnson, E.M., Gordon, R.E., Negri-Bell, H., Evans, M.T., Coolbaugh, J., Rosario-Peralta, Y., Samet, J., Slusser, E., Birkenbach, M.P., & Daniel, D.C. (2008). Colocalization of MCM8 and MCM7 with proteins involved in distinct aspects of DNA replication. *Microsc Res Tech* 71, 288-297, 1059-910X (Print) 1059-910X (Linking).
- Kneissl, M., Putter, V., Szalay, A.A., & Grummt, F. (2003). Interaction and assembly of murine pre-replicative complex proteins in yeast and mouse cells. *J Mol Biol* 327, 111-128, 0022-2836 (Print) 0022-2836 (Linking).
- Koonin, E.V. (1993). A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. *Nucleic Acids Res* 21, 2541-2547, 0305-1048 (Print) 0305-1048 (Linking).
- Kubota, Y., Mimura, S., Nishimoto, S., Takisawa, H., & Nojima, H. (1995). Identification of the yeast MCM3-related protein as a component of *Xenopus* DNA replication licensing factor. *Cell* 81, 601-609, 0092-8674 (Print) 0092-8674 (Linking).
- Labib, K., Tercero, J.A., & Diffley, J.F. (2000). Uninterrupted MCM2-7 function required for DNA replication fork progression. *Science* 288, 1643-1647.

- Lei, M. & Tye, B.K. (2001). Initiating DNA synthesis: from recruiting to activating the MCM complex. *J Cell Sci* 114, 1447-1454, 0021-9533 (Print) 0021-9533 (Linking).
- Liu, Y., Richards, T.A., & Aves, S.J. (2009). Ancient diversification of eukaryotic MCM DNA replication proteins. *BMC Evol Biol* 9, 60, 1471-2148 (Electronic) 1471-2148 (Linking).
- Lutzmann, M., Maiorano, D., & Mechali, M. (2005). Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2-8 protein family. *Gene* 362, 51-56, 0378-1119 (Print) 0378-1119 (Linking).
- Maine, G.T., Sinha, P., & Tye, B.K. (1984). Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics* 106, 365-385, 0016-6731 (Print) 0016-6731 (Linking).
- Maiorano, D., Cuvier, O., Danis, E., & Mechali, M. (2005). MCM8 is an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. *Cell* 120, 315-328, 0092-8674 (Print).
- Matsubayashi, H. & Yamamoto, M.T. (2003). REC, a new member of the MCM-related protein family, is required for meiotic recombination in *Drosophila*. *Genes Genet Syst* 78, 363-371, 1341-7568 (Print) 1341-7568 (Linking).
- Neuwald, A.F., Aravind, L., Spouge, J.L., & Koonin, E.V. (1999). AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* 9, 27-43, 1088-9051 (Print) 1088-9051 (Linking).
- Saraste, M., Sibbald, P.R., & Wittinghofer, A. (1990). The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* 15, 430-434, 0968-0004 (Print) 0968-0004 (Linking).
- Shultz, R.W., Tatineni, V.M., Hanley-Bowdoin, L., & Thompson, W.F. (2007). Genome-wide analysis of the core DNA replication machinery in the higher plants *Arabidopsis* and rice. *Plant Physiol* 144, 1697-1714, 0032-0889 (Print) 0032-0889 (Linking).
- Thommes, P., Kubota, Y., Takisawa, H., & Blow, J.J. (1997). The RLF-M component of the replication licensing system forms complexes containing all six MCM/P1 polypeptides. *Embo J* 16, 3312-3319, 0261-4189 (Print) 0261-4189 (Linking).
- Tye, B.K. (1999). MCM proteins in DNA replication. *Annu Rev Biochem* 68, 649-686, 0066-4154 (Print) 0066-4154 (Linking).
- Vassilev, L. & Johnson, E.M. (1990). An initiation zone of chromosomal DNA replication located upstream of the c-myc gene in proliferating HeLa cells. *Mol Cell Biol* 10, 4899-4904, 0270-7306 (Print) 0270-7306 (Linking).
- Volkening, M. & Hoffmann, I. (2005). Involvement of human MCM8 in prereplication complex assembly by recruiting hcdc6 to chromatin. *Mol Cell Biol* 25, 1560-1568, 0270-7306 (Print) 0270-7306 (Linking).
- Walker, J.E., Saraste, M., Runswick, M.J., & Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *Embo J* 1, 945-951, 0261-4189 (Print) 0261-4189 (Linking).

Yan, H., Merchant, A.M., & Tye, B.K. (1993). Cell cycle-regulated nuclear localization of MCM2 and MCM3, which are required for the initiation of DNA synthesis at chromosomal replication origins in yeast. *Genes Dev* 7, 2149-2160, 0890-9369 (Print) 0890-9369 (Linking).

# Regulation of MCM7 DNA Replication Licensing Activity

Jian-Hua Luo and Yan P. Yu

*Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh  
USA*

## 1. Introduction

Miniature chromosome maintenance (MCM) proteins were initially identified from autonomously replicating sequence in *Saccharomyces cerevisiae*. Mutations of some of these proteins such as MCM7 or MCM3 in yeast result in loss of the large chunk of yeast chromosomes. MCM7 cDNA encodes a 543-amino acid protein and is ubiquitously expressed in all tissues. Initiation of DNA replication is a complex process involving the concerted action of many proteins. A large body of studies indicate that MCM7 is a critical component of DNA replication licensing complex in the yeast and xenopus (Chong et al., 1996; Coxon et al., 1992; Dalton and Whitbread, 1995; Kearsey et al., 1996). Some studies suggest that MCM4, 6 and 7 complex contains DNA helicase activity (Ishimi, 1997; You et al., 1999). DNA replication licensing complex is multimeric and phase specific. In the yeast, DNA replication licensing proteins such as MCM2-7 and several replication origin binding proteins such as Cdc6 and Cdt1 forms DNA replication licensing complex in G1 phase to enable DNA replication and to promote cell cycle entry into S phase. Such complex, however, dissipates in the S, G2 and M phases to prevent re-firing of DNA replication, and thus protect the integrity of genomes. There is little interest in MCM complex as target for oncogenic or tumor suppressor pathway until the links of MCM7 over-expression and amplification to several human malignancies were found (Brake et al., 2003; Honeycutt et al., 2006; Kan et al., 2009; Ren et al., 2006).

## 2. MCM7 transforming oncogenic activity

Initial implication of MCM7 involvement in human malignancies is positive immunostaining of MCM7 in several human malignancies, including endometrial carcinoma (Li et al., 2005), melanoma (Gambichler et al., 2009), esophageal adenocarcinoma (Kan et al., 2009), colorectal adenocarcinoma (Nishihara et al., 2008), oral squamous cell carcinoma (Feng et al., 2008), glioblastoma (Facoetti et al., 2006), and thyroid cancer (Kebebew et al., 2006). Most of these studies used MCM7 as a proliferation marker to compare with the existing markers such as Ki-67 or PCNA. The first study addressing the oncogenic role of MCM7 came from genome analysis of prostate cancer. By performing a genome wide copy number analysis using biotin-labeled genome DNA on Affymetrix U133 2.0 chip (Ren et al., 2006) (figure 1). Ninety-two genes and expressed sequenced tags (ESTs) consistently had a two-fold or greater number of genome copies in the prostate cancer

specimens relative to matched blood cells ( $p < 0.01$ , using baseline analysis by Microarray Suite 5.0™). To determine whether these gene amplifications were associated with increased expression of the same loci, mRNA expression arrays were performed with the same tumor samples and matched normal prostate tissue adjacent to but separate from the tumor. Only five of the 92 potential amplifications had increased mRNA expression. Using SYBR-green quantitative PCR and fluorescence *in situ* hybridization (FISH) with bacterial artificial chromosome containing these genes, it has been confirmed that two of these genes were amplified in the prostate cancer specimens. One of these genes, MCM7, located at 7q21.3 and a component of DNA replication licensing complex, was found to be amplified 4 and 15 fold, respectively.

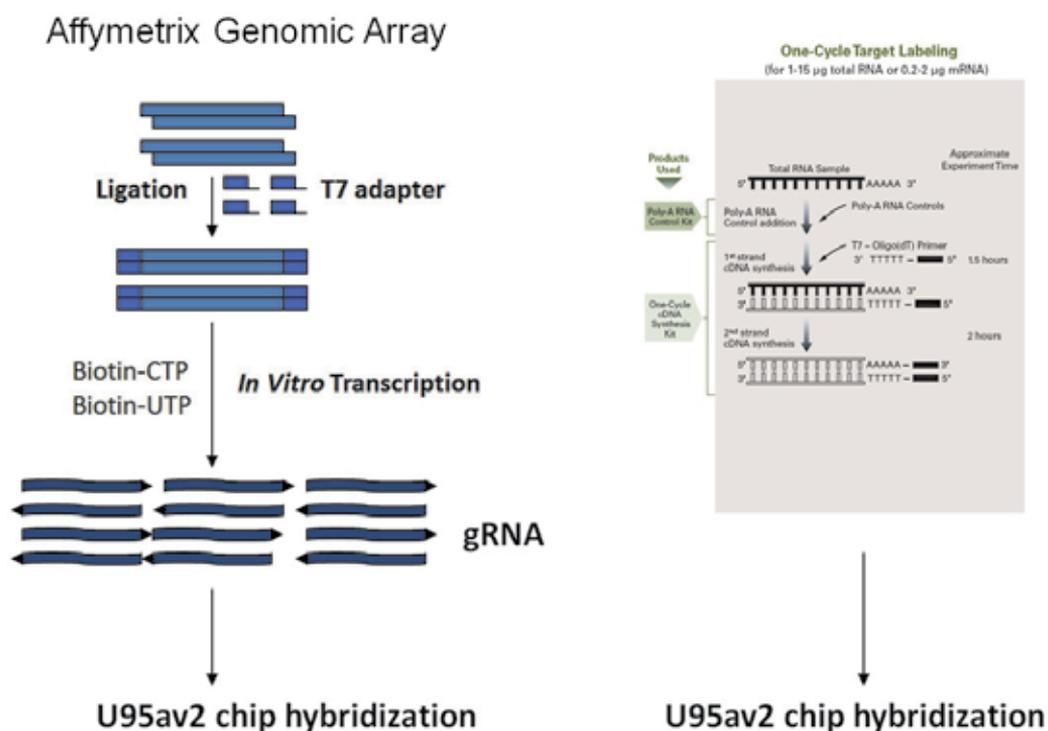


Fig. 1. Schematic diagram of genome hybridization on U133 2.0 chip. Left panel: Genome DNA from prostate cancer samples with matched blood were digested with Tsp509I restriction enzyme and ligated with an adaptor containing T7 promoter. In vitro transcription was then carried out to generate biotin-labeled gRNA. The gRNA was then fragmented and hybridized to HU133 2.0 chip for signal analysis. Right panel: Total RNA from matched samples of the left were reverse transcribed using T7-oligo dT (24) primer. After double-stranded cDNA synthesis, in vitro transcription using GTP, ATP, biotin-labeled UTP and CTP was then performed. The resulting cRNA was then fragmented and hybridized to U95av2 chip.

Subsequent validation analyses suggest that either copy number and/or protein level increase of MCM7 are associated with prostate cancer relapse and metastasis. To determine the prevalence of MCM7 amplification in prostate cancer, FISH analyses were performed on

74 prostate samples, including 58 cancers and 16 benign tissues (Figure 2). Nearly half, (26/58, 45%) of the prostate cancers had at least a doubling in the number of genome copies (compared to the centromere of chromosome 7) (Table 1). Meanwhile, none of the benign prostate tissues demonstrated a similar amplification. As further validation of these findings, SYBR-green quantitative PCR was performed on an additional 133 prostate samples, including 119 cancers and 14 normal tissues. In this analysis, fifty percent (59/119) of the prostate cancers had at least a 2-fold increase in the MCM7/ $\beta$ -actin ratio relative to DNA from benign prostate specimens. In the most up-regulated case, the MCM7/ $\beta$ -actin ratio in the cancer sample was 16-fold greater than control levels.

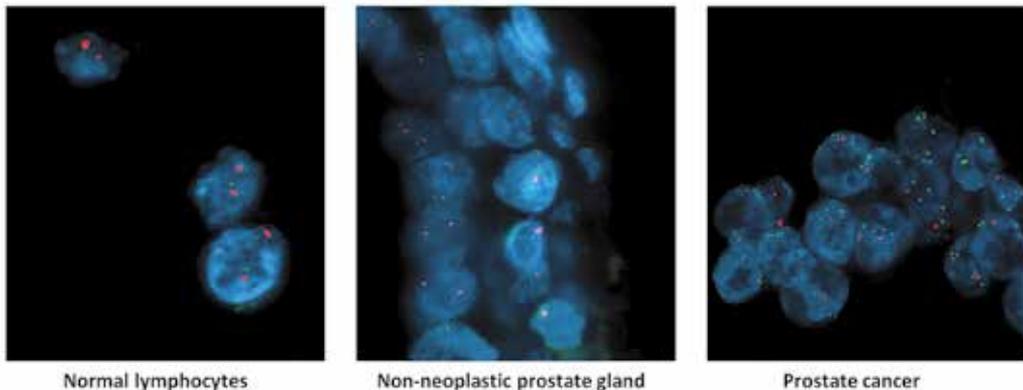


Fig. 2. MCM7 amplification and overexpression in prostate cancer. Images of fluorescence *in situ* hybridization (FISH) on normal lymphocytes, non-neoplastic prostate gland and prostate cancer. Chromosome 7 centromere is labeled with orange spectrum (orange), and MCM7 is labeled with fluorescein (green).

Amplification of MCM7 was associated with a significantly increased rate of prostate cancer relapse within 5 years of radical prostatectomy (defined as a detectable PSA): 76.5% (52/68) patients with MCM7 amplification relapsed compared to only 12.3% (7/57) of patients without MCM7 amplification ( $p < 0.0001$ , log-rank test) (figure 3A). Similar results were seen with MCM7 protein expression: 76.3% patients with  $\geq 2+$  MCM7 expression experienced a recurrence within 5 years after radical prostatectomy compared with only 26.5% patients with weak MCM7 expression ( $p < 0.006$ ) (figure 3B). The combination of MCM7 amplification and overexpression generated additional improvement in separating the two groups (Figure 3C).

Eleven clones of DU145 transfected with a constitutive MCM7 cDNA overexpression vector were characterized by Western blot analysis, and two with high MCM7 expression were chosen for subsequent analysis. Over a 6-h experimental period DU145 pCMV-MCM7 clones had twice the number of cells enter S phase (figure 3B), with cell proliferation being 50% faster than vector only controls (figure 3C). DU145 pCMV-MCM7 clones had a 2-fold higher level of invasiveness in Matrigel transmigration analysis compared to vector-only controls. *In vivo*, DU145 pCMV-MCM7 xenograft tumors were twelve times larger than vector-only control tumors, 2.58 cm<sup>3</sup> versus 0.22 cm<sup>3</sup>, respectively ( $p < 0.0001$ ), with a 50% mortality within 6 weeks of inoculation ( $p = 0.0063$ ). These findings are consistent with the clinical findings that increased MCM7 expression is associated with higher rates of local tumor invasion.

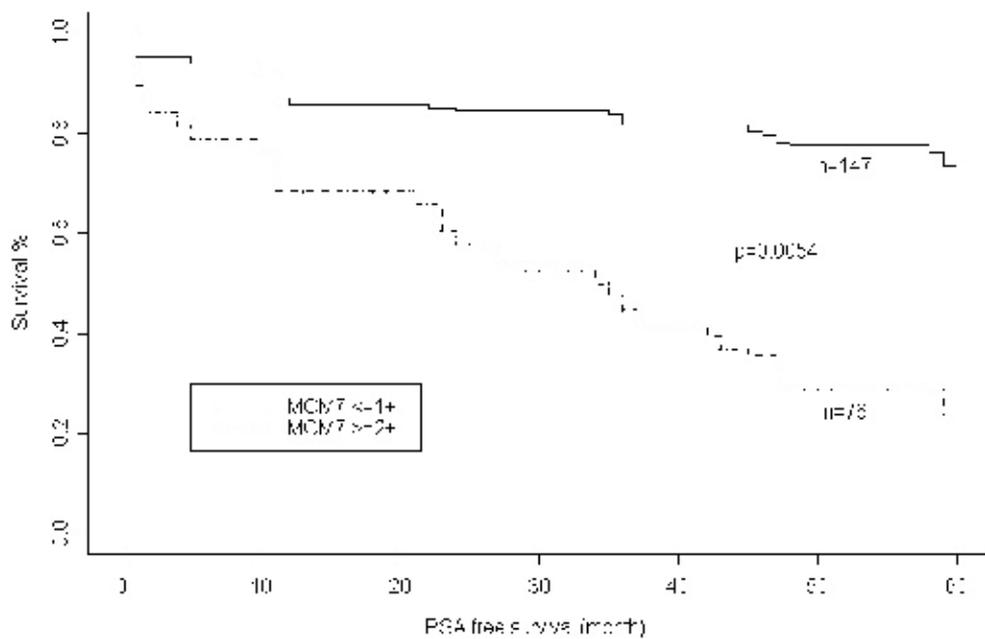
Case number	Gleason score	Pathology stage	MCM7/ centromere	Relapse*
95T	7	T3a	6.67	Y
72T	7	T3a	6.42	Y
54T	9	T3b	6.1	Y
45T	6	T2b	5.49	Y
139T	8	T2b	4.2	Y
69T	6	T3a	4.15	Y
5T	7	T2b	3.97	Y
85T	7	T2b	3.77	Y
48T	8	T3a	3.5	Y
4T	9	T3b	3.3	Y
101T	6	T3a	3.11	Y
107T	6	T3a	3.11	Y
2213T	6	T3a	3.06	Y
42T	7	T3a	2.69	Y
214T	6	T3a	2.66	Y
3T	9	T3b	2.52	Y
2T	7	T3a	2.49	Y
147T	6	T3a	2.46	Y
68T	6	T2b	2.1	Y
22T	6	T3a	1.86	Y
7T	7	T3a	1.62	Y
93T	9	T3b	1.52	Y
6T	6	T2b	1.5	Y
641T	6	T2b	1.1	Y
56T	6	T2b	0.95	Y
39T	7	T3a	0.84	Y
40T	6	T3a	5.76	N
75T	6	T2b	5.55	N
911T	6	T3a	4.48	N
46T	7	T3a	4.47	N
83T	7	T2b	1.94	N
99T	8	T3a	1.82	N
13T	7	T2b	1.74	N
67T	9	T3b	1.64	N
1127T	6	T3a	1.58	N
M35	6	T2b	1.55	N
26T	6	T2a	1.54	N
11T	7	T2b	1.54	N
194T	6	T2b	1.37	N

82T	7	T3a	1.24	N
55T	6	T2b	1.18	N
91T	7	T2b	1.14	N
94T	7	T3a	1.13	N
96T	6	T2a	0.97	N
941T	6	T2b	2.17	U
M33	8	T3b	2.13	U
M29	7	T3b	2.05	U
217T	7	T3a	1.96	U
121T	6	T2b	1.86	U
M21	6	T2b	1.68	U
M24	6	T2b	1.64	U
853T	6	T2b	1.49	U
M28	8	T3b	1.39	U
998T	6	T2b	1.29	U
M36	7	T2b	1.14	U
M31	7	T3a	1.11	U
828T	6	T2b	1	U
M23	7	T2b	0.94	U
M2	Normal Prostate		1.2	N/A
A4	Normal Prostate		1.01	N/A
M4	Normal Prostate		0.9	N/A
P28	Normal Prostate		1.2	N/A
P91	Normal Prostate		0.89	N/A
P93	Normal Prostate		0.95	N/A
U-2	Normal Prostate		1.3	N/A
D-41	Normal Prostate		0.8	N/A
U-4	Normal Prostate		1.2	N/A
9	Normal Prostate		1	N/A
19	Normal Prostate		1.09	N/A
99	Normal Prostate		0.93	N/A
11	Normal Prostate		1.02	N/A
12	Normal Prostate		1.08	N/A
13	Normal Prostate		1.03	N/A
14	Normal Prostate		1.05	N/A

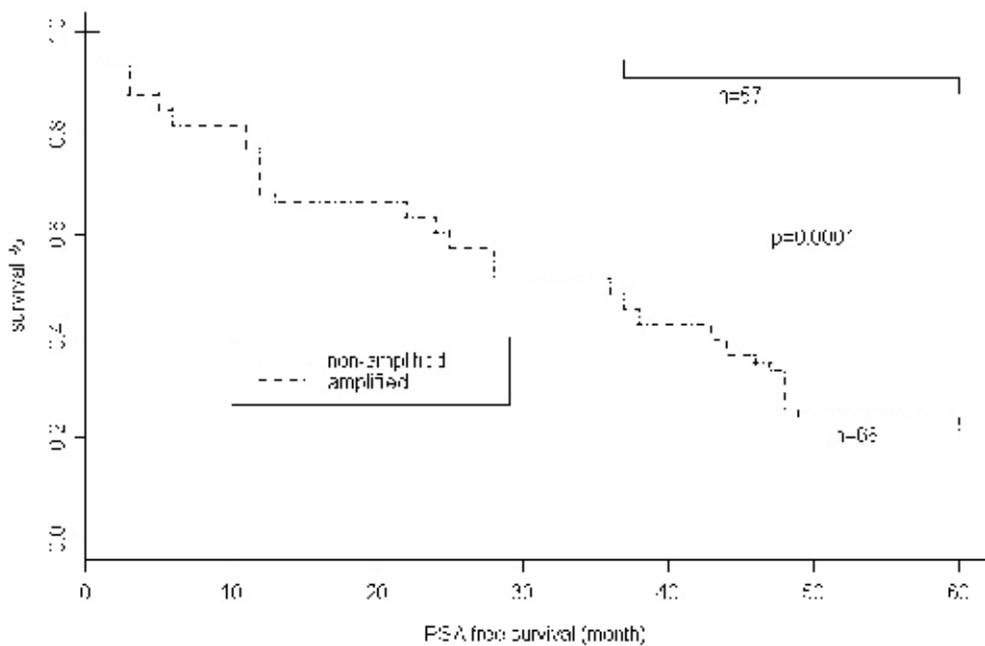
Y=yes; N-No; U-undetermined due to lack of follow-up; N/A-not applicable.

\*Determined by PSA relapse or physical evidences of metastasis within 5 years after prostatectomy.

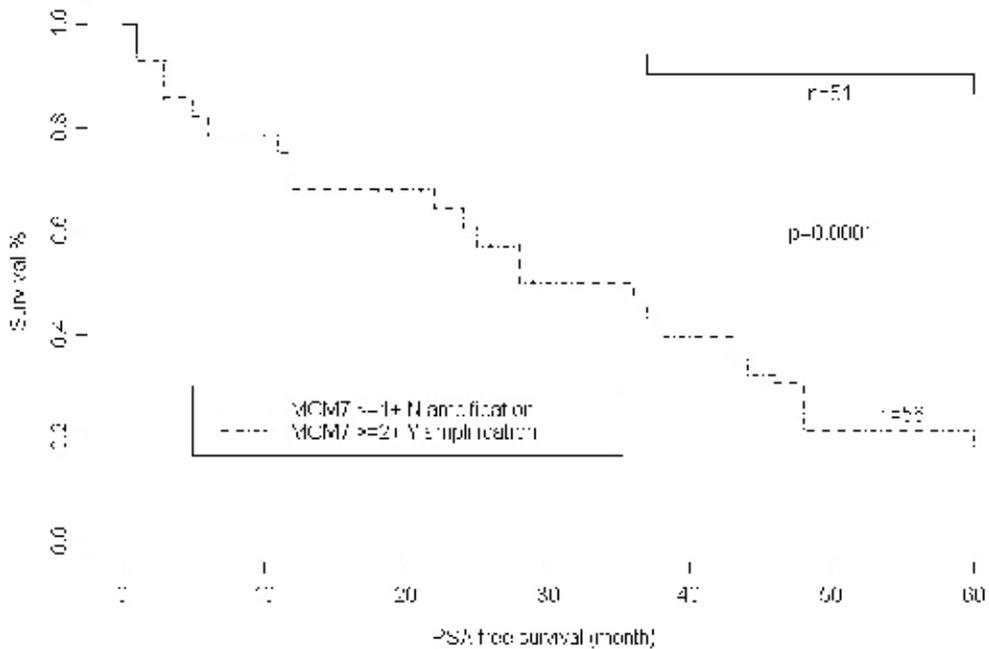
Table 1. FISH analysis of MCM7 amplification in primary prostate cancer



(a)



(b)



(c)

Fig. 3. MCM7 amplification and overexpression associated with prostate cancer relapse. A. Kaplan curves of samples from patients with clinical follow-up for 5-year period. The samples were divided by presence (Y) or absence (N) of MCM7 amplification based on FISH and/or quantitative PCR. Ratios of MCM7/Centromere or MCM7/ $\beta$ -actin  $\geq 2$  are considered an increase of MCM7 DNA. B. Kaplan curves of samples from patients with clinical follow-up for 5-year period. The samples were divided by MCM7 expression scores: scoring  $\geq 2+$  or  $\leq 1$ . C. Kaplan curves of samples from patients with clinical follow-up for 5-year period. The samples were divided by presence (Y) or absence (N) of MCM7 amplification and MCM7 overexpression.

Amplification of MCM7 was also found in esophageal carcinoma (Kan et al., 2009). The magnitude of MCM7 amplification correlates with the expression of MCM7, tumor grades and aggressiveness of the esophageal cancer (Kan et al., 2009). It is presumed that amplification of MCM7 is the driving force of MCM7 over-expression in primary human malignancies. When MCM7 was transgened in basal cells of skin of mice utilizing keratin promoter, the animals developed squamous cell carcinoma upon DMBA/TPA challenge versus complete negative results from the WT controls (Honeycutt et al., 2006). However, MCM7 does not play an initiator role in cancer development because organ specific MCM7 transgene mice develop no spontaneous cancer either in skin nor prostate model (Honeycutt et al., 2006; Poliseno et al.). The inability of MCM7 expression alone to initiate carcinogenesis could result from other negative feedback mechanisms that neutralize MCM DNA replication licensing in the cell, such as Rb or ILK signaling discussed later. Even though MCM7's transforming activity is clearly observed, strictly speaking, it does not fall into the

category of a typical proto-oncogene because it does not have a viral counterpart nor sequence mutations that render its oncogenic potential. MCM7 probably falls into the broadly defined oncogene category whose gain of oncogenic function is generated by epigenome or chromosomal numerical alterations.

### 3. MCM7 as target of oncogenic or tumor suppressor signaling pathways

The first significant signaling pathway targeting MCM7 was found by yeast two-hybrid screening analysis where the N-terminus of Rb bait probe binds with MCM7 (Sterner et al., 1998). Additional analysis suggests that other Rb homologues p107 and p130 also bind with MCM7 (Sterner et al., 1998). *In vitro* analysis indicates that the binding of MCM7 and Rb inhibits DNA replication. However, the biological significance of such interaction was not elucidated until 11 years later that interaction of Rb-MCM7 is essential for TGF $\beta$  induced blockade of entry into S phase (Mukherjee et al.). These studies suggest that MCM7 could be the main target of Rb in controlling cycle S phase check point (figure 4). This is because over-expression of MCM7 can reverse Rb inhibition effect, and peptide that interferes with MCM7/Rb binding but not other activity of Rb also reverses the Rb check point blockade. Another salient example of MCM7 as a target of a signaling pathway is androgen receptor signaling. It is well known that androgen receptor regulates cell growth and proliferation. However, most of the studies have been focusing on gene expression regulation which may play secondary role in controlling cell cycle progression. It was found later that AR interacts with MCM7 directly, and inactivates or activates MCM DNA replication licensing depending on the nature of ligands or their concentrations (Shi et al., 2008). Mutation of MCM7 that abrogates its interaction with androgen receptor but not its DNA replication licensing activity eliminates the pleiotropic effect of testosterone. Furthermore, androgen receptor mutant that does not bind with MCM7 can translocate into nucleus upon androgen stimulation but fails to induce cell proliferation or to enhance transcription of androgen dependent genes. It appears that both DNA replication activity and transcription activity of AR is dependent on its binding with MCM7. It is likely that AR serves as a co-replication factor that directs the MCM complex DNA replication licensing through its interaction with MCM7 (figure 4). One surprising finding in the analysis is that MCM7 also serves as a co-transcription factor for AR. There are several well read studies showing that transcription activity enhances DNA replication, and that transcription activity is dependent on DNA replication in eukaryotic cells (Marahrens and Stillman, 1992; Veldman et al., 1985). It remains to be seen whether MCM7 dependent transcription activity holds truth with other MCM7 interacting transcription factors. The MCM7 dependency of androgen receptor transcription activity suggests that androgen dependent gene expression could only occur in actively proliferating cells, since MCM7 is excluded from the nucleus during S, G2 and M phases, and it only re-enters the nucleus in G1 phase.

MCM7 also plays some critical role in mediating the function of cell membrane receptor. This is demonstrated in its interaction with integrin linked kinase (figure 4). It appears that MCM7 is a substrate of integrin linked kinase. The binding and phosphorylation of MCM7 N-terminus by integrin linked kinase reduce the binding of MCM7 with other DNA replication licensing factors, and lead to slow-down of cell growth (Han et al., 2010). The phosphorylation of MCM7 by integrin linked kinase proves to be a critical link to the tumor suppressor activity of integrin  $\alpha 7$ . Dominant negative mutant of integrin linked kinase

interrupts the integrin linked kinase/MCM7 interaction, and partly blocks the tumor suppression activity of integrin  $\alpha 7$ . MCM7 mutant that lacks the integrin linked kinase binding motif operates DNA replication licensing similar to that of wild type but is unresponsive to integrin  $\alpha 7$  signaling. These findings suggest that MCM7 could be the end target of many oncogenic or tumor suppressing signaling pathways. Depending on the nature of these interactions or modification of MCM7, it may lead to increased or decreased DNA replication licensing activity of MCM complex, and guides the cells into either higher level of proliferation or cell growth arrest.

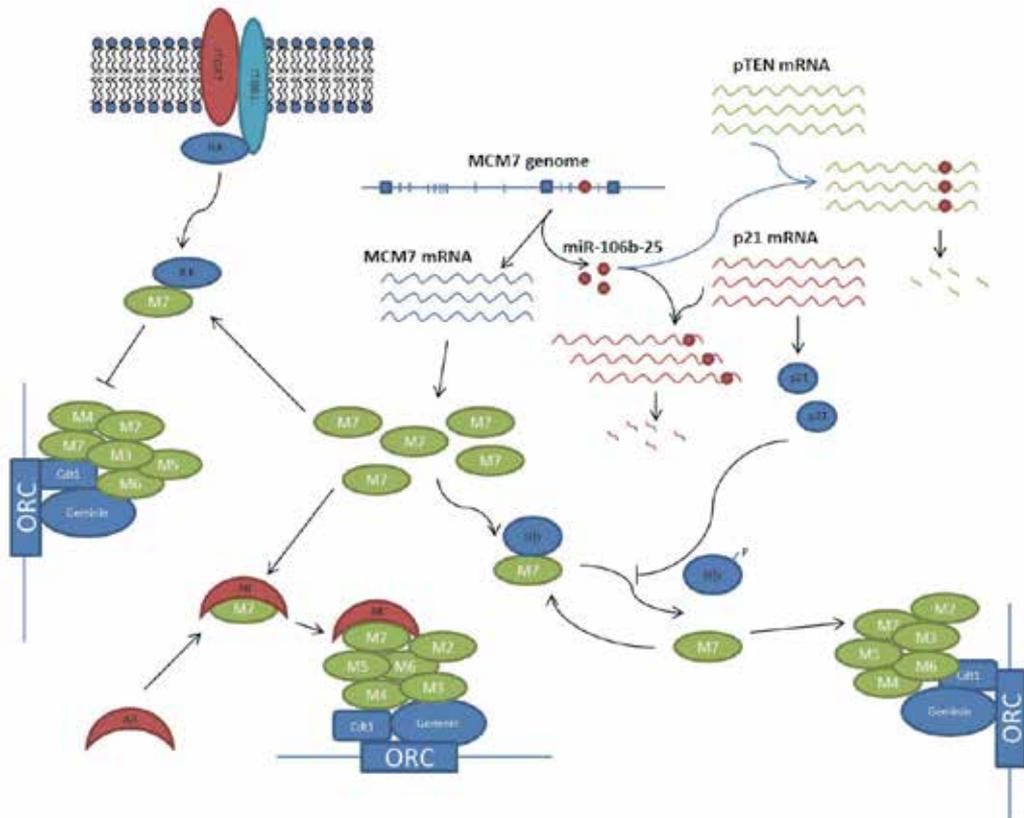


Fig. 4. Diagram of MCM7 oncogenic signaling. Fifteen exons and miR106b-25 oncogenic miRNA cluster of MCM7 genome were transcribed into MCM7 mRNA and miRNA 106b, 93 and 25. Two prong pathways were directed at promoting DNA replication: Degradation of critical tumor suppressor genes p21 and pTEN mRNA by miR106b-25 cluster, and increase of DNA replication licensing by MCM7 protein. Androgen receptor (AR) promotes DNA replication by enhancing MCM7 replication licensing activity. Integrin  $\alpha 7$  inhibits DNA replication through inhibition of MCM7 activity.

#### 4. Transforming miRNA cluster in MCM7 genome

A unique feature of MCM7 genome is that it contains an intronic miRNA miR-106b-25 cluster in intron 13, which includes miR-106b, miR93 and miR-25. miR-25 is highly

homologous to miR-32, an onco-miRNA, while miR-106b and miR-93 belong to the family of miR-17. All three members of miR-106b-25 cluster are abundantly expressed in most of the prostate cancer cell lines. Similar to their host gene MCM7, amplification and up-regulation of all members of miR-106b-25 cluster are found in several human malignancies including esophageal and prostate cancers (Ambs et al., 2008; Kan et al., 2009; Petrocca et al., 2008a; Poliseno et al., ; Sikand et al., 2009). These miRNAs target multiple tumor suppressor genes, and shut down their expression levels (figure 4). One of the most notable examples is pTEN gene expression. miR-25 and miR-93 expression decrease pTEN protein levels, and result in activation of Akt pathway (Poliseno et al.). The expression of miR106-25 cluster increase tumorigenesis in both anchorage-independent assay and xenografted tumors in animal model. The mice with knock-in prostate specific MCM7 and miR-106-25 cluster develop cancer like dysplasia in mice, while mice with pure MCM7 knock in do not seem to develop dysplasia without carcinogen challenge. One may interpret such observation as miR-106-25 cluster function as tumor initiator by knocking down pTEN in mice, while over-expression of MCM7 serves as a mechanism for the development of invasive phenotype. miR-106b-25 cluster appears more versatile than knocking down just one tumor suppressor gene. Several studies suggest that miR-106 and miR-25 target p21, Bim and E2F1 (Ambs et al., 2008; Kan et al., 2009; Petrocca et al., 2008a; Petrocca et al., 2008b). These targets and their relationship have been clearly demonstrated in prostate cancer, esophageal adenocarcinoma and gastric cancer. Inhibition of E2F1 by miR-106b and miR-93 in gastric cancer cell lines partially blocks TGF $\beta$  induced apoptosis (Petrocca et al., 2008b). Interestingly, E2F1 is a strong stimulator of MCM7 locus transcription. As a result, a negative feedback loop of MCM7/miR-106b-25 is formed. Such feedback loop may serve to limit the excessive cell death induced by TGF $\beta$  in physiological condition, and thus achieves a balance of cell growth and cell death. In the event of malignancies, amplification of MCM7 locus may bypass such negative feedback loop, and tips the balance to cell survival. Targeting on p21waf and bim by miR-106b and miR-25 were demonstrated in both esophageal and prostate cancers. Application of inhibitors of miR-106, miR-93 or miR-25 inhibits tumorigenesis both *in vitro* and *in vivo*. With these targets that all appear critical for the oncogenic activity of miR-106b-25 cluster, it is not easy to analyze the contribution of each of these pathways to carcinogenesis since there is a lack of experiments to block off each of the pathways to offset the oncogenic activity of miR-106b-25 cluster. Generally speaking, miR-25/93-pTEN pathway is compelling because of the inverse correlation of pTEN expression and miR-106b-25 miRNA levels in primary and mouse tumor samples, the animal model initiating prostate cancer with MCM7/miR-106b-25 cluster transgene showing clear down-regulation of pTEN expression and concomitant increased level of miR-106b-25 cluster expression, and consequent Akt pathway activation. It would be of interest to see how much oncogenic phenotype will be reversed if these animals are transgened with pTEN, or p21 or BIM or E2F1 gene construct lacking the miR-106b-25 target sequences in their 3' untranslated regions.

## 5. Potential therapeutic target

MCM7 is essential for any cell that undergoes proliferation. This poses a dilemma for MCM7 gene targeted therapy. Nevertheless, a report indicates that shRNA targeting MCM7 in xenografted PC3 and DU145 tumors in mice dramatically reduces tumor volume, rate of metastasis and fatality (Shi et al., 2010). The drawback of this analysis is that the shRNA

target may not recognize the MCM7 sequence from mice since the target sequence is intended for human MCM7. Nevertheless, the study is a proof of principle that MCM7 knockdown is a potential effective approach in combating prostate cancer, particularly those with MCM7 amplification and over-expression. In light of double oncogenic effect of MCM7 genome cluster, it may be more effective if a shRNA is designed to neutralize the effect of miR-106b, miR-93 and miR-25 onco-miRNA. This can be accomplished with large dosage of morpholino oligonucleotide specific for these onco-miRNAs. These RNase resistant oligonucleotides have a clear advantage over genome approach that it will not interfere with the genome structure of a cell but have long lasting presence once they are taken. Furthermore, there is no evidence suggesting that knocking down of these miRNAs adversely affects the survival of normal cells.

## 6. Abbreviation

M7-MCM7; M6-MCM6; M5-MCM5; M4-MCM4; M3-MCM3; M2-MCM2; ITGA7- Integrin  $\alpha$ 7; ILK-integrin linked kinase; AR-androgen receptor; ORC-origin recognition complex.

## 7. Acknowledgement

This work was supported by grants from National Cancer Institute (RO1 CA098249 and R56 CA098249 to JHL), American Cancer Society (RSG-08-137-01-CNE to YPY) and Department of Defense (W81 XWH-09-1-0376 to JHL).

## 8. References

- Ambs, S., Prueitt, R.L., Yi, M., Hudson, R.S., Howe, T.M., Petrocca, F., Wallace, T.A., Liu, C.G., Volinia, S., Calin, G.A., *et al.* (2008). Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. *Cancer research* 68, 6162-6170.
- Brake, T., Connor, J.P., Petereit, D.G., and Lambert, P.F. (2003). Comparative analysis of cervical cancer in women and in a human papillomavirus-transgenic mouse model: identification of minichromosome maintenance protein 7 as an informative biomarker for human cervical cancer. *Cancer research* 63, 8173-8180.
- Chong, J.P., Thommes, P., and Blow, J.J. (1996). The role of MCM/P1 proteins in the licensing of DNA replication. *Trends Biochem Sci* 21, 102-106.
- Coxon, A., Maundrell, K., and Kearsley, S.E. (1992). Fission yeast *cdc21+* belongs to a family of proteins involved in an early step of chromosome replication. *Nucleic acids research* 20, 5571-5577.
- Dalton, S., and Whitbread, L. (1995). Cell cycle-regulated nuclear import and export of Cdc47, a protein essential for initiation of DNA replication in budding yeast. *Proc Natl Acad Sci U S A* 92, 2514-2518.
- Han, Y.C., Yu, Y.P., Nelson, J., Wu, C., Wang, H., Michalopoulos, G.K., and Luo, J.H. (2010). Interaction of integrin-linked kinase and miniature chromosome maintenance 7-mediating integrin  $\{\alpha\}7$  induced cell growth suppression. *Cancer research* 70, 4375-4384.

- Honeycutt, K.A., Chen, Z., Koster, M.I., Miers, M., Nuchtern, J., Hicks, J., Roop, D.R., and Shohet, J.M. (2006). Deregulated minichromosomal maintenance protein MCM7 contributes to oncogene driven tumorigenesis. *Oncogene* 25, 4027-4032.
- Ishimi, Y. (1997). A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *The Journal of biological chemistry* 272, 24508-24513.
- Kan, T., Sato, F., Ito, T., Matsumura, N., David, S., Cheng, Y., Agarwal, R., Paun, B.C., Jin, Z., Oлару, A.V., *et al.* (2009). The miR-106b-25 polycistron, activated by genomic amplification, functions as an oncogene by suppressing p21 and Bim. *Gastroenterology* 136, 1689-1700.
- Kearsey, S.E., Maiorano, D., Holmes, E.C., and Todorov, I.T. (1996). The role of MCM proteins in the cell cycle control of genome duplication. *Bioessays* 18, 183-190.
- Marahrens, Y., and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science (New York, NY)* 255, 817-823.
- Mukherjee, P., Winter, S.L., and Alexandrow, M.G. Cell cycle arrest by transforming growth factor beta1 near G1/S is mediated by acute abrogation of prereplication complex activation involving an Rb-MCM interaction. *Molecular and cellular biology* 30, 845-856.
- Petrocca, F., Vecchione, A., and Croce, C.M. (2008a). Emerging role of miR-106b-25/miR-17-92 clusters in the control of transforming growth factor beta signaling. *Cancer research* 68, 8191-8194.
- Petrocca, F., Visone, R., Onelli, M.R., Shah, M.H., Nicoloso, M.S., de Martino, I., Iliopoulos, D., Pilozi, E., Liu, C.G., Negrini, M., *et al.* (2008b). E2F1-regulated microRNAs impair TGFbeta-dependent cell-cycle arrest and apoptosis in gastric cancer. *Cancer cell* 13, 272-286.
- Poliseno, L., Salmena, L., Riccardi, L., Fornari, A., Song, M.S., Hobbs, R.M., Sportoletti, P., Varmeh, S., Egia, A., Fedele, G., *et al.* Identification of the miR-106b~25 microRNA cluster as a proto-oncogenic PTEN-targeting intron that cooperates with its host gene MCM7 in transformation. *Science signaling* 3, ra29.
- Ren, B., Yu, G., Tseng, G.C., Cieply, K., Gavel, T., Nelson, J., Michalopoulos, G., Yu, Y.P., and Luo, J.H. (2006). MCM7 amplification and overexpression are associated with prostate cancer progression. *Oncogene* 25, 1090-1098.
- Shi, Y.K., Yu, Y.P., Tseng, G.C., and Luo, J.H. (2010). Inhibition of prostate cancer growth and metastasis using small interference RNA specific for minichromosome complex maintenance component 7. *Cancer gene therapy* 17, 694-699.
- Shi, Y.K., Yu, Y.P., Zhu, Z.H., Han, Y.C., Ren, B., Nelson, J.B., and Luo, J.H. (2008). MCM7 Interacts with Androgen Receptor. *The American journal of pathology* 173, 1758-1767.
- Sikand, K., Slane, S.D., and Shukla, G.C. (2009). Intrinsic expression of host genes and intronic miRNAs in prostate carcinoma cells. *Cancer cell international* 9, 21.
- Sterner, J.M., Dew-Knight, S., Musahl, C., Kornbluth, S., and Horowitz, J.M. (1998). Negative regulation of DNA replication by the retinoblastoma protein is mediated by its association with MCM7. *Molecular and cellular biology* 18, 2748-2757.
- Veldman, G.M., Lupton, S., and Kamen, R. (1985). Polyomavirus enhancer contains multiple redundant sequence elements that activate both DNA replication and gene expression. *Molecular and cellular biology* 5, 649-658.
- You, Z., Komamura, Y., and Ishimi, Y. (1999). Biochemical analysis of the intrinsic Mcm4-Mcm6-mcm7 DNA helicase activity. *Molecular and cellular biology* 19, 8003-8015.

# Regulation of Minichromosome Maintenance (MCM) Helicase in Response to Replication Stress

Faria Zafar and Takuro Nakagawa

*Department of Biological Sciences, Graduate School of Science, Osaka University  
Japan*

## 1. Introduction

Progression of DNA replication forks is hindered by various reasons such as DNA-binding proteins, DNA damage, and dNTP depletion. The minichromosome maintenance (MCM) complex, which comprises of Mcm2-7 subunits, is the DNA helicase that plays the central role in the progression of replication forks. MCMs are loaded onto specific sites of chromosomes called origins to create pre-replication complexes (pre-RCs). At the onset of the S-phase, MCM forms a complex with GINS and Cdc45, and starts unwinding the double-stranded DNA providing template strands to polymerases. Even after the fork encounters replication stress, the DNA unwinding continues to some extent and causes an extended length of single-stranded DNA to be exposed, which in turn induces the activation of checkpoint pathways. However, unregulated translocation of MCM may cause re-annealing of two complementary single-strands behind the fork and destabilize the stalled replication machinery. Thus, DNA helicase activity needs to be tightly regulated under the stress conditions in order to preserve the replication machinery. It appears that the proteins associated with the replication forks including Mrc1/Claspin, Tof1/Swi1/Tim1, Csm3/Swi3/Tipin, Ctf4/Mcl1/AND-1 mediate the interaction between the MCM helicase and DNA polymerases, and are required for coupling of DNA unwinding with DNA synthesis. The evolutionally conserved C-terminal domain (CTD) of Mcm4 is also involved in the regulation of the MCM helicase activity. The checkpoint kinase, cyclin-dependent (CDK), Dbf4-dependent kinase (DDK) and Polo-like kinase (PLK) kinase phosphorylate and regulate the function of the MCM helicase. Here we review the recent findings in regard to the MCM regulation in response to replication stress, and discuss how DNA synthesis and unwinding are coordinated to maintain the genome integrity.

## 2. Formation of the replication machineries on eukaryotic chromosomes

### 2.1 Assembly of pre-Replication Complex (pre-RC)

A single round of replication of chromosomal DNA in every cell cycle is important for faithful transmission of the genetic information to daughter cells. G1-phase of the cell cycle is marked as the growth phase, which prepares the cells for the S-phase (replication) (Pardee, 1989). The initiation sites of DNA replication on chromosomes are called replication

origins. A series of proteins that are specifically required for chromosome replication, shown in Table 1, are assembled on each origin for the initiation. Essentially, all of the replication proteins are conserved from yeasts to humans, making yeast an excellent model system to understand the molecular mechanisms of DNA replication. The origin recognition complex (ORC), which consists of Orc1-6 subunits, has been identified as the protein complex that binds to the origin DNA in an ATP-dependent manner (Bell & Stillman, 1992). The minichromosome maintenance (MCM) complex, that forms a ring-like structure consisting of Mcm2-7 subunits each of which contains the AAA+ motif (Neuwald et al, 1999), has been originally identified by a genetic screening for the budding yeast mutants that affect autonomous replication sequence (ARS) activity on episomal plasmids (Maine et al, 1984). MCM proteins were independently identified from *Xenopus* egg extracts as the factor that binds to chromosomes, licensing the initiation of DNA replication (Blow & Laskey, 1986; Kubota et al, 1995). In 1997, DNA helicase activity associated with MCM was first reported by using Mcm4-6-7 sub-complex in an *in vitro* assay (Ishimi, 1997). Using conditional degron system, it has been shown that MCM is essential for the initiation as well as the elongation phases of DNA replication (Labib et al, 2000). From late M to G1 phase, MCMs are loaded onto the origins depending on Cdt1 and Cdc6, resulting in the formation of pre-replicative complexes (pre-RCs) (Diffley et al, 1994; Donovan et al, 1997; Liang et al, 1995; Nishitani et

<b>Budding yeast</b>	<b>Fission yeast</b>	<b>Frog</b>	<b>Mammal</b>
<b>pre-RC components</b>			
Orc1-6 (ORC)	Orc1-6 (ORC)	Orc1-6 (ORC)	Orc1-6 (ORC)
Mcm2-7 (MCM)	Mcm2-7 (MCM)	Mcm2-7 (MCM)	Mcm2-7 (MCM)
Cdc6	Cdc18	Cdc6	Cdc6
Cdt1	Cdt1	Cdt1	Cdt1
<b>Initiation factors</b>			
Sld2	Drc1	RTS/RecQ4	RTS/RecQL4
Sld3	Sld3	Treslin/Ticrr	Treslin
Dpb11	Cut5/Rad4	TopBP1/Cut5/Mus101	TopBP1
<b>Initiation and elongation factors</b>			
Sld5-Psf1-Psf2-Psf3 (GINS)	Sld5-Psf1-Psf2-Psf3 (GINS)	Sld5-Psf1-Psf2-Psf3 (GINS)	Sld5-Psf1-Psf2-Psf3 (GINS)
Cdc45/Sld4	Cdc45/Sna11	Cdc45	Cdc45
Mcm10/Dna43	Mcm10/Cdc23	Mcm10	Mcm10
<b>Luxury components in replication fork</b>			
Tof1	Swi1	TIM1	Tim1
Csm3	Swi3	TIPIN	Tipin
Mrc1	Mrc1	CLASPIN	Claspin/CLSPN
Ctf1	Mcl1	AND-1	AND-1

Table 1. DNA replication proteins are conserved from yeasts to humans.

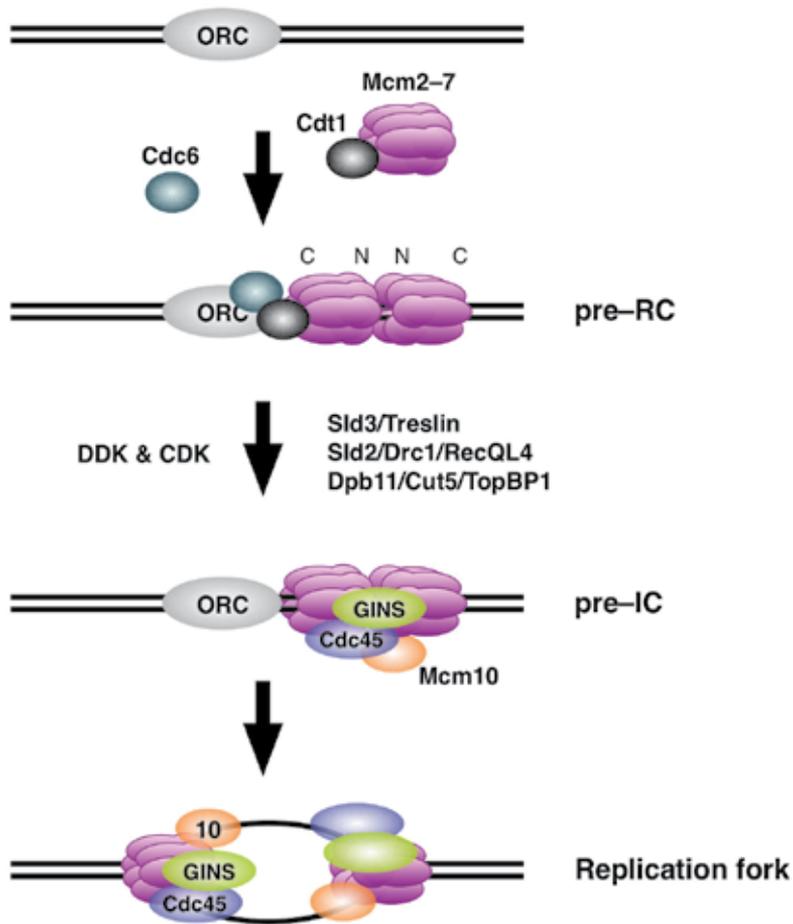


Fig. 1. Stepwise formation of DNA replication forks in eukaryotes.

al, 2000; Remus et al, 2009). The pre-RC is an important landing pad for the replication initiation, because it is prerequisite for the assembly of other factors that are essential for replication. Being connected via their N-terminal domains, head-to-head double hexamers of Mcm2-7 are stably and exclusively formed on the origin DNA (Evrin et al, 2009; Gambus et al, 2011; Remus et al, 2009). Each of the double hexamers may be involved in bi-directional replication from a single initiation site. Although DNA replication initiates from the single replication origin *oriC* in *E. coli* (Kaguni et al, 1982), there are large number of origins where pre-RCs are assembled onto each chromosome of eukaryotes. In both budding and fission yeasts, there are hundreds of origins per genome of about 14-Mb (Hayashi et al, 2007; Wyrick et al, 2001). Under the unperturbed condition, replication starts from only a subset of these origins. The dormant origins serve as the replication initiation sites under the stressed conditions (see below). In multicellular organisms, the origins that fire in early S phase differ during development or between different cell types (Goldman et al, 1984). It is also important for normal development that many pre-RCs are produced in the genome,

because mutations in ORC, Cdc6, or Cdt1 have been found in the patients of Meier-Gorlin syndrome (Bicknell et al, 2011a; Bicknell et al, 2011b; Guernsey et al, 2011), which is a rare autosomal recessive genetic condition whose primary clinical hallmarks include small stature, small external ears and small or absent patellae.

## 2.2 Assembly of pre-Initiation Complex (pre-IC)

The choice of the origins to be used is determined by the loading of additional replication proteins onto pre-RCs to form pre-initiation complexes (pre-ICs). Among the components of pre-IC, it seems that Sld2/Drc1/RecQ4/RecQL4 (Kamimura et al, 1998; Sangrithi et al, 2005; Wang & Elledge, 1999), Sld3/Treslin (Kamimura et al, 2001; Kumagai et al, 2010; Nakajima & Masukata, 2002; Sansam et al, 2010), and Dpb11/Cut5/Rad4/TopBP1 (Araki et al, 1995; Saka & Yanagida, 1993; Yamane et al, 1997) are essential for the initiation but not for the elongation phase of replication. Two classes of protein kinase that are active in the S phase: Dbf4-dependent kinase (DDK) and cyclin-dependent kinase (CDK) play important roles in the assembly of pre-IC and ensure that only a single round of DNA replication takes place in each cell cycle. Mcm2, 3, 4, and 6 subunits can be phosphorylated by DDK *in vitro* (Lei et al, 1997), and the phosphorylation of Mcm4 appears to be critical for loading of Cdc45 onto pre-RCs (Masai et al, 2006; Sheu & Stillman, 2006; Sheu & Stillman, 2010). In fission yeast, loading of Sld3 onto origins occurs in a DDK-dependent but CDK-independent manner (Yabuuchi et al, 2006). And, Sld3 and DDK are required for the loading of Cdc45 as well as GINS complex (Yabuuchi et al, 2006; Yamada et al, 2004). Thus, it is tempting to argue that the loading of Sld3 is mediated by the phosphorylation of MCM proteins by DDK. It has been established that Sld2 and Sld3 are the two major substrates of CDK-dependent phosphorylation that are essential for replication in budding yeast (Masumoto et al, 2002; Tanaka et al, 2007; Zegerman & Diffley, 2007). The phosphorylation of Sld2 and Sld3 enhances their interaction with Dpb11 that contains multiple copies of the BRCT motif, which is the phosphopeptide-binding module found in many other proteins including BRCA2 (Glover et al, 2004). As both Sld2 and Sld3 are essential for stable association of Dpb11 to replication origins, the phosphorylation-mediated interaction between them is required for the Dpb11 loading onto pre-RCs. The CDK-mediated regulation for the assembly of Sld2 and Sld3 with the BRCT-motif-containing protein, Dpb11/Cut5/TopBP1 appears to be an evolutionally conserved mechanism, as the fission yeast homolog of Sld2, Drc1 is phosphorylated by Cdc2 kinase, and the Drc1 phosphorylation appears to be important for interaction with the homolog of Dpb11, Cut5/Rad4 in a CDK-dependent manner (Fukuura et al, 2011; Noguchi et al, 2002). Furthermore, it has been shown that the mammalian homolog of Sld3, Treslin associates with TopBP1 in a Cdk2-dependent manner and is essential for the initiation of DNA replication (Kumagai et al, 2010). Dpb11 forms a complex with DNA polymerase  $\epsilon$  (Masumoto et al, 2000), and is required for the loading of the Sld5-Psf1-Psf2-Psf3 (GINS) complex (Kanemaki et al, 2003; Kubota et al, 2003; Takayama et al, 2003) and Cdc45 (Muramatsu et al, 2010) onto pre-RCs. Mcm10 protein also binds to subunits of Mcm2-7 complex, and is essential for the initiation and elongation of replication, although its mechanism of action remains elusive (Aves et al, 1998; Heller et al, 2011; Izumi et al, 2000; Solomon et al, 1992). In the S phase, MCM proteins and Cdc45 dissociate from origin and travel along DNA, concerted with DNA polymerase  $\epsilon$  (Aparicio et al, 1997). The traveling along DNA has also been observed for GINS (Kanemaki et al, 2003; Takayama et al, 2003). Purification of proteins from yeast cells in the S phase identified the complexes

containing Cdc45, GINS and Mcm2-7 (Gambus et al, 2006). *Xenopus* Cdc45, GINS and Mcm2-7 were also identified at the replication fork (Pacek et al, 2006). In vitro experiments showed that the Cdc45/Mcm2-7/GINS (CMG) complexes purified from *Drosophila* cells exhibit robust DNA unwinding activity, and that the CMG complex translocates on DNA in a 3' to 5' direction (Moyer et al, 2006). Association with Cdc45 and GINS enhances ATP hydrolysis, DNA binding and the helicase activity of Mcm2-7 (Ilves et al, 2010). These studies show that the CMG complex functions as the DNA unwinding complex in the replication machinery. Replication may begin such that two forks originate from a single origin moving in opposite directions (bi-directionally) behind the CMG complex, which unwinds DNA to provide single-stranded DNA (ssDNA) templates to the polymerases for duplication. To sum up, CMG complex is the motor of the replication fork that unwinds double-stranded DNA (dsDNA), providing template ssDNA for polymerases that synthesize new strands. Therefore, it is important to understand the formation and the regulation of the CMG complex to elucidate the molecular mechanism of DNA replication.

### 3. Blockage to the progression of DNA replication forks

DNA replication forks may encounter certain impediments such as DNA damage, dNTP depletion, the proteins that tightly bind to DNA, or epigenetic status of nucleosomes. In order to replicate the entire genome within a limited period of time, such aforementioned blocks are necessary to be removed or tolerated. Failures to respond to these replication fork blockages lead to genome instability such as gross chromosomal rearrangements (GCRs), consequently leading to cell death or genetic diseases including cancer in multicellular organisms.

#### 3.1 Inhibitors to the replication fork progression

Replication is an eventful process liable to encounter odds during its procession. Various exogenous substances have been recognized to obstruct this process in their own specific mode of action. Hydroxyurea (HU) is a specific inhibitor of the enzyme ribonucleotide reductase (RNR) that is essential for production of deoxyribonucleotides (dNTPs) (Young et al, 1967). Thus, treatment of cells with HU inhibits DNA synthesis by depleting dNTPs. Aphidicolin is a tetracyclic diterpene tetraol, obtained from *C. aphidicola* and certain other fungi (Bucknall et al, 1973). It is a specific and direct inhibitor of DNA polymerase  $\alpha$  and also  $\delta$ , two of the three DNA polymerases (i.e. DNA Pol  $\alpha$ ,  $\delta$ , and  $\epsilon$ ) that are essential for chromosomal DNA replication (Goscin & Byrnes, 1982; Ikegami et al, 1978). Thus, treatment of cells with either HU or aphidicolin inhibits the progression of replication forks. Chemical adducts on DNA have the potential to impede the fork progression. Methyl methanesulfonate (MMS) is one such well-characterized DNA alkylating agent. The predominant adduct in dsDNA resulting from MMS exposure is 7-methylguanine (N7-meG) followed by 3-methyladenine (N3-meA) (Lawley & Brookes, 1963). The methylation of DNA physically impedes the progression of replication forks, leading to the formation of DNA double-strand breaks (DSBs) during S phase (Groth et al, 2010). Camptothecin (CPT) is a specific inhibitor of DNA topoisomerase I (Top1) capable of removing DNA supercoils during replication as well as transcription. CPT blocks both the DNA cleavage and religation reactions of Top1 (Kjeldsen et al, 1992). Inhibition of the initial cleavage step leads to the accumulation of supercoils ahead of the replication fork that induces potentially lethal

DNA lesions (Koster et al, 2007). Blocking the rejoining step accumulates the reaction intermediates in which Top1 is covalently attached to the end of DNA (Hsiang et al, 1985). When replication forks reach these DNA nicks, they will be converted to one-ended DSBs. CPT analogues have significant activity against solid tumours, and have gained US Food and Drug Administration approval for the treatment of ovarian and lung cancer (Stewart, 2004). DNA interstrand cross-links (ICLs) also block the fork progression, by preventing the two DNA strands from separating. Antitumor drugs such as cisplatin, psoralen, nitrogen mustard or mitomycin C as well as endogenous agents formed by lipid peroxidation induce ICLs (Niedernhofer et al, 2003). There are different types of chemicals that block the progression of replication forks, and they do so by distinct mechanisms.

### 3.2 Replication fork pausing sites on chromosomes

Progression of replication forks is affected either by protein-DNA complexes or the torsional stress around the fork. It has been estimated that there are >1,000 discrete sites in the budding yeast genome that impede normal fork progression, including tRNA and rDNA genes, dormant origins, silent mating-type loci, centromeres and telomeres (Ivessa et al, 2003). Non-nucleosomal and nucleosomal protein-DNA complexes impede the fork progression (Deshpande & Newlon, 1996; Ivessa et al, 2003). Replication fork barrier (RFB) present in the rDNA gene is one of the well-characterized pausing sites in the genome. The RFBs block fork progression in an orientation dependent manner. The 3' end of the 35S rRNA transcription unit in rDNA acts as a barrier to replication forks moving in the direction opposite to RNA polymerase I (Brewer & Fangman, 1988). The arrest of replication forks at the RFB site occurs independently of transcription but is mediated by the Fob1 protein that binds to the specific sequence of ~100 bp in the RFB site (Brewer et al, 1992; Kobayashi et al, 1992; Kobayashi & Horiuchi, 1996). However, the Fob1 binding to the RFB site is not sufficient for the transcription-independent fork arrest (Calzada et al, 2005). Tof1, Csm3, and Mrc1 are the luxury members of the replication machinery that are associated with the CMG complex (Katou et al, 2003). Among them Tof1 and Csm3 are required for the pausing at the RFB in rDNA of budding yeast (Calzada et al, 2005). The fission yeast homologs, Swi1 and Swi3, are also required for the fork arrest at the RFB site in rDNA and at another orientation-dependent fork barrier site present in the mating-type switch locus, RTS1 (Dalgaard & Klar, 2000). Thus, it is possible that the proteins that bind to the barrier sites are not simple obstacles to the fork movement rather they negatively regulate the CMG activity via a specific interaction with the fork components.

In the absence of RFB in the rDNA locus, DNA and RNA syntheses simultaneously advance on the same template DNA, the 35S rRNA gene, making a collision between oppositely traveling replication and transcription machineries. When transcription occurs frequently, this collision slows down the progression of replication forks (Takeuchi et al, 2003). The fork pausing by the collision between transcription and replication machineries is also observed in other regions of chromosomes. Some of the highly transcribed tRNA genes and RNA polymerase II genes also impede the progression of replication forks. Genome-wide analyses of DNA Pol2-binding sites, which are indicative of the position of replication forks, showed that the fork arrest takes place regardless of whether replication and transcription move in the same or opposite directions (Azvolinsky et al, 2009). Direction-independent collision was confirmed by similar experiments that mapped the position of the Psf2 subunit of the GINS complex (Sekedat et al, 2010), suggesting that replication forks arrest at highly transcribed genes because the transcription and replication machineries are not allowed to occupy the same DNA at the same time.

DNA supercoils produced ahead of advancing replication forks are resolved by a coordinated action of Top1 and Top2. In the absence of functional topoisomerases, the replication fork-related topological constraints are accumulated, leading to fork collapse and DNA damage checkpoint activation (Bermejo et al, 2007; Brill et al, 1987; Kim & Wang, 1989). Behind the replication fork, intertwining of sister chromatid DNA takes place. Top2 activity is essential for the separation of intertwined chromosomal DNA molecules before the onset of anaphase (Holm et al, 1985). Expression of catalytically inactive Top2 prevents completion of DNA replication and induces DNA damage checkpoint response (Baxter & Diffley, 2008), suggesting that the processing of catenation of sister chromatids behind the fork may also affect the fork progression.

The replication pausing sites are chromosome fragile sites.  $\gamma$ -H2A or  $\gamma$ -H2AX is one of the well-characterized histone modifications that occur around DNA damage sites in the checkpoint kinase dependent manner (Downs et al, 2000). By using DNA microarray, recent studies mapped the localization of  $\gamma$ -H2A on budding or fission yeast genomes in unperturbed S phase (Rozenzhak et al, 2010; Szilard et al, 2010). In budding yeast, they found the accumulation of  $\gamma$ -H2A to occur in repressed genes and that is dependent on the activity of a histone deacetylase (HDAC). In fission yeast,  $\gamma$ -H2A and the Brc1 protein that recognizes  $\gamma$ -H2A through a pair of BRCT domains were localized at heterochromatin regions of chromosomes such as silent-mating type loci, centromeres and telomeres. They also showed that the S-phase specific localization of  $\gamma$ -H2A and Brc1 is dependent on Ctr4, which is responsible for the methylation of histone H3 9<sup>th</sup> Lys. Thus, HDAC- and/or H3K9<sup>me</sup>-mediated heterochromatin seems to impair the stability of replication forks when they pass through, although how the chromatin status affects the fork progression remains to be elucidated.

#### 4. Checkpoint response to replication problems

Replication problems such as replication fork stalling or collapse are detected by the surveillance system called checkpoint. The replication checkpoint pathway is activated in response to replication stalling, while the DNA damage checkpoint will be activated when DSBs are formed. These checkpoint pathways temporarily halt the cell cycle progression, giving time for cells to solve the replication problems before entering into mitosis.

##### 4.1 Checkpoint activation in response to replication problems

Extended lengths of ssDNA formed at stalled forks is the key DNA structure that induces the replication checkpoint response. The ssDNA never exists naked in vivo and is immediately coated by ssDNA-binding proteins: replication protein A (RPA). The ssDNA-RPA complex recruits the most upstream checkpoint kinase the ATR-ATRIP complex in mammals (Choi et al, 2010; Zou & Elledge, 2003). Primer-template junctions present right next to the ssDNA-RPA complex in the context of stalled forks are also important for the ATR activation (MacDougall et al, 2007). The proliferating cell nuclear antigen (PCNA)-related checkpoint clamp, the Rad9-Rad1-Hus1 (9-1-1) complex is recruited to stalled forks with the aid of the checkpoint-specific clamp loader (Bermudez et al, 2003; Ellison & Stillman, 2003; Zou et al, 2002). Intriguingly, the essential replication initiation protein Dpb11/Cut5/TopBP1 is again recruited to the stalled fork depending on the protein phosphorylation induced by the ATR kinase, and is required for the activation of replication checkpoint (Delacroix et al, 2007; Furuya et al, 2004; Kumagai et al, 2006; Lee et al, 2007),

suggesting a mechanistic similarity between the replication initiation and the checkpoint activation. Tof1/Swi1, Csm3/Swi3, and Mrc1 form the protein complex associated with the CMG complex (Bando et al, 2009; Katou et al, 2003; Noguchi et al, 2004; Shimmoto et al, 2009). Tof1/Swi1 and Csm3/Swi3 form a relatively stable complex and are required for the association of Mrc1 with the replication fork. Although this fork protection complex is dispensable for DNA replication, it is required for the activation of the replication checkpoint effector kinase such as fission yeast Cds1 (Alcasabas et al, 2001; Murakami & Okayama, 1995; Noguchi et al, 2003; Tanaka & Russell, 2001; Unsal-Kacmaz et al, 2007), indicating that the substrate specificity of the most upstream checkpoint kinase is modulated by the proteins associated with the CMG complex.

As mentioned above, the extension of single-stranded region at times of replication stress is crucial to induce a cascade of checkpoint responses. Even after DNA polymerases come to a halt, the CMG complex continues to execute its job as a helicase to a considerable length to produce the single-strand region sufficient to activate the checkpoint. Inhibition of Cdc45 or Mcm7 after the replication initiation blocks accumulation of RPA on chromatin and the checkpoint activation (Byun et al, 2005). Thus, DNA unwinding activity of the CMG complex is required not only for DNA synthesis under unperturbed condition, but also for the activation of the replication checkpoint at times of replication stress. The interactions between the CMG complex and the checkpoint proteins may also contribute to the activation of replication checkpoint (Tsao et al, 2004). There is an intimate link between the CMG complex and the replication checkpoint. However, when stalled forks are collapsed to create DSBs, the damage checkpoint will be activated. The Mre11-Rad50-Nbs1/Xrs2 (MRN/MRX) complex recognizes DSBs and leads to the activation of the damage checkpoint kinase ATM/Tel1 (Lee & Paull, 2004; Lee & Paull, 2005; Usui et al, 2001).

#### **4.2 Checkpoint response to replication stress**

In addition to the cell cycle regulation, the activated checkpoint pathways regulate gene expression through modification of transcription factors (Huang et al, 1998), and inhibit the replication initiation from late origins (Santocanale & Diffley, 1998; Shirahige et al, 1998). The inhibition of late origins is mediated by the checkpoint kinase-dependent phosphorylation of two of the essential replication initiators, DDK and Sld3 (Lopez-Mosqueda et al, 2010; Zegerman & Diffley, 2010). This output of the checkpoint response may contribute to the genome integrity by preventing the formation of additional number of stalled forks (Lopez-Mosqueda et al, 2010). The replication checkpoint kinases prevent stalled replication forks from breaking down (Desany et al, 1998; Lopes et al, 2001; Tercero & Diffley, 2001). In fission yeast, the Rad3 kinase is activated in response to replication stress, and it phosphorylates the downstream Cds1 kinase, which in turn phosphorylates various downstream target proteins including the structure-specific nuclease Mus81 (Boddy et al, 2001; Kai et al, 2005). The phosphorylation of Mus81 by the Cds1 kinase results in dissociation of Mus81 from chromatin, preventing it from cleaving the stalled fork. Another example of the checkpoint target is Rad60, which is required for recombinational repair probably through the regulation of the Smc5/6 complex (Miyabe et al, 2006; Morishita et al, 2002). The checkpoint kinase phosphorylates Rad60, leading to its delocalization from the nucleus after replication stress (Boddy et al, 2003; Miyabe et al, 2009). These observations are consistent to the idea that the replication checkpoint prevents recombinational repair to occur as long as the stalled forks are able to keep the checkpoint activated. Intriguingly, the

MCM proteins are phosphorylated in response to replication stress through the checkpoint kinases (Bailis et al, 2008; Cortez et al, 2004b; Ishimi et al, 2003; Randell et al, 2010; Yoo et al, 2004b). The phosphorylation of the MCM proteins may increase the stability of the complex in the stalled condition (Randell et al, 2010).

## 5. Regulation of the replication initiation to maintain the genome integrity

Many pre-RCs are assembled on each chromosome in late M to G1 phases, but replication initiates recurrently from only a subset of them leaving the rest of them dormant. However, when replication forks are collapsed, replication initiates from the nearby dormant origins (Doksani et al, 2009), indicating the importance of the dormant origins under the stressed condition. Even under the normal growth condition, the damage checkpoint operates to stabilize a chromosome when all or almost all of the efficient replication origins on the chromosome are deleted (Theis et al, 2010). It is likely that during the long travel on DNA, replication forks spontaneously collapse to create DSBs. Consistent to this idea, decreasing the replication initiation by partial inhibition of the assembly of pre-RCs causes gross chromosomal rearrangements (GCRs) (Tanaka & Diffley, 2002). The most common fragile site in human lymphocytes FRA3B is caused by a paucity of replication initiation events in that region (Letessier et al, 2011). A hypomorphic mutation of *Mcm4<sup>Chaos3</sup>*, that decreases the *Mcm2-7* protein levels, in mice causes various types of chromosome instability and shows predisposition to cancer (Chuang et al, 2010; Kawabata et al, 2011; Shima et al, 2007). These findings demonstrate that the number and the distribution of replication initiation along a chromosome are important for maintaining the genome stability.

Polo-like kinase (PLK) is involved in various important cellular events such as regulation of mitotic entry, chromosome segregation, centrosome maturation, and mitotic exit. PLK contains two tandem Polo boxes, termed as the Polo-box domain that interacts with phosphoproteins (Elia et al, 2003). The *Xenopus* PLK homolog *Plx1* binds to Claspin in a manner dependent on the phosphorylation of Claspin by the checkpoint kinase (Yoo et al, 2004a). The *Plx1* phosphorylates Claspin and causes its dissociation from chromatin, resulting in the inactivation of the replication checkpoint kinase *Chk1* after a prolonged checkpoint arrest. The PLK homolog in budding yeast *Cdc5* is also required for the down regulation of the replication checkpoint: the adaptation, that is the resumption of the cell cycle in the presence of a single unrepaired DSB after a prolonged arrest (Donnianni et al, 2010; Toczyski et al, 1997; Vidanes et al, 2010). These findings indicate that PLK has an inhibitory effect on the checkpoint response. Members of the MCM helicase are phosphorylated by the checkpoint kinases, suggesting a regulation of the MCM helicase by the ATR/ATM checkpoint (Cortez et al, 2004a; Yoo et al, 2004b). Mammalian PLK binds to the *Mcm* proteins in the phosphorylation dependent manner (Lowery et al, 2007; Tsvetkov & Stern, 2005). In *Xenopus*, the phosphorylation of *Mcm2* by ATR stimulates the interaction with *Plx1*, probably recruiting *Plx1* to the damage sites (Trenz et al, 2008). Although *Plx1* is dispensable for DNA replication under the normal condition, it is essential to complete DNA replication when there are only a limited number of pre-RCs assembled on chromatin or in the presence of a low dose of replication inhibitors such as aphidicolin. Thus, it seems that, even though the overall level of the initiation of DNA replication is prohibited in the presence of replication problem, the inhibitory effect caused by the checkpoint is relieved by the Polo-like kinases so that replication initiation occurs from the dormant origins nearby the stalled or collapsed fork.

## 6. Coupling DNA unwinding with DNA synthesis under the stressed condition

In bacteria, the replication fork DNA helicase (DnaB) and the replicative DNA polymerase (Pol III holoenzyme) are associated with each other, and the interaction is essential for a high rate of replication fork movement of about 1,000 nt/s (Kim et al, 1996). In eukaryotes, the CMG complex unwinds template DNA strands and specific DNA polymerases (Pol  $\alpha$ ,  $\delta$  and  $\epsilon$ ) synthesize their respective leading and lagging strands. When the polymerase function is interfered by replication stress, the CMG complex continues to unwind dsDNA to some extent, and induce the activation of the replication checkpoint. However, unregulated continuous translocation of the CMG helicase exposes longer stretch of ssDNA (Figure 2). The uncoupling of the helicase from the polymerase may cause re-annealing of the two complementary single-strands behind the helicase, which is a serious obstacle at the time of resumption of DNA synthesis. In addition, the long ssDNA is likely to be fragile compared with the short one.

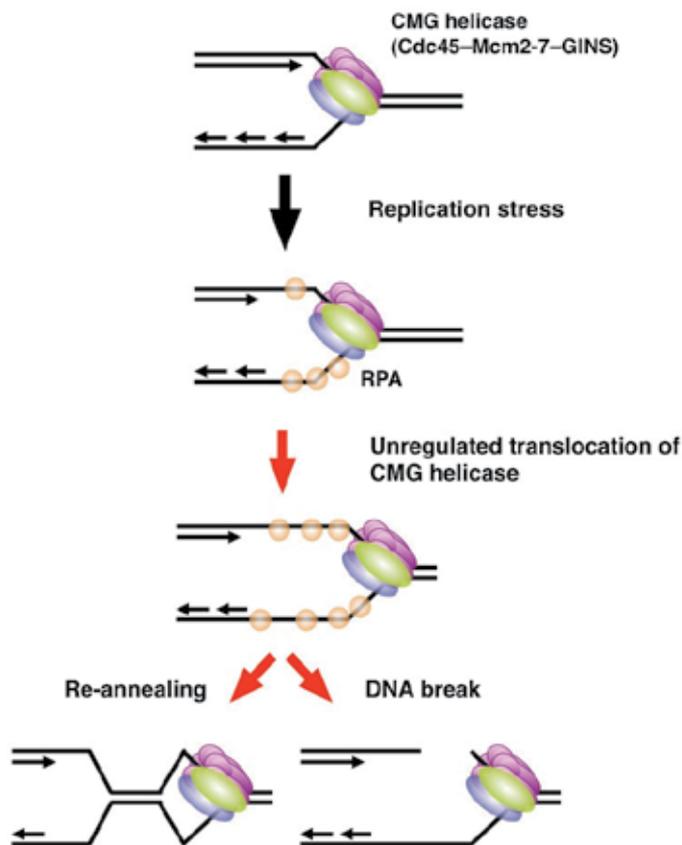


Fig. 2. Unregulated translocation of CMG helicase disrupts the replication fork.

No direct functional interaction between DNA polymerases and the CMG complexes has been reported, so far. However, in budding yeast, elimination of Tof1, Csm3 or Mrc1 causes uncoupling of the CMG helicase from DNA synthesis (Katou et al, 2003). Mrc1 may bridge

the gap between the leading strand polymerase and the CMG helicase, as Mrc1 interacts with both DNA pol  $\epsilon$  (Lou et al, 2008) and the MCM protein (Bando et al, 2009; Komata et al, 2009). On the other hand, Ctf4 may bridge the gap between the lagging strand polymerase and the CMG helicase as Ctf4 interacts with DNA Pol  $\alpha$  (Miles & Formosa, 1992) and with MCM and GINS (Gambus et al, 2009). The connection between the polymerase and helicase mediated by Mrc1 and Ctf4 may be important to couple the two important reactions of replication (Figure 3A). When DNA synthesis is inhibited, higher-order conformation around DNA polymerases would change. The conformational change emanates a signal that may be transmitted via the bridge molecules to the CMG helicase, to prevent uncoupling of the unwinding from the synthesis. Consistent to this model, mutations in the CMG complex suppress the hypersensitivity of fission yeast *mrc1* cells to HU (Nitani et al, 2006).

Another model is also proposed, in which the checkpoint kinase controls the activity of the CMG helicase (Figure 3B). The MCM proteins are phosphorylated dependent on the replication checkpoint kinases (Bailis et al, 2008; Cortez et al, 2004b; Ishimi et al, 2003; Randell et al, 2010; Yoo et al, 2004b). As phosphorylation of MCM down regulates the DNA unwinding activity of Mcm4-6-7 (Ishimi et al, 2003), it is possible that once the replication checkpoint kinase becomes activated with the aid of the CMG helicase, it turns off the activity of CMG.

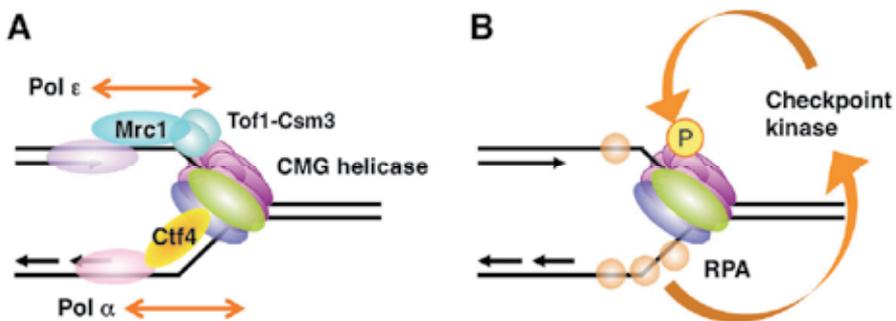


Fig. 3. Down regulation of the CMG helicase in response to replication stress.

The evolutionally conserved C-terminal domain (CTD) of Mcm4 plays an important role in the regulation of DNA unwinding carried out by the MCM helicase, as the deletion of the C-terminal in fission yeast accumulates RPA on the stalled replication forks while binding of MCM and GINS remains unaffected (Nitani et al, 2008). A recent study has shown that among Mcm2-7 subunits Mcm4 has the strongest affinity to both GINS and Cdc45 (Ilves et al, 2010). Thus, it is possible that the C-terminal region of Mmc4 affects the configuration of the CMG complex thereby regulating its helicase activity when DNA synthesis is inhibited. It is important to understand the interaction between the components of CMG, as is modulated in response to replication stress.

There is increasing body of evidence that regulation of the CMG helicase is important not only for DNA replication but also for maintaining the genome integrity. Recent findings that hypomorphic mutations in the replication proteins such as MCM in mammals cause growth abnormality and or predisposition to cancer, underscore the importance of the tight regulation of DNA replication. Genome-wide analysis of replication origins in different

types of tissues or cells would reveal the specificity of cancer to certain types of tissues. Detailed studies of spatiotemporal regulation of the integral component of the replication fork, the CMG helicase, would provide great insights into the mechanism by which chromosomal DNA is faithfully replicated in the S phase, which is one of the essential events during cell proliferation.

## 7. Acknowledgement

We thank Takahiro Inoue for his comments on the manuscript. This work was supported by the project for promotion of strategic creative research from MEXT, Japan to TN.

## 8. References

- Alcasabas AA, Osborn AJ, Bachant J, Hu F, Werler PJ, Bousset K, Furuya K, Diffley JF, Carr AM, Elledge SJ (2001) Mrc1 transduces signals of DNA replication stress to activate Rad53. *Nat Cell Biol* 3: 958-965
- Aparicio OM, Weinstein DM, Bell SP (1997) Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. *Cell* 91: 59-69
- Araki H, Leem SH, Phongdara A, Sugino A (1995) Dpb11, which interacts with DNA polymerase II(epsilon) in *Saccharomyces cerevisiae*, has a dual role in S-phase progression and at a cell cycle checkpoint. *Proc Natl Acad Sci U S A* 92: 11791-11795
- Aves SJ, Tongue N, Foster AJ, Hart EA (1998) The essential *Schizosaccharomyces pombe* *cdc23* DNA replication gene shares structural and functional homology with the *Saccharomyces cerevisiae* DNA43 (MCM10) gene. *Curr Genet* 34: 164-171
- Azvolinsky A, Giresi PG, Lieb JD, Zakian VA (2009) Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae*. *Mol Cell* 34: 722-734
- Bailis JM, Luche DD, Hunter T, Forsburg SL (2008) Minichromosome maintenance proteins interact with checkpoint and recombination proteins to promote S-phase genome stability. *Mol Cell Biol* 28: 1724-1738
- Bando M, Katou Y, Komata M, Tanaka H, Itoh T, Sutani T, Shirahige K (2009) Csm3, Tof1, and Mrc1 form a heterotrimeric mediator complex that associates with DNA replication forks. *J Biol Chem* 284: 34355-34365
- Baxter J, Diffley JF (2008) Topoisomerase II inactivation prevents the completion of DNA replication in budding yeast. *Mol Cell* 30: 790-802
- Bell SP, Stillman B (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* 357: 128-134
- Bermejo R, Doksani Y, Capra T, Katou YM, Tanaka H, Shirahige K, Foiani M (2007) Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes & Dev* 21: 1921-1936

- Bermudez VP, Lindsey-Boltz LA, Cesare AJ, Maniwa Y, Griffith JD, Hurwitz J, Sancar A (2003) Loading of the human 9-1-1 checkpoint complex onto DNA by the checkpoint clamp loader hRad17-replication factor C complex in vitro. *Proc Natl Acad Sci USA* 100: 1633-1638
- Bicknell LS, Bongers EM, Leitch A, Brown S, Schoots J, Harley ME, Aftimos S, Al-Aama JY, Bober M, Brown PA, van Bokhoven H, Dean J, Edrees AY, Feingold M, Fryer A, Hoefsloot LH, Kau N, Knoers NV, Mackenzie J, Opitz JM, Sarda P, Ross A, Temple IK, Toutain A, Wise CA, Wright M, Jackson AP (2011a) Mutations in the pre-replication complex cause Meier-Gorlin syndrome. *Nat Genet* 43:350-355
- Bicknell LS, Walker S, Klingseisen A, Stiff T, Leitch A, Kerzendorfer C, Martin CA, Yeyati P, Al Sanna N, Bober M, Johnson D, Wise C, Jackson AP, O'Driscoll M, Jeggo PA (2011b) Mutations in ORC1, encoding the largest subunit of the origin recognition complex, cause microcephalic primordial dwarfism resembling Meier-Gorlin syndrome. *Nat Genet* 43:356-359
- Blow JJ, Laskey RA (1986) Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of *Xenopus* eggs. *Cell* 47: 577-587
- Boddy MN, Gaillard PH, McDonald WH, Shanahan P, Yates JR, 3rd, Russell P (2001) Mus81-Eme1 are essential components of a Holliday junction resolvase. *Cell* 107: 537-548
- Boddy MN, Shanahan P, McDonald WH, Lopez-Girona A, Noguchi E, Yates IJ, Russell P (2003) Replication checkpoint kinase Cds1 regulates recombinational repair protein Rad60. *Mol Cell Biol* 23: 5939-5946
- Brewer BJ, Fangman WL (1988) A replication fork barrier at the 3' end of yeast ribosomal RNA genes. *Cell* 55: 637-643
- Brewer BJ, Lockshon D, Fangman WL (1992) The arrest of replication forks in the rDNA of yeast occurs independently of transcription. *Cell* 71: 267-276
- Brill SJ, DiNardo S, Voelkel-Meiman K, Sternglanz R (1987) Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA. *Nature* 326: 414-416
- Bucknall RA, Moores H, Simms R, Hesp B (1973) Antiviral effects of aphidicolin, a new antibiotic produced by *Cephalosporium aphidicola*. *Antimicrob Agents Chemother* 4: 294-298
- Byun TS, Pacek M, Yee MC, Walter JC, Cimprich KA (2005) Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev* 19: 1040-1052
- Calzada A, Hodgson B, Kanemaki M, Bueno A, Labib K (2005) Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. *Genes & Dev* 19: 1905-1919
- Choi JH, Lindsey-Boltz LA, Kemp M, Mason AC, Wold MS, Sancar A (2010) Reconstitution of RPA-covered single-stranded DNA-activated ATR-Chk1 signaling. *Proc Natl Acad Sci U S A* 107: 13660-13665

- Chuang CH, Wallace MD, Abratte C, Southard T, Schimenti JC (2010) Incremental genetic perturbations to MCM2-7 expression and subcellular distribution reveal exquisite sensitivity of mice to DNA replication stress. *PLoS Genet* 6:e1001110
- Cortez D, Glick G, Elledge SJ (2004a) Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proc Natl Acad Sci USA* 101: 10078-10083
- Cortez D, Glick G, Elledge SJ (2004b) Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proc Natl Acad Sci U S A* 101: 10078-10083
- Dalgaard JZ, Klar AJ (2000) *swi1* and *swi3* perform imprinting, pausing, and termination of DNA replication in *S. pombe*. *Cell* 102: 745-751
- Delacroix S, Wagner JM, Kobayashi M, Yamamoto K, Karnitz LM (2007) The Rad9-Hus1-Rad1 (9-1-1) clamp activates checkpoint signaling via TopBP1. *Genes & Dev* 21: 1472-1477
- Desany BA, Alcasabas AA, Bachant JB, Elledge SJ (1998) Recovery from DNA replicational stress is the essential function of the S-phase checkpoint pathway. *Genes & Dev* 12: 2956-2970
- Deshpande AM, Newlon CS (1996) DNA replication fork pause sites dependent on transcription. *Science* 272: 1030-1033
- Diffley JF, Cocker JH, Dowell SJ, Rowley A (1994) Two steps in the assembly of complexes at yeast replication origins in vivo. *Cell* 78: 303-316
- Doksani Y, Bermejo R, Fiorani S, Haber JE, Foiani M (2009) Replicon dynamics, dormant origin firing, and terminal fork integrity after double-strand break formation. *Cell* 137: 247-258
- Donnianni RA, Ferrari M, Lazzaro F, Clerici M, Tamilselvan Nachimuthu B, Plevani P, Muzi-Falconi M, Pelliccioli A (2010) Elevated levels of the polo kinase Cdc5 override the Mec1/ATR checkpoint in budding yeast by acting at different steps of the signaling pathway. *PLoS Genet* 6: e1000763
- Donovan S, Harwood J, Drury LS, Diffley JF (1997) Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *Proc Natl Acad Sci U S A* 94: 5611-5616
- Downs JA, Lowndes NF, Jackson SP (2000) A role for *Saccharomyces cerevisiae* histone H2A in DNA repair. *Nature* 408: 1001-1004
- Elia AE, Rellos P, Haire LF, Chao JW, Ivins FJ, Hoepker K, Mohammad D, Cantley LC, Smerdon SJ, Yaffe MB (2003) The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain. *Cell* 115: 83-95
- Ellison V, Stillman B (2003) Biochemical characterization of DNA damage checkpoint complexes: clamp loader and clamp complexes with specificity for 5' recessed DNA. *PLoS Biol* 1: E33
- Evrin C, Clarke P, Zech J, Lurz R, Sun J, Uhle S, Li H, Stillman B, Speck C (2009) A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc Natl Acad Sci USA* 106: 20240-20245

- Fukuura M, Nagao K, Obuse S, Takahashi TS, Nakagawa T, Masukata H (2011) CDK promotes interactions of Sld3 and Drc1 with Cut5 for initiation of DNA replication in fission yeast. *Mol Biol Cell* 22: 2620-2633
- Furuya K, Poitelea M, Guo L, Caspari T, Carr AM (2004) Chk1 activation requires Rad9 S/TQ-site phosphorylation to promote association with C-terminal BRCT domains of Rad4TOPBP1. *Genes & Dev* 18: 1154-1164
- Gambus A, Jones RC, Sanchez-Diaz A, Kanemaki M, van Deursen F, Edmondson RD, Labib K (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol* 8: 358-366
- Gambus A, Khoudoli GA, Jones RC, Blow JJ (2011) MCM2-7 form double hexamers at licensed origins in xenopus egg extract. *J Biol Chem* 286: 11855-11864.
- Gambus A, van Deursen F, Polychronopoulos D, Foltman M, Jones RC, Edmondson RD, Calzada A, Labib K (2009) A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. *EMBO J* 28: 2992-3004
- Glover JN, Williams RS, Lee MS (2004) Interactions between BRCT repeats and phosphoproteins: tangled up in two. *Trends Biochem Sci* 29: 579-585
- Goldman MA, Holmquist GP, Gray MC, Caston LA, Nag A (1984) Replication timing of genes and middle repetitive sequences. *Science* 224: 686-692
- Goscin LP, Byrnes JJ (1982) DNA polymerase delta: one polypeptide, two activities. *Biochemistry* 21: 2513-2518
- Groth P, Ausländer S, Majumder MM, Schultz N, Johansson F, Petermann E, Helleday T (2010) Methylated DNA causes a physical block to replication forks independently of damage signalling, O<sup>6</sup>-methylguanine or DNA single-strand breaks and results in DNA damage. *J Mol Biol* 402: 70-82
- Guernsey DL, Matsuoka M, Jiang H, Evans S, Macgillivray C, Nightingale M, Perry S, Ferguson M, Leblanc M, Paquette J, Patry L, Rideout AL, Thomas A, Orr A, McMaster CR, Michaud JL, Deal C, Langlois S, Superneau DW, Parkash S, Ludman M, Skidmore DL, Samuels ME (2011) Mutations in origin recognition complex gene ORC4 cause Meier-Gorlin syndrome. *Nat Genet* 43: 360-364
- Hayashi M, Katou Y, Itoh T, Tazumi A, Yamada Y, Takahashi T, Nakagawa T, Shirahige K, Masukata H (2007) Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast. *EMBO J* 26: 1327-1339
- Heller RC, Kang S, Lam WM, Chen S, Chan CS, Bell SP (2011) Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. *Cell* 146: 80-91
- Holm C, Goto T, Wang JC, Botstein D (1985) DNA topoisomerase II is required at the time of mitosis in yeast. *Cell* 41: 553-563
- Hsiang YH, Hertzberg R, Hecht S, Liu LF (1985) Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I. *J Biol Chem* 260: 14873-14878
- Huang M, Zhou Z, Elledge SJ (1998) The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor. *Cell* 94: 595-605

- Ikegami S, Taguchi T, Ohashi M, Oguro M, Nagano H, Mano Y (1978) Aphidicolin prevents mitotic cell division by interfering with the activity of DNA polymerase-alpha. *Nature* 275: 458-460
- Ilves I, Petojevic T, Pesavento JJ, Botchan MR (2010) Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell* 37: 247-258
- Ishimi Y (1997) A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J Biol Chem* 272: 24508-24513
- Ishimi Y, Komamura-Kohno Y, Kwon HJ, Yamada K, Nakanishi M (2003) Identification of MCM4 as a target of the DNA replication block checkpoint system. *J Biol Chem* 278: 24644-24650
- Ivessa AS, Lenzmeier BA, Bessler JB, Goudsouzian LK, Schnakenberg SL, Zakian VA (2003) The *Saccharomyces cerevisiae* helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol cell* 12: 1525-1536
- Izumi M, Yanagi K, Mizuno T, Yokoi M, Kawasaki Y, Moon KY, Hurwitz J, Yatagai F, Hanaoka F (2000) The human homolog of *Saccharomyces cerevisiae* Mcm10 interacts with replication factors and dissociates from nuclease-resistant nuclear structures in G(2) phase. *Nucleic Acids Res* 28: 4769-4777
- Kaguni JM, Fuller RS, Kornberg A (1982) Enzymatic replication of *E. coli* chromosomal origin is bidirectional. *Nature* 296: 623-627
- Kai M, Boddy MN, Russell P, Wang TS (2005) Replication checkpoint kinase Cds1 regulates Mus81 to preserve genome integrity during replication stress. *Genes Dev* 19: 919-932
- Kamimura Y, Masumoto H, Sugino A, Araki H (1998) Sld2, which interacts with Dpb11 in *Saccharomyces cerevisiae*, is required for chromosomal DNA replication. *Mol Cell Biol* 18: 6102-6109
- Kamimura Y, Tak YS, Sugino A, Araki H (2001) Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in *Saccharomyces cerevisiae*. *The EMBO J* 20: 2097-2107
- Kanemaki M, Sanchez-Diaz A, Gambus A, Labib K (2003) Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. *Nature* 423: 720-724
- Katou Y, Kanoh Y, Bando M, Noguchi H, Tanaka H, Ashikari T, Sugimoto K, Shirahige K (2003) S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. *Nature* 424: 1078-1083
- Kawabata T, Luebben SW, Yamaguchi S, Ilves I, Matise I, Buske T, Botchan MR, Shima N (2011) Stalled Fork Rescue via Dormant Replication Origins in Unchallenged S Phase Promotes Proper Chromosome Segregation and Tumor Suppression. *Mol Cell* 41: 543-553
- Kim RA, Wang JC (1989) Function of DNA topoisomerases as replication swivels in *Saccharomyces cerevisiae*. *J Mol Biol* 208: 257-267
- Kim S, Dallmann HG, McHenry CS, Marians KJ (1996) Coupling of a replicative polymerase and helicase: a tau-DnaB interaction mediates rapid replication fork movement. *Cell* 84: 643-650

- Kjeldsen E, Svejstrup JQ, Gromova, II, Alsner J, Westergaard O (1992) Camptothecin inhibits both the cleavage and religation reactions of eukaryotic DNA topoisomerase I. *J Mol Biol* 228: 1025-1030
- Kobayashi T, Hidaka M, Nishizawa M, Horiuchi T (1992) Identification of a site required for DNA replication fork blocking activity in the rRNA gene cluster in *Saccharomyces cerevisiae*. *Mol Gen Genet* 233: 355-362
- Kobayashi T, Horiuchi T (1996) A yeast gene product, Fob1 protein, required for both replication fork blocking and recombinational hotspot activities. *Genes Cells* 1: 465-474
- Komata M, Bando M, Araki H, Shirahige K (2009) The direct binding of Mrc1, a checkpoint mediator, to Mcm6, a replication helicase, is essential for the replication checkpoint against methyl methanesulfonate-induced stress. *Mol Cell Biol* 29: 5008-5019
- Koster DA, Palle K, Bot ES, Bjornsti MA, Dekker NH (2007) Antitumour drugs impede DNA uncoiling by topoisomerase I. *Nature* 448: 213-217
- Kubota Y, Mimura S, Nishimoto S, Takisawa H, Nojima H (1995) Identification of the yeast MCM3-related protein as a component of *Xenopus* DNA replication licensing factor. *Cell* 81: 601-609
- Kubota Y, Takase Y, Komori Y, Hashimoto Y, Arata T, Kamimura Y, Araki H, Takisawa H (2003) A novel ring-like complex of *Xenopus* proteins essential for the initiation of DNA replication. *Genes Dev* 17: 1141-1152
- Kumagai A, Lee J, Yoo HY, Dunphy WG (2006) TopBP1 activates the ATR-ATRIP complex. *Cell* 124: 943-955
- Kumagai A, Shevchenko A, Dunphy WG (2010) Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. *Cell* 140: 349-359
- Labib K, Tercero JA, Diffley JF (2000) Uninterrupted MCM2-7 function required for DNA replication fork progression. *Science* 288: 1643-1647
- Lawley PD, Brookes P (1963) Further Studies on the Alkylation of Nucleic Acids and Their Constituent Nucleotides. *Biochem J* 89: 127-138
- Lee J, Kumagai A, Dunphy WG (2007) The Rad9-Hus1-Rad1 checkpoint clamp regulates interaction of TopBP1 with ATR. *J Biol Chem* 282: 28036-28044
- Lee JH, Paull TT (2004) Direct activation of the ATM protein kinase by the Mre11/Rad50/Nbs1 complex. *Science* 304: 93-96
- Lee JH, Paull TT (2005) ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science* 308: 551-554
- Lei M, Kawasaki Y, Young MR, Kihara M, Sugino A, Tye BK (1997) Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis. *Genes Dev* 11: 3365-3374
- Letessier A, Millot GA, Koundrioukoff S, Lachages AM, Vogt N, Hansen RS, Malfoy B, Brison O, Debatisse M (2011) Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature* 470: 120-123
- Liang C, Weinreich M, Stillman B (1995) ORC and Cdc6p interact and determine the frequency of initiation of DNA replication in the genome. *Cell* 81: 667-676

- Lopes M, Cotta-Ramusino C, Pelliccioli A, Liberi G, Plevani P, Muzi-Falconi M, Newlon CS, Foiani M (2001) The DNA replication checkpoint response stabilizes stalled replication forks. *Nature* 412: 557-561
- Lopez-Mosqueda J, Maas NL, Jonsson ZO, Defazio-Eli LG, Wohlschlegel J, Toczyski DP (2010) Damage-induced phosphorylation of Sld3 is important to block late origin firing. *Nature* 467: 479-483
- Lou H, Komata M, Katou Y, Guan Z, Reis CC, Budd M, Shirahige K, Campbell JL (2008) Mrc1 and DNA polymerase epsilon function together in linking DNA replication and the S phase checkpoint. *Mol Cell* 32: 106-117
- Lowery DM, Clauser KR, Hjerrild M, Lim D, Alexander J, Kishi K, Ong SE, Gammeltoft S, Carr SA, Yaffe MB (2007) Proteomic screen defines the Polo-box domain interactome and identifies Rock2 as a Plk1 substrate. *EMBO J* 26: 2262-2273
- MacDougall CA, Byun TS, Van C, Yee MC, Cimprich KA (2007) The structural determinants of checkpoint activation. *Genes & Dev* 21: 898-903
- Maine GT, Sinha P, Tye BK (1984) Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics* 106: 365-385
- Masai H, Taniyama C, Ogino K, Matsui E, Kakusho N, Matsumoto S, Kim JM, Ishii A, Tanaka T, Kobayashi T, Tamai K, Ohtani K, Arai K (2006) Phosphorylation of MCM4 by Cdc7 kinase facilitates its interaction with Cdc45 on the chromatin. *J Biol Chem* 281: 39249-39261
- Masumoto H, Muramatsu S, Kamimura Y, Araki H (2002) S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature* 415: 651-655
- Masumoto H, Sugino A, Araki H (2000) Dpb11 controls the association between DNA polymerases alpha and epsilon and the autonomously replicating sequence region of budding yeast. *Mol Cell Biol* 20: 2809-2817
- Miles J, Formosa T (1992) Evidence that POB1, a *Saccharomyces cerevisiae* protein that binds to DNA polymerase alpha, acts in DNA metabolism in vivo. *Mol Cell Biol* 12: 5724-5735
- Miyabe I, Morishita T, Hishida T, Yonei S, Shinagawa H (2006) Rhp51-dependent recombination intermediates that do not generate checkpoint signal are accumulated in *Schizosaccharomyces pombe* rad60 and smc5/6 mutants after release from replication arrest. *Mol Cell Biol* 26: 343-353
- Miyabe I, Morishita T, Shinagawa H, Carr AM (2009) *Schizosaccharomyces pombe* Cds1Chk2 regulates homologous recombination at stalled replication forks through the phosphorylation of recombination protein Rad60. *J Cell Sci* 122: 3638-3643
- Morishita T, Tsutsui Y, Iwasaki H, Shinagawa H (2002) The *Schizosaccharomyces pombe* rad60 gene is essential for repairing double-strand DNA breaks spontaneously occurring during replication and induced by DNA-damaging agents. *Mol Cell Biol* 22: 3537-3548
- Moyer SE, Lewis PW, Botchan MR (2006) Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci U S A* 103: 10236-10241
- Murakami H, Okayama H (1995) A kinase from fission yeast responsible for blocking mitosis in S phase. *Nature* 374: 817-819

- Muramatsu S, Hirai K, Tak YS, Kamimura Y, Araki H (2010) CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon), and GINS in budding yeast. *Genes Dev* 24: 602-612
- Nakajima R, Masukata H (2002) SpSld3 is required for loading and maintenance of SpCdc45 on chromatin in DNA replication in fission yeast. *Mol Biol Cell* 13: 1462-1472
- Neuwald AF, Aravind L, Spouge JL, Koonin EV (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome research* 9: 27-43
- Niedernhofer LJ, Daniels JS, Rouzer CA, Greene RE, Marnett LJ (2003) Malondialdehyde, a product of lipid peroxidation, is mutagenic in human cells. *J Biol Chem* 278: 31426-31433
- Nishitani H, Lygerou Z, Nishimoto T, Nurse P (2000) The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature* 404: 625-628
- Nitani N, Nakamura K, Nakagawa C, Masukata H, Nakagawa T (2006) Regulation of DNA replication machinery by Mrc1 in fission yeast. *Genetics* 174: 155-165
- Nitani N, Yadani C, Yabuuchi H, Masukata H, Nakagawa T (2008) Mcm4 C-terminal domain of MCM helicase prevents excessive formation of single-stranded DNA at stalled replication forks. *Proc Natl Acad Sci U S A* 105: 12973-12978
- Noguchi E, Noguchi C, Du LL, Russell P (2003) Swi1 prevents replication fork collapse and controls checkpoint kinase Cds1. *Mol Cell Biol* 23: 7861-7874
- Noguchi E, Noguchi C, McDonald WH, Yates JR, 3rd, Russell P (2004) Swi1 and Swi3 are components of a replication fork protection complex in fission yeast. *Mol Cell Biol* 24: 8342-8355
- Noguchi E, Shanahan P, Noguchi C, Russell P (2002) CDK phosphorylation of Drc1 regulates DNA replication in fission yeast. *Curr Biol* 12: 599-605
- Pacek M, Tutter AV, Kubota Y, Takisawa H, Walter JC (2006) Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol Cell* 21: 581-587
- Pardee AB (1989) G1 events and regulation of cell proliferation. *Science* 246: 603-608
- Randell JC, Fan A, Chan C, Francis LI, Heller RC, Galani K, Bell SP (2010) Mec1 Is One of Multiple Kinases that Prime the Mcm2-7 Helicase for Phosphorylation by Cdc7. *Mol Cell* 40: 353-363
- Remus D, Beuron F, Tolun G, Griffith JD, Morris EP, Diffley JF (2009) Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* 139: 719-730
- Rozenzhak S, Mejia-Ramirez E, Williams JS, Schaffer L, Hammond JA, Head SR, Russell P (2010) Rad3 decorates critical chromosomal domains with gammaH2A to protect genome integrity during S-Phase in fission yeast. *PLoS Genet* 6: e1001032
- Saka Y, Yanagida M (1993) Fission yeast cut5+, required for S phase onset and M phase restraint, is identical to the radiation-damage repair gene rad4+. *Cell* 74: 383-393
- Sangrithi MN, Bernal JA, Madine M, Philpott A, Lee J, Dunphy WG, Venkitaraman AR (2005) Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. *Cell* 121: 887-898

- Sansam CL, Cruz NM, Danielian PS, Amsterdam A, Lau ML, Hopkins N, Lees JA (2010) A vertebrate gene, *ticrr*, is an essential checkpoint and replication regulator. *Genes & Dev* 24: 183-194
- Santocanale C, Diffley JF (1998) A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. *Nature* 395: 615-618
- Sekedat MD, Fenyo D, Rogers RS, Tackett AJ, Aitchison JD, Chait BT (2010) GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. *Mol Syst Biol* 6: 353
- Sheu YJ, Stillman B (2006) Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Mol Cell* 24: 101-113
- Sheu YJ, Stillman B (2010) The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature* 463: 113-117
- Shima N, Alcaraz A, Liachko I, Buske TR, Andrews CA, Munroe RJ, Hartford SA, Tye BK, Schimenti JC (2007) A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* 39: 93-98
- Shimmoto M, Matsumoto S, Odagiri Y, Noguchi E, Russell P, Masai H (2009) Interactions between Swi1-Swi3, Mrc1 and S phase kinase, Hsk1 may regulate cellular responses to stalled replication forks in fission yeast. *Genes Cells* 14: 669-682
- Shirahige K, Hori Y, Shiraishi K, Yamashita M, Takahashi K, Obuse C, Tsurimoto T, Yoshikawa H (1998) Regulation of DNA-replication origins during cell-cycle progression. *Nature* 395: 618-621
- Solomon NA, Wright MB, Chang S, Buckley AM, Dumas LB, Gaber RF (1992) Genetic and molecular analysis of DNA43 and DNA52: two new cell-cycle genes in *Saccharomyces cerevisiae*. *Yeast* 8: 273-289
- Stewart DJ (2004) Topotecan in the first-line treatment of small cell lung cancer. *Oncologist* 9 Suppl 6: 33-42
- Szilard RK, Jacques PE, Laramee L, Cheng B, Galicia S, Bataille AR, Yeung M, Mendez M, Bergeron M, Robert F, Durocher D (2010) Systematic identification of fragile sites via genome-wide location analysis of  $\gamma$ -H2AX. *Nat Struct Mol Biol* 17: 299-305
- Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, Araki H (2003) GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev* 17: 1153-1165
- Takeuchi Y, Horiuchi T, Kobayashi T (2003) Transcription-dependent recombination and the role of fork collision in yeast rDNA. *Genes & Dev* 17: 1497-1506
- Tanaka K, Russell P (2001) Mrc1 channels the DNA replication arrest signal to checkpoint kinase Cds1. *Nat Cell Biol* 3: 966-972
- Tanaka S, Diffley JF (2002) Deregulated G1-cyclin expression induces genomic instability by preventing efficient pre-RC formation. *Genes Dev* 16: 2639-2649
- Tanaka S, Umemori T, Hirai K, Muramatsu S, Kamimura Y, Araki H (2007) CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* 445: 328-332
- Tercero JA, Diffley JF (2001) Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint. *Nature* 412: 553-557
- Theis JF, Irene C, Dershowitz A, Brost RL, Tobin ML, di Sanzo FM, Wang JY, Boone C, Newlon CS (2010) The DNA damage response pathway contributes to the

- stability of chromosome III derivatives lacking efficient replicators. *PLoS Genet* 6: e1001227
- Toczyski DP, Galgoczy DJ, Hartwell LH (1997) CDC5 and CKII control adaptation to the yeast DNA damage checkpoint. *Cell* 90: 1097-1106
- Trenz K, Errico A, Costanzo V (2008) Plx1 is required for chromosomal DNA replication under stressful conditions. *EMBO J* 27: 876-885
- Tsao CC, Geisen C, Abraham RT (2004) Interaction between human MCM7 and Rad17 proteins is required for replication checkpoint signaling. *EMBO J* 23: 4660-4669
- Tsvetkov L, Stern DF (2005) Interaction of chromatin-associated Plk1 and Mcm7. *J Biol Chem* 280: 11943-11947
- Unsal-Kacmaz K, Chastain PD, Qu PP, Mino P, Cordeiro-Stone M, Sancar A, Kaufmann WK (2007) The human Tim/Tipin complex coordinates an Intra-S checkpoint response to UV that slows replication fork displacement. *Mol Cell Biol* 27: 3131-3142
- Usui T, Ogawa H, Petrini JH (2001) A DNA damage response pathway controlled by Tel1 and the Mre11 complex. *Mol Cell* 7: 1255-1266
- Vidaness GM, Sweeney FD, Galicia S, Cheung S, Doyle JP, Durocher D, Toczyski DP (2010) CDC5 inhibits the hyperphosphorylation of the checkpoint kinase Rad53, leading to checkpoint adaptation. *PLoS Biol* 8: e1000286
- Wang H, Elledge SJ (1999) DRC1, DNA replication and checkpoint protein 1, functions with DPB11 to control DNA replication and the S-phase checkpoint in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 96: 3824-3829
- Wyrick JJ, Aparicio JG, Chen T, Barnett JD, Jennings EG, Young RA, Bell SP, Aparicio OM (2001) Genome-wide distribution of ORC and MCM proteins in *Saccharomyces cerevisiae*: high-resolution mapping of replication origins. *Science* 294: 2357-2360
- Yabuuchi H, Yamada Y, Uchida T, Sunathvanichkul T, Nakagawa T, Masukata H (2006) Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins. *EMBO J* 25: 4663-4674
- Yamada Y, Nakagawa T, Masukata H (2004) A novel intermediate in initiation complex assembly for fission yeast DNA replication. *Mol Biol Cell* 15: 3740-3750
- Yamane K, Kawabata M, Tsuruo T (1997) A DNA-topoisomerase-II-binding protein with eight repeating regions similar to DNA-repair enzymes and to a cell-cycle regulator. *Eur J Biochem* 250: 794-799
- Yoo HY, Kumagai A, Shevchenko A, Dunphy WG (2004a) Adaptation of a DNA replication checkpoint response depends upon inactivation of Claspin by the Polo-like kinase. *Cell* 117: 575-588
- Yoo HY, Shevchenko A, Dunphy WG (2004b) Mcm2 is a direct substrate of ATM and ATR during DNA damage and DNA replication checkpoint responses. *J Biol Chem* 279: 53353-53364
- Young CW, Schochetman G, Karnofsky DA (1967) Hydroxyurea-induced inhibition of deoxyribonucleotide synthesis: studies in intact cells. *Cancer Res* 27: 526-534
- Zegerman P, Diffley JF (2007) Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* 445: 281-285
- Zegerman P, Diffley JF (2010) Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. *Nature* 467: 474-478

Zou L, Cortez D, Elledge SJ (2002) Regulation of ATR substrate selection by Rad17-dependent loading of Rad9 complexes onto chromatin. *Genes & Dev* 16: 198-208

Zou L, Elledge SJ (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* 300: 1542-1548

# Cell Cycle Control of DNA Replication by Phosphorylation and Dephosphorylation of Replication-Initiation Proteins in Budding Yeast

Yuanliang Zhai\*, Philip Y.K. Yung\* and Chun Liang

*Division of Life Science and Center for Cancer Research*

*Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong  
China*

## 1. Introduction

Eukaryotic DNA replication initiates in a two-step fashion (Diffley, 2004), and both steps are tightly regulated by phosphorylation and dephosphorylation of replication-initiation proteins. First, a multi-protein complex known as pre-RC (pre-replicative complex) is assembled on replication origins from late M to early G1 phase; this process is referred to as replication licensing. Subsequently, origin activation (firing) occurs in accordance to the rise of S-CDK (cyclin dependent kinase) and DDK (Dbf4p dependent kinase) activities during which replication forks are established and DNA synthesis begins.

### 1.1 Components of replication licensing

Pre-RC is assembled onto replication origins in a stepwise manner, and it includes the following components:

#### 1.1.1 Replicators and replication origins

Replicators and replication origins in budding yeast were together identified as some conserved DNA sequences, autonomously replicating sequences (ARSs). There are about 400 ARSs distributed throughout the yeast genome to promote the association of initiation proteins and the unwinding of DNA double helices (Raghuraman et al., 2001). Despite the conservation of the initiation proteins in eukaryotes, the origin sequences are highly divergent, and poorly characterized in many organisms (Machida et al., 2005; Sclafani & Holzen, 2007)

#### 1.1.2 ORC (Origin recognition complex)

Origin recognition complex (ORC) can specifically recognize and associate with ARSs directly in an ATP dependent manner throughout the cell cycle, serving as a landing pad for other factors to be assembled. ORC utilizes the energy from ATP hydrolysis to drive pre-RC assembly (Bowers et al., 2004; Randell et al., 2006). Although the six subunits of ORC (Orc1-6p) form a tight complex *in vivo*, only Orc1-5p are required for origin recognition and DNA binding (Stillman, 2005). Nevertheless, Orc6p plays an essential role in pre-RC assembly

---

\* These authors contributed equally to this chapter

through interacting with Cdt1p (Cdc10p-dependent transcript) to promote multiple rounds of MCM (Minichromosome maintenance) loading onto replication origins (Chen et al., 2007).

ORC is functionally related to DnaA, the initiator protein in bacteria. Structural analysis showed that DnaA displays a helical conformation that, when bound to the replication origin, induces superhelical tension which causes localized DNA melting to facilitate helicase loading (Erzberger et al., 2006). Despite detailed biochemical analysis, the exact mechanism of ORC as the initiator awaits further investigations. However, some studies highlighted its role in manipulating local chromatin environment (Lipford & Bell, 2001; Kan et al., 2008; Espinosa et al., 2010).

### 1.1.3 Cdc6p (Cell division cycle)

The budding yeast Cdc6p was identified as a multicopy suppressor of the *orc5-1* ts mutant (Liang et al., 1995). It interacts with ORC, and is required for loading the hexameric helicase MCM (Minichromosome maintenance) proteins onto chromatin (Cocker et al., 1996; Donovan et al., 1997; Tanaka et al., 1997). Analysis with EM micrograph reconstruction showed that the ORC-Cdc6p complex displayed a ring-shaped structure, which possesses six AAA+ ATPase domains (five from ORC and one from Cdc6p) (Speck et al., 2005). *In vitro* studies suggested that the ATPase activities from both initiation proteins are needed for robust MCM loading (Seki & Diffley, 2000). Importantly, ATP hydrolysis by Cdc6p destabilizes the ORC-Cdc6p complex, and because ARS-specific DNA binding to the complex inhibits the Cdc6p ATPase activity (Speck & Stillman, 2007), it is thus proposed that the ATPase domain of Cdc6p confers specificity on the site of pre-RC assembly.

### 1.1.4 Cdt1p (Cdc10p-dependent transcript)

In contrast to other initiation proteins which were first identified in budding yeast, the helicase loader component Cdt1p was first isolated from fission yeast (Nishitani et al., 2000), *Xenopus* (Maiorano et al., 2000) and *Drosophila* (Whittaker et al., 2000) as a protein involved in DNA replication. Based on homology search, and despite low level of sequence homology, the budding yeast Cdt1p was identified as the gene *TAH11* (Topo-A hypersensitive) (Tanaka & Diffley, 2002), previously isolated by virtue of its genetic interaction with a topoisomerase I mutant. Cdt1p in budding yeast, like the homologs in higher eukaryotes, is essential for loading MCM onto replication origins, acting as a bridge between Mcm2-7p and ORC. On the one hand, Cdt1p interacts with the MCM complex specifically through the Mcm6p subunit (Wei et al., 2010); on the other hand, Cdt1p associates with ORC via transient contact with Orc6p, resulting in the incorporation of Mcm2-7p complex into pre-RC (Chen et al., 2007). Unlike higher eukaryotes and fission yeast where Cdt1p is regulated by degradation, ScCdt1p is stable throughout the cell cycle but is regulated by nuclear import and export (Tanaka & Diffley, 2002).

### 1.1.5 MCM (Minichromosome maintenance) proteins

MCM genes were first isolated in a genetic screen from mutants defective in maintaining the stability of ARS-based plasmids (Maine et al., 1984). The six subunits are highly related to one another in sequence, but each MCM protein is essential for cell growth. MCM proteins migrate with the replication fork (Aparicio et al., 1997) and function as the putative replicative helicase. *In vitro* studies strongly suggest that Mcm2-7 proteins act as DNA helicase (Bochman & Schwacha, 2009). The Mcm 4/6/7 complex purified from HeLa cells

displayed weak DNA helicase activity. Similar studies in mouse and fission yeast also identified a DNA helicase activity dependent on the Mcm4/6/7 complex. Surprisingly, although all six Mcm proteins function at the replication fork in budding yeast, the intact Mcm2-7p complex showed no helicase activity *in vitro* (Takahashi et al., 2005). A recent study with biochemical reconstitution of *Drosophila* CMG (Cdc45p, MCM and GINS) complex suggested that the MCM complex is loaded in an inactive form as a pre-RC component (Ilves et al., 2010), and it displays robust ATPase and helicase activity only when it is associated with Cdc45p and GINS. Apart from the association with Cdc45p and GINS, modifications on MCM also play an important role in the helicase activation.

In order to perform its helicase activity, the Mcm2-7p complex needs the action of both DDK and CDKs to phosphorylate some MCM subunits as well as other initiation proteins (Labib, 2010). The N-terminal tails of Mcm2/Mcm4/Mcm6 appears to be the major substrates of DDK for DNA replication, among which Mcm4p phosphorylation by DDK is particularly important. N-terminal serine/threonine-rich domain (NSD) of Mcm4p contains multiple motifs that are targets for DDK, and phosphorylations of these motifs by DDK relieve the inhibitory effect of Mcm4-NSD on DNA replication (Sheu & Stillman, 2006; 2010).

At the heart of the regulation of replication initiation is the loading of the MCM helicase onto origins and its subsequent activation to unwind the double helix. The loading process is well organized spatially and temporally. Oscillation of CDK activity during cell cycle restricts this process only within late anaphase and G1 phase at low CDK activity (Diffley, 2004). At the M-to-G1 transition, the MCM complex is allowed to be imported into the nucleus. Only after the helicase loaders Cdc6p and Cdt1p associate with ORC can the MCM complex be loaded onto chromatin to form pre-RC (Cocker et al., 1996; Donovan et al., 1997; Tanaka & Diffley, 2002). Recent studies using reconstituted yeast pre-RC and EM microscopy demonstrated that MCM is loaded onto replication origins as a double-hexamer, connected in a head-to-head configuration via their N-terminal tails (Evrin et al., 2009; Remus et al., 2009).

### 1.1.6 Noc3p (Nucleolar-associated complex)

Noc3p was isolated as a multicopy suppressor of the *mcm5-1* ts mutant (Zhang et al., 2002). Noc3p mutants display classical ARS-number suppressible plasmid loss phenotype. Noc3p binds to ARS DNA and interacts with ORC and MCM proteins. And inactivation of Noc3p results in failure of pre-RC assembly and G1/S transition (Zhang et al., 2002). Because both ORC and Noc3p bind to the chromatin throughout the cell cycle and depletion of Noc3p only affect the binding of initiation proteins downstream of ORC onto the chromatin, ORC and Noc3p represent ARS-bound scaffold proteins that mark the early step of pre-RC assembly. Besides, since both ORC and Noc3p have been implicated in ribosome biogenesis (Milkereit et al., 2001; Du and Stillman, 2002), they might be important regulator coordinating both cell growth and division (Jorgensen & Tyers, 2004). A study in budding yeast also shows that overexpression of Noc3p caused severe growth defects in *orc* and *cdc6* mutants, just as overexpression of some MCM subunits and some other replication-initiation proteins did (Honey and Futcher, 2007), supporting the role of Noc3p in replication initiation together with ORC and MCM proteins. Noc3p interacts with Cdt1p, Mcm6p and four of the six subunits of ORC that are closely connected with one another (Wu, R., J. Yeung and CL, unpublished data), apparently making a snug fit into the main body of the ORC architecture. We have also isolated separation-of-function mutants in *NOC3* (Wu, R. and CL, unpublished data), indicating that the Noc3p's functions in DNA replication and ribosome biogenesis are separable.

Consistent with the role of Noc3p as an initiation protein, *fad24* (Factor for adipocyte differentiation, the human homolog of Noc3p), is found to regulate differentiation (Tominaga et al., 2004) and is essential for DNA replication in human (Johmura et al., 2008a) and mouse cells (Johmura et al., 2008b). Importantly, *fad24* interacts with HBO1 (Histone acetyltransferase-binding to ORC), which is an Orc1p interacting protein in human cells required for acetylating origin-proximal histone H4 to promote replication licensing (Iizuka & Stillman, 1999). If these interactions are conserved in yeast, it is tempting to speculate that Noc3p might also recruit histone acetyltransferase (HAT) to ARSs. Gcn5p, a recently identified HAT with important role in replication licensing (Espinosa et al., 2010), would be an attractive candidate.

Data from a fission yeast study are also consistent with Noc3p being required for replication initiation, as Noc3p inactivation (probably incompletely) caused delay in S phase entry, slow progression through S phase, and cell cycle arrest in late S and/or G2/M phases, although the authors interpreted the data in favor of Noc3p's role in cell division instead of replication (Houchens et al., 2008). These phenotypes are typical of the majority of the published ORC and MCM ts mutants in both budding and fission yeasts, as partial initiation of DNA replication by the residual mutant proteins even under the restrictive conditions for cell growth allows the cells to replicate most, but not all of the DNA, resulting in late S and/or G2/M arrest. Consistent with this interpretation, these mutants would not arrest in the first G2/M phase if the cells were released into the cell cycle after the mutant proteins were inactivated in early S phase or G2/M phase instead of G1 phase (Zhang et al., 2002; Gibson et al., 2006). Therefore, the apparent cell cycle arrest in G2/M phases is the result of incomplete DNA replication, not because of a cell division defect *per se*.

## 1.2 A model of replication licensing

Summing up the current data, a model of pre-RC assembly is proposed as shown in Fig. 1. ORC and Noc3p bind to ARS elements throughout the cell cycle, and they act as a scaffold to recruit other initiators. During mitotic exit, *de novo* synthesized and stabilized Cdc6p binds with ORC/Noc3p and together they form a ring complex on ARS. At the same time MCM and Cdt1p are imported into the nucleus in an inter-dependent manner. And through the concerted action of Cdc6p and Cdt1p, the MCM helicase is loaded as a double hexamer onto the chromatin to license the replication origin. This complex composed of ORC, Noc3p, Cdc6p, Cdt1p, MCM and possibly other unidentified proteins, is referred to as the pre-RC, which is assembled only at low CDK activity upon mitotic exit to early G1 of the next cell cycle. Worth mentioning, all of the known pre-RC components are conserved from yeasts to humans despite that some of them are regulated differently in different organisms.

## 1.3 Origin activation (Origin firing)

Origin firing refers to the activation of licensed replication origins by the concerted activities of CDK and DDK during which helicase activation is coupled to the recruitment of DNA polymerases.

### 1.3.1 GINS (Go, Ichi, Nii, and San; five, one, two, and three in Japanese)

The GINS complex is composed of Sld5p, Psf1p, -2p and -3p discovered by several groups (Kanemaki et al., 2003; Kubota et al., 2003; Takayama et al., 2003). A functional proteomic

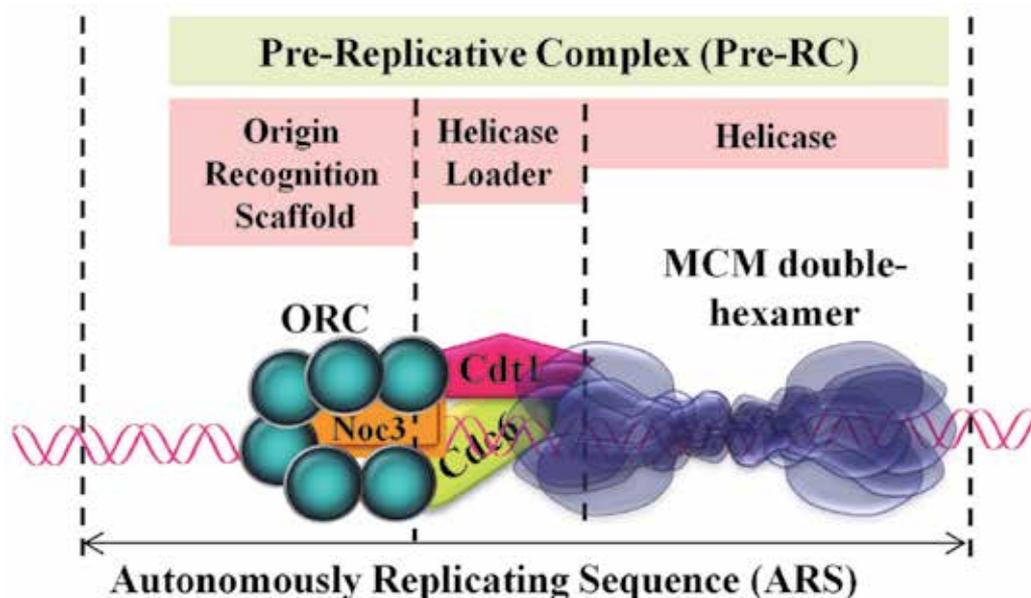


Fig. 1. The pre-replication complex in budding yeast.

The pre-RC is assembled step-wise on ARS during late M to early G1 phase of the cell cycle. It is composed of several groups of proteins including the origin recognition scaffold (ORC and Noc3p), the helicase loaders (Cdc6p and Cdt1p) and the helicase (the MCM complex). See text for details.

screen in budding yeast identified the GINS components that, upon depletion, would affect DNA replication (Kanemaki et al., 2003). *PSF1* and *PSF2* (Partner of Sld Five) were identified as multicopy suppressors of the *sdl5-12* ts mutant, and *PSF3* was identified as a multicopy suppressor of the *psf1-1* ts mutant (Takayama et al., 2003). These proteins interact with one another and are essential for DNA replication in budding yeast. By sequence homology search, they were identified and purified as a ring-like complex from *Xenopus* egg extracts (Kubota et al., 2003). Crystallography analysis reveals that the human GINS complex adopts a trapezium-like structure (Boskovic et al., 2007), rather than the ring-like shape counterpart in *Xenopus* as observed on EM micrograph (Kubota et al., 2003). GINS is only recruited to origins upon S-phase onset, and it travels along with the replication fork as replication proceeds (Kanemaki et al., 2003).

### 1.3.2 Cdc45p

*CDC45* was initially characterized as one of a group of genes, together with *MCM4*, *5*, and *7* that display extensive genetic interactions with one another (Hennessy et al., 1991). Cdc45p forms a complex with Sld3p, and together they are recruited in an inter-dependent manner to licensed origins throughout S-phase (Kamimura et al., 2001). Notably, early-firing origins are marked with Sld3p-Cdc45p even in G1 phase (Kamimura et al., 2001). And because Cdc45p origin binding profile coincides with the activation timing of origins, Cdc45p is considered as a marker for active origins (Aparicio et al., 1997; Zou & Stillman, 1998; Kamimura et al., 2001). In addition, stable recruitment of Cdc45p to pre-RC requires DDK phosphorylation of the Mcm4p N-terminal tail (Sheu & Stillman, 2006).

### 1.3.3 Sld2p, Sld3p and Dpb11p

*DPB11* (DNA polymerase B) was first identified as a multicopy suppressor of the *ts* alleles *dpb2-1* and *pol12-1* (Araki et al., 1995), which encode two different subunits of the DNA polymerase  $\epsilon$  in budding yeast. Identified by the same genetic screen strategy, *SLD2*, *SLD3* and other *SLD* genes (including *SLD5* in the GINS complex) were found to suppress the *dpb11-1 ts* allele when they are carried on high copy plasmids (Kamimura et al., 1998).

Dpb11p contains two pairs of tandem BRCT (BRCA1 C-terminal) domains, which are sites known as phosphopeptide binding modules. Dpb11p is proposed to be a bridging factor to promote DNA replication. Both Sld2p and -3p are substrates of S-CDK, showing an S-phase specific phosphorylation profile, and they only bind Dpb11p in the phosphorylated forms (Tanaka et al., 2007; Zegerman & Diffley, 2007). Such interactions are elegantly partitioned over Dpb11p. Notably, the N-terminal proximal BRCT repeats of Dpb11p are required for Sld3p binding while those located in the C-terminal region interact with Sld2p. Mutagenesis study and synthetic phospho-peptide mapping revealed that phosphorylation at T84 on Sld2p, and that T600 and S622 in Sld3p are responsible for mediating Dpb11p binding (Zegerman & Diffley, 2007). By a series of rescue experiments and genetic screens, it was found that the lethality of cells carrying non-phosphorylatable Sld3p (i.e. Sld3p-600A, 609A, 622A) could be rescued when the mutant Sld3p was fused with the N-terminal BRCT repeats-truncated Dpb11p (referred to as SD-fusion) (Zegerman & Diffley, 2007); while cells with non-phosphorylatable Sld2p could be rescued by a Sld2p phospho-mimic mutant (either Sld2-T84D or Sld2-11D) (Tanaka et al., 2007; Zegerman & Diffley, 2007). Moreover, ectopic expression of Sld2-T84D in an SD-fusion genetic background led to DNA replication independent of S-CDK. Remarkably, when these CDK-bypass mutants were combined with the *bob1* allele (an *MCM5* allele that suppresses *cdc7 $\Delta$*  lethality and bypasses the requirement of DDK in replication) (Tanaka et al., 2007; Zegerman & Diffley, 2007) or the *mcm4 $\Delta$ 74-174* allele (a DDK-bypass allele of *MCM4* lacking a self-inhibitory domain) (Sheu & Stillman, 2010), DNA replication still occurred even when the yeast cells were arrested in G1 phase lacking both CDK and DDK activities. Taking together, Sld2p, Sld3p and Mcm4p represent the minimal targets of the S-phase kinases in promoting DNA replication.

### 1.3.4 Mcm10p, Ctf4p and DNA polymerase $\alpha$

Budding yeast Mcm10p is a recruitment factor that targets polymerase  $\alpha$  to the licensed origins in G1 phase and the replication forks during S-phase (Ricke & Bielsky, 2004; Lee et al., 2010). It binds to origins in G1 cells in an MCM-dependent manner and co-migrates with the fork as part of the replisome progression complex (RPC) (Gambus et al., 2006). Study in *Xenopus* suggested that Mcm10p channels its function in part through interacting with a protein known as And1 (Ctf4p in budding yeast), which is important for efficient DNA synthesis and acts as a bridging factor between Mcm10p and Pol  $\alpha$  (Zhu et al., 2007). The budding yeast Ctf4p was initially isolated as a mutant defective in chromosome segregation (Spencer et al., 1990). It travels with the replication fork and is known to participate in sister chromatids cohesion (Hanna et al., 2001; Gambus et al., 2006). Consistent with the *Xenopus* ortholog, Ctf4p was also identified as a RPC component, and it interacts with Mcm10p for stable chromatin association during S-phase (Tanaka et al., 2009). Besides, it helps to stabilize both Mcm10p and Pol  $\alpha$  (Gambus et al., 2009; Wang et al., 2010). And cells without Ctf4p display a delay onset and slow progression of S-phase (Wang et al., 2010).

### 1.4 Current model of origin activation

Both CDK and DDK activities are required for origin firing, and genetic evidence suggested that CDKs work in advanced of DDK to promote DNA replication. As illustrated in Fig. 2, binding of Sld3p and Cdc45p to pre-RC in G1 cells marks early-firing origins (Kamimura et al., 2001). Upon S-phase onset, rising S-CDK activity phosphorylates both Sld2p and Sld3p, creating binding sites for the BRCT repeats in Dpb11p (Tanaka et al., 2007; Zegerman & Diffley, 2007). The Sld2p-Dpb11p interaction nucleates the formation of an unstable complex known as pre-LC (*pre-loading complex*), which contains the Pol  $\epsilon$  subunit Dpb2p, Mcm10p and GINS, in addition to Sld2p and Dpb11p (Muramatsu et al., 2010). Since phospho-Sld3p binds directly to Dpb11p and both Mcm10p and Sld3p interact with pre-RC components, this transient complex might be formed at the vicinity of licensed origins during S-phase. At the same time, DDK phosphorylates the N-terminal tail of Mcm4p at pre-RC to alleviate its auto-inhibitory activity (Sheu & Stillman, 2010) and to help stabilize Cdc45p binding to pre-RC (Sheu & Stillman, 2006). Stable association of Cdc45p is also facilitated by origin-bound Mcm10p (Sawyer et al., 2004). Since Mcm10p is a pre-LC component, origin-targeting of Pol  $\alpha$  is coupled to the rise of S-CDK activity.

In summary, GINS and Cdc45p are recruited for CMG complex assembly and helicase activation, and both Pol  $\alpha$  and Pol  $\epsilon$  are targeted to licensed origins by the joint actions of the S-phase kinases upon origin activation.

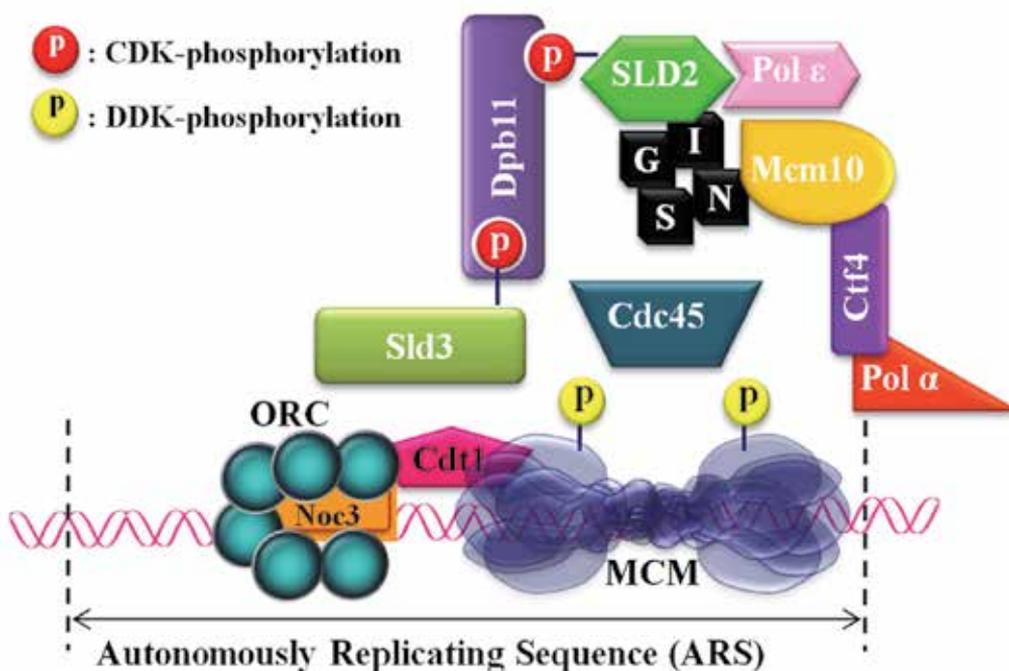


Fig. 2. The current model of origin activation in budding yeast.

At the beginning of S-phase, the concerted actions of CDK and DDK trigger origin-recruitment of both Cdc45p and GINS which bind MCM to form the active helicase. At the same time, polymerase  $\epsilon$  and  $\alpha$  are recruited for DNA synthesis. See text for details. Yellow and red circles represent DDK and CDK phosphorylation respectively.

## 2. CDK-dependent anti-rereplication mechanisms

Uncontrolled overreplication of the genome within the same cell cycle is detrimental to cell survival. Hence organisms possess multiple redundant mechanisms to prevent overreplication (Diffley, 2004). As shown in Fig. 3, central to this regulation lies at inhibitory CDK phosphorylations on Orc2p, Orc6p, Cdc6p and Mcm3p, which render these pre-RC components fail to relicense replication origins within the same cell cycle.

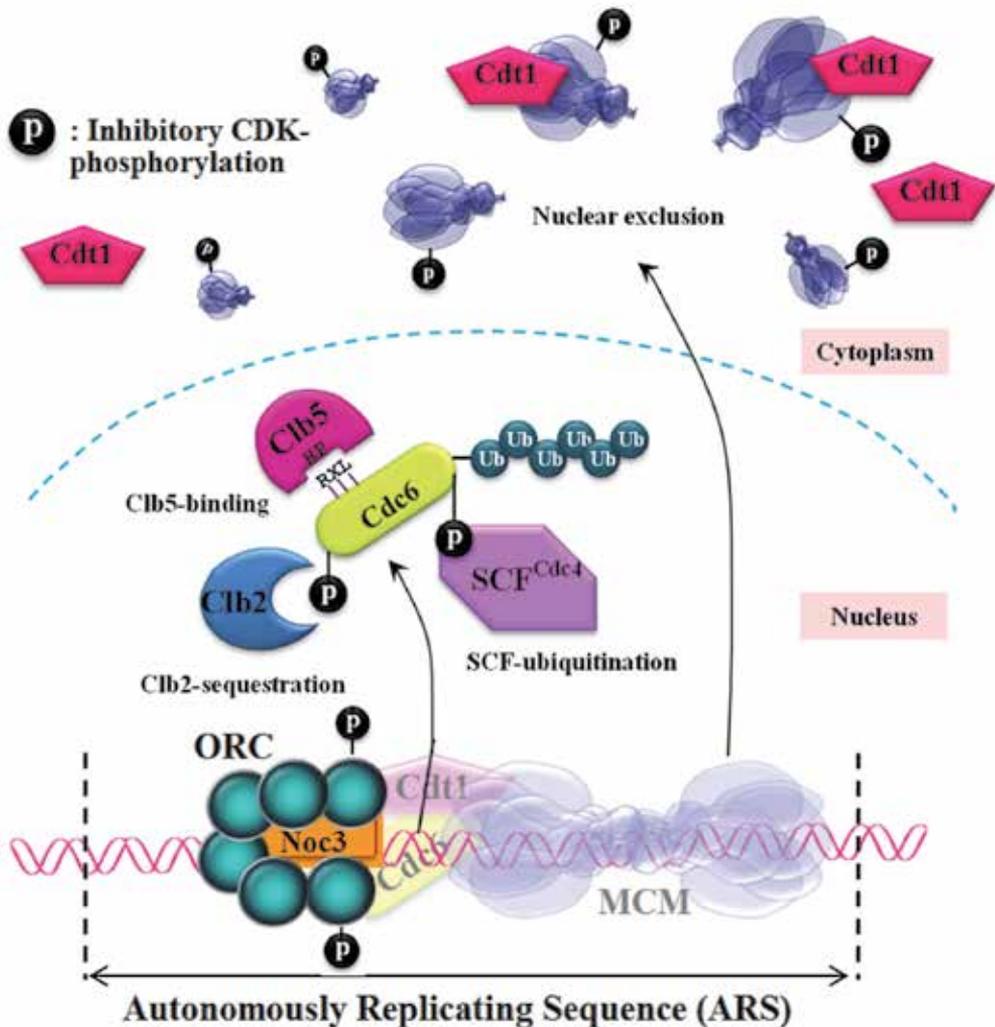


Fig. 3. Anti-rereplication mechanisms in budding yeast.

Anti-rereplication mechanisms in budding yeast targets multiple pre-RC components to render them unable to relicense replication origins within the same cell cycle. All of these mechanisms require CDK-phosphorylations. HP, hydrophobic patch. See text for details

### 2.1 Orc2p

Orc2p is a subunit of the hexameric ORC complex that binds replication origins throughout the cell cycle. Orc2p has six full CDK phosphorylation sites, and it is phosphorylated by Clb-CDK from S-phase until mitotic exit. *In vitro* studies suggest that phosphorylated Orc2p inhibits the ATP binding activity of Orc5p (Makise et al., 2009), and this might affect the overall activity of ORC to promote pre-RC assembly.

### 2.2 Orc6p

Orc6p shares the same phosphorylation profile as Orc2p during the cell cycle. In addition to the four full CDK consensus sites, Orc6p is found to be phosphorylated by CDK at a sub-optimal TP site at amino acid residue 114. Orc6p phosphorylation by CDK blocks the interaction between Cdt1p and Orc6p, thus largely suppressing the loading of Mcm2-7p complex onto replication origins. Furthermore, S-CDKs also bind to the RXL motif of Orc6p to produce a steric hindrance for MCM loading (Chen & Bell, 2011).

### 2.3 Cdc6p

Cdc6p is part of the MCM loader required for pre-RC assembly. From late M to early G1 phase, *de novo* synthesized Cdc6p binds to ORC and activates the ATPase activity of ORC to recruit the Mcm2-7p helicase onto replication origins (Randell et al., 2006; Chen et al., 2007). Cdc6p is a very unstable protein whose protein level is regulated by CDK phosphorylation and ubiquitin mediated proteolysis (Perkins et al., 2001). There are eight CDK consensus sites on Cdc6p; those at the N-terminal region are phosphorylated by Cln-CDK and targeted for robust SCF<sup>Cdc4p</sup>-mediated degradation during the G1/S transition (Drury et al., 1997). In G2/M cells, Cdc6p is phosphorylated by Clb-CDK and degraded through the SCF<sup>Cdc4p</sup> pathway in a much slower rate. And a moderate rate of Cdc6p proteolysis was also observed in alpha-factor blocked cells independent of CDK phosphorylation and the SCF complex (Drury et al., 2000). Although phosphorylation of the N-terminal CDK sites of Cdc6p does not form an SCF phospho-degron in G2/M phase, it creates a strong affinity site for Clb2p-binding that would dislodge Cdc6p from chromatin (Mimura et al., 2004). As such, Cdc6p is incompetent for pre-RC assembly whenever CDK activity persists.

### 2.4 Mcm3p

Mcm3p is a component of the hexameric MCM helicase that is loaded onto chromatin in the last step of pre-RC assembly. Subcellular localization of the MCM complex is cell cycle regulated (Labib et al., 1999). MCM is targeted to the nucleus from late M phase to G1 phase and largely exported to the cytoplasm from S-phase onwards until mitotic exit. Central to this regulation is the phosphorylation status of Mcm3p, which possesses five conserved CDK sites spanning over a NLS-NES (nuclear localization signal-nuclear exit signal) module near the N-terminus. The entire MCM complex contains a single bipartite NLS, which is distributed on two distinct MCM proteins Mcm2p and Mcm3p. Therefore, it is necessary for all six MCM subunits to be assembled into a whole complex in order to generate a functional NLS for nuclear import (Liku et al., 2005). On the other hand, there is an NES in Mcm3p adjacent to the NLS sequence. CDK phosphorylation of Mcm3p is required for nuclear export of the MCM complex, which probably downregulates the NLS activity and upregulates the NES activity (Liku et al., 2005). Therefore, nuclear export of the Mcm2-7p complex also provides a significant contribution to replication control (Nguyen et al., 2001; Liku et al., 2005).

Persistence of high CDK activity from late G1 to M phase ensures that activation of each origin occurs no more than once per cell cycle. The mechanisms that CDKs target multiple pre-RC components for inhibitory phosphorylations contribute to multiple overlaying strategies in cells to antagonize inappropriate DNA rereplication (Blow & Dutta, 2005). In a mutant strain (Nguyen et al., 2001) where all CDK sites of Orc2p and Orc6p were mutated to non-phosphorylatable alanine residues and the MCM complex was constitutively targeted into the nucleus, genome rereplication occurred upon overexpression of stabilized Cdc6p. This indicates that CDK inhibitions on ORC, Cdc6p and Mcm3p act in a redundant manner and each of these mechanisms is mostly sufficient to block the re-assembly of pre-RC during S-phase.

Although the strategies against rereplication vary in metazoans (Arias & Walter, 2005), most of them target pre-RC components. In mammalian cells, Orc1p is degraded by the ubiquitin-proteasome pathway during S-phase via the SCF<sup>Skp2p</sup> E3 ligase (Méndez et al., 2002). Besides, CDK phosphorylation on Orc1p inhibits its chromatin association during mitosis (Li & DePamphilis, 2002). Another initiation protein, the mammalian Cdt1p is targeted by multiple pathways and kept inactive during S and G2-phase. Cdt1p is targeted for degradation by the Cullin-based E3 ligase, Cul4p-Ddb1p<sup>Cdt2p</sup>. Because the degradation depends on the Cdt1p-PCNA interaction (Arias & Walter, 2006), this limits the timing of proteolysis only upon origin firing. Besides, Cdt1p is phosphorylated by CDK in S and G2 phases, which creates a phospho-degron recognized by the SCF<sup>Skp2p</sup> complex and is targeted for degradation (Takeda et al., 2005). Finally a protein known as Geminin was discovered as a binding inhibitor of Cdt1p. Geminin forms a complex with Cdt1p and renders it unable to load the MCM complex onto chromatin, thereby inhibiting pre-RC assembly (Wohlschlegel et al., 2000). Furthermore, a study suggested that Geminin also targets HBO1, which is a histone modifier required to acetylate origin-proximal H4 to promote MCM loading (Iizuka & Stillman, 1999). Geminin, when complexed with Cdt1p, acts as a potent inhibitor against the HBO1 acetyltransferase activity and hence prevents replication licensing (Miotto & Struhl, 2010).

### **3. The quest for a phosphatase to counteract anti-rereplication mechanisms and reset licensing competence during mitotic exit – studies on Cdc14p**

As discussed above, eukaryotic cells utilize elaborate intrinsic mechanisms to prevent DNA rereplication and ensure faithful inheritance of the genetic material. It is of equal importance to remove the blocks by dephosphorylating pre-RC proteins before cells undergo another cell cycle. Otherwise, existence of obstacles to DNA replication from the previous cell cycle may cause genome under-replication and chromosome instability. Cdc14p, a master phosphatase that promotes mitotic exit, fulfills the requirements of such a key regulator of DNA replication.

#### **3.1 Linkage between Cdc14p and DNA replication initiation**

Linkage between Cdc14p and DNA replication initiation has been implicated in a series of genetic data. Similar to all known initiation mutants reported, the plasmid loss of *cdc14-1* cells could be suppressed by adding multiple copies of an ARS to the plasmid (Hogan & Koshland, 1992; Ma et al., 2010). Overexpression of the initiation protein Orc6p caused synthetic dosage lethality in *cdc14-1* cells (Kroll et al., 1996), and synthetic lethality was observed when *cdc14-1* was combined with *cdc6-1*, *orc2-1* or *orc5-1* (Loo et al., 1995; Kroll et al., 1996; Yuste-Rojas & Cross, 2000). Orc6p was found to be dephosphorylated by

recombinant Cdc14p *in vitro* (Bloom & Cross, 2007). A recent study reported that genome-wide under-replication was observed in *cdc14-1* cells (Dulev et al., 2009). Recently, we show that Cdc14p, a dual-specificity phosphatase essential for mitotic exit, is responsible for resetting the competency of replication licensing by dephosphorylating multiple pre-RC proteins (Zhai et al., 2010).

### 3.2 Regulation of Cdc14p activation

Cdc14p is sequestered in the nucleolus for most part of the cell cycle, and it is only released and activated upon anaphase onset sequentially through the FEAR (Cdc fourteen early release) and MEN (*mitotic exit network*) pathways to counteract several CDK-dependent phosphorylations of mitotic substrates (Stegmeier & Amon, 2004; Sullivan & Morgan, 2007). As such, the functional window of Cdc14p coincides with the temporal profile of pre-RC assembly in the cell cycle.

During mitotic exit, a series of coordinated cellular events occur from anaphase onset to early G1 of the next cell cycle during which CDK activity is inhibited in a graded manner (Sullivan & Morgan, 2007). Cdc14p is the phosphatase responsible for driving these cell cycle events. Cdc14p released by FEAR at early anaphase mainly contributes to promoting onset of sister chromatids segregation, stabilization of anaphase spindle, establishment of spindle midzone, nucleolar segregation and MEN activation (Amon, 2008). The release of Cdc14p by the FEAR pathway is transient and restricted within the nucleus, and Cdc14p will be resequenced to the nucleolus if MEN pathway is defective (Rock & Amon, 2009). The MEN pathway, a Ras-like GTPase signaling cascade, results in complete activation of Cdc14p with the phosphatase located throughout the cell (Stegmeier & Amon, 2004). The main role of Cdc14p activated by the MEN pathway is to trigger CDK inactivation and drive mitotic exit (Amon, 2008). The FEAR and MEN pathways help to ensure a complete segregation of all chromosomes before the onset of mitotic exit and cytokinesis.

The important substrates of the MEN-activated Cdc14p include Cdh1p and Swi5p (*Switching deficient*) (Visintin et al., 1998). Cdh1p is the late mitotic co-activator of the APC E3 ligase, and it is inactive when phosphorylated by CDK. Dephosphorylation of Cdh1p by Cdc14p activates the APC<sup>Cdh1p</sup> activity which in turn targets mitotic cyclins for destruction. Parallel to this pathway, dephosphorylation of the transcription factor Swi5p triggers its nuclear import to activate *CDC6* and *SIC1* transcription. As a result, high level of APC activity and Sic1p accumulation lead to CDK inhibition.

### 3.3 The role of Cdc14p in resetting replication licensing

In a recent study, our group provided genetic and biochemical evidence to demonstrate the essential role of Cdc14p in dephosphorylating the initiation proteins Orc2p, Orc6p, Cdc6p and Mcm3p and resetting the competency of replication licensing during mitotic exit in budding yeast (Zhai et al., 2010).

In an effort to identify novel factors that are involved in or regulate DNA replication initiation, our lab has carried out a yeast phenotypic screen with randomly mutagenized yeast cells and have obtained several new hypomorphic alleles of *cdc14* mutants as well as mutants in other genes (Ma et al., 2010). These new *cdc14* mutants, like the previous *cdc14-1* (Hogan & Koshland, 1992) and all known replication-initiation mutants reported, have high rates of plasmid loss that can be suppressed by the presence of multiple copies of an ARS on the plasmid (Zou et al., 1997; Tye, 1999; Zhang et al., 2002; Ma et al., 2010). These results and

the previous genetic data described above prompted us to further examine the role of Cdc14p in DNA replication licensing.

Because of the parallel temporal profiles of pre-RC assembly and mitotic exit, it is difficult to perform loss-of-function experiments to examine the function of Cdc14p in replication licensing during the M/G1 transition. To circumvent the requirement of Cdc14p for mitotic exit, a mitotic rereplication system based on pulsed expression of an N-terminally truncated, relatively stable form of Sic1p (Sic1 $\Delta$ NT) was employed (Noton & Diffley, 2000). Sic1 $\Delta$ NT can inhibit mitotic CDK activity and hence drive pre-RC assembly when overexpressed from a galactose-inducible promoter (Dahmann et al., 1995). Resumption of CDK activity through subsequent repression of the *GAL* promoter and turnover of Sic1 $\Delta$ NT activates replication origins and causes rereplication when cells are still blocked in mitosis. By using this mitotic replication system, we showed that Cdc14p is essential for DNA replication by promoting pre-RC assembly.

Orc2p, Orc6p, Cdc6p and Mcm3p are known targets of CDKs (Ubersax et al., 2003). We showed that Cdc14p physically interacts with and dephosphorylates Orc2p, Orc6p, Cdc6p and Mcm3p *in vivo* and *in vitro*. Consistent with this, overexpression of Orc2p, Orc6p, Cdc6p or Mcm3p, but not the corresponding non-phosphorylatable (S/T to A) forms, resulted in severe growth defects in *cdc14-1* cells, probably by overloading the weakened Cdc14-1 phosphatase with one of its substrates. Inactivation of the Cdc14-3 protein also largely inhibited ORC dephosphorylation, nuclear localization of Mcm4-EGFP, and chromatin association of Cdc6p and MCM proteins.

To demonstrate that Orc2p, Orc6p, Cdc6p and Mcm3p are the major set of Cdc14p substrates whose dephosphorylations during mitotic exit promote replication licensing, we tested if combined non-phosphorylatable and/or phosphorylation-insensitive mutant alleles of these initiation proteins could bypass the function of Cdc14p in pre-RC assembly. A previous study showed that pre-RC assembly, origin firing and partial mitotic genome reduplication occurred upon ectopic expression of Cdc6 $\Delta$ NT in a quadruple mutant strain (*orc2-6A orc6-4A MCM7-NLS GAL-cdc6 $\Delta$ NT*) where the CDK consensus motifs in Orc2p and Orc6p were mutated to non-phosphorylatable forms, the MCM complex was constitutively nuclear-targeted and Cdc6p was expressed in a stabilized form (Nguyen et al., 2001). We compared DNA rereplication in *CDC14* wild-type and *cdc14-3* mutant cells in the quadruple (*orc2-6A orc6-4A MCM7-NLS GAL-cdc6 $\Delta$ NT*) mutant background upon Cdc6 $\Delta$ NT ectopic expression after Cdc14-3 inactivation in cells blocked in mitosis. Efficient DNA rereplication occurred in both *CDC14* and *cdc14-3* cells even when the Cdc14-3 protein was inactivated. These results strongly suggest that, of all initiation proteins, Orc2p, Orc6p, Cdc6p and Mcm3p represent the major set of Cdc14p substrates whose dephosphorylation is necessary for pre-RC assembly and DNA replication.

#### 4. Summary and discussion

In eukaryotes, replication licensing is achieved through sequential loading of several replication-initiation proteins onto replication origins to form pre-replicative complexes (pre-RCs). On the other hand, unscheduled replication licensing is prevented by cyclin-dependent kinases (CDKs) through inhibitory phosphorylations of multiple initiation proteins. It is known that CDK inactivation during mitotic exit promotes pre-RC formation for the next cell cycle. We have recently shown in budding yeast that Cdc14p dephosphorylates Orc2p, Orc6p, Cdc6p and Mcm3p to restore their competence for pre-RC

assembly (Zhai et al., 2010). Cells without functional Cdc14p failed to dephosphorylate the initiation proteins and to form pre-RCs even when CDK activities are inhibited, and they could not replicate DNA in the mitotic rereplication systems. On the other hand, pulsed ectopic expression of Cdc14p in mitotic cells resulted in efficient pre-RC assembly and DNA rereplication. Furthermore, Cdc14p becomes dispensable for DNA rereplication in mitotic cells with combined non-phosphorylatable / phosphorylation-insensitive alleles of the initiation proteins. These results unravel the essential role of Cdc14p in replication licensing besides its established functions for mitotic exit, providing new insight into the intricate regulation of DNA replication by the interplay between CDKs and the Cdc14p phosphatase.

#### 4.1 Proposed model of Cdc14p-driven origin licensing

Identification of multiple pre-RC components as a new set of Cdc14p substrates extends the physiological roles of this phosphatase outside the scope of mitotic exit. Based on the available data, we propose that Cdc14p provides a favorable environment for resetting the competency of replication licensing during mitotic exit by dephosphorylating multiple initiation proteins as well as other factors that promote the expression and/or stabilization of the initiation proteins (see diagrams in Fig. 4). During mitotic exit, Cdc14p dephosphorylates Swi5p, triggering nuclear localization of Swi5p which in turn induces the expression of Sic1p and Cdc6p (Piatti et al., 1995; Visintin et al., 1998). Cdc14p also stabilizes Sic1p and Cdc6p by keeping them in the dephosphorylated forms which are recognized by the SCF ubiquitination system (Verma et al., 1997; Perkins et al., 2001). Dephosphorylation of Cdh1p by Cdc14p activates APC<sup>Cdh1p</sup> which in turn mediates Clb2p degradation to free Cdc6p for pre-RC assembly (Visintin et al., 1998; Jaspersen et al., 1999; Mimura et al., 2004). Clb2p degradation and Sic1p accumulation also contribute to CDK inactivation (Stegmeier & Amon, 2004). Dephosphorylation of ORC by Cdc14p restores the ability of ORC to bind ATP (Makise et al., 2009) and other initiation proteins, and nuclear import of MCM complex is permitted when Mcm3p is dephosphorylated by Cdc14p. Altogether, these events promote pre-RC assembly in a step-wise manner onto replication origins for replication initiation.

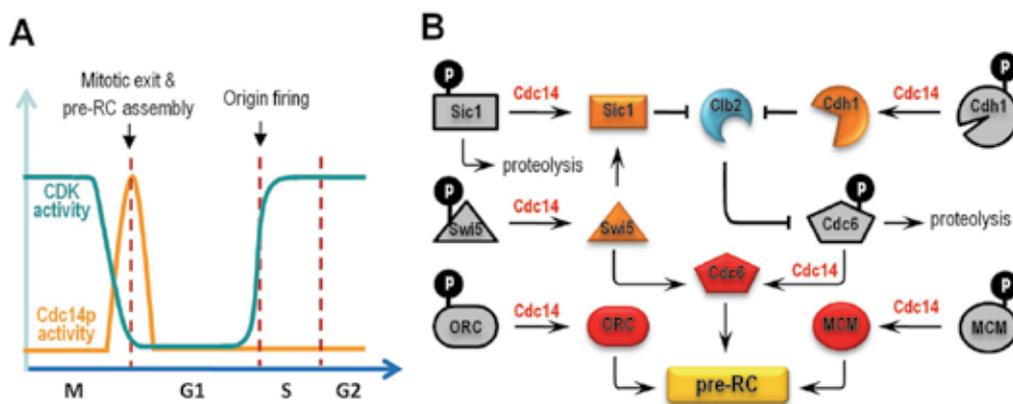


Fig. 4. Model depicting the essential role of Cdc14p in pre-RC assembly during mitotic exit. (A) Schematic diagram illustrating the change of CDK activity and the Cdc14p phosphatase activity in the cell cycle. (B) Diagram of Cdc14p substrates related to pre-RC assembly.

Phosphorylated, inactive Cdc14p substrates are shaded gray, while their dephosphorylated, active forms are highlighted in red (substrates identified by us in Zhai et al., 2010) or orange (previously identified substrates). Modified from Zhai et al., 2010. See text for details.

#### **4.2 Coordinating replication licensing with other mitotic events by Cdc14p**

Given the broad spectrum of Cdc14p targets and the sophisticated biphasic control of Cdc14p activation, Cdc14p may help to constitute a framework of a spatiotemporal program to coordinate different phosphorylation-regulated mitotic events. For example, it has been recently demonstrated that the histone acetyltransferase Gcn5p may facilitate replication licensing by inducing local decondensation at ARS-proximal regions (Espinosa et al., 2010). It might be worth testing whether the histone modifier Gcn5p and/or its origin-targeting factor are subject directly or indirectly to Cdc14p control. Genome-wide screening of Cdc14p substrates will likely shed light on this and other issues related to the coordination between replication licensing and mitotic exit.

#### **4.3 Conserved origin-resetting mechanism in other higher eukaryotes?**

Although CDK-dependent mechanisms against genome reduplication are diverse in eukaryotes (Arias & Walter, 2007), our findings in budding yeast underscore the importance of looking for phosphatases or other CDK-antagonizing activities that extinguish the inhibitory phosphorylations on initiation proteins for replication licensing in other organisms. There are two homologs of Cdc14p in humans known as Cdc14A and Cdc14B (Amon, 2008). However, their roles in mitotic exit remain controversial. Besides, it was suggested that the protein phosphatase-1 (PP1) might be the responsible phosphatase for driving mitotic exit (Wu et al., 2009). Hence it remains unknown whether or not Cdc14p homologs in higher eukaryotes also control replication licensing.

If the function of Cdc14p is conserved in human cells, the dual roles of Cdc14p in mitosis and DNA replication, and the intimate association of cancer with replication licensing (Shima et al., 2007) and the cell cycle (Hook et al., 2007), Cdc14p may promise to be an important protein to study in both normal and cancer cells.

#### **4.4 Implications in DNA amplification**

DNA amplification represents the most extreme case of re-replication in which only specific genomic loci undergo repeated origin firing while replication of the rest of the genome is inhibited by anti-rereplication mechanisms (Claycomb & Orr-Weaver, 2005). DNA amplification provides extra DNA templates to boost gene expression in response to developmental needs and in tumorigenesis. If Cdc14p or a functional analog exists in higher eukaryotes, targeting such phosphatase to amplification loci would contribute to establishing a restricted zone of dephosphorylating activity for pre-RC assembly. If this is true, one could imagine that such phosphatase should only selectively act against inhibitory phosphorylations of pre-RC proteins, but not affect origin activation by CDK.

### **5. References**

Amon, A. (2008). A decade of Cdc14 - A personal perspective. Delivered on 9 July 2007 at the 32nd FEBS Congress in Vienna, Austria. *FEBS Journal* 275, 5774-5784.

- Aparicio, O. M.; Weinstein, D. M. & Bell, S. P. (1997). Components and dynamics of DNA replication complexes in *S. cerevisiae*: Redistribution of MCM proteins and Cdc45p during S phase. *Cell* 91, 59-69.
- Araki, H.; Leem, S. H.; Phongdara, A. & Sugino, A. (1995). Dpb11, which interacts with DNA polymerase II([epsi]) in *Saccharomyces cerevisiae*, has a dual role in S-phase progression and at a cell cycle checkpoint. *Proceedings of the National Academy of Sciences of the United States of America* 92, 11791-11795.
- Arias, E. E. & Walter, J. C. (2005). Replication-dependent destruction of Cdt1 limits DNA replication to a single round per cell cycle in *Xenopus* egg extracts. *Genes and Development* 19, 114-126.
- Arias, E. E. & Walter, J. C. (2006). PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. *Nature Cell Biology* 8, 84-90.
- Arias, E. E. & Walter, J. C. (2007). Strength in numbers: Preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev.* 21, 497-518.
- Bloom, J. & Cross, F. R. (2007). Novel role for Cdc14 sequestration: Cdc14 dephosphorylates factors that promote DNA replication. *Mol. Biol. Cell* 27, 842-853.
- Blow, J. J. & Dutta, A. (2005). Preventing re-replication of chromosomal DNA. *Nature Rev. Mol. Cell Biol.* 6, 476-486.
- Bochman, M. L. & Schwacha, A. (2009). The Mcm complex: Unwinding the mechanism of a replicative helicase. *Microbiology and Molecular Biology Reviews* 73, 652-683.
- Boskovic, J.; Coloma, J.; Aparicio, T.; Zhou, M.; Robinson, C. V.; Méndez, J. & Montoya, G. (2007). Molecular architecture of the human GINS complex. *EMBO Reports* 8, 678-684.
- Bowers, J. L.; Randell, J. C. W.; Chen, S. & Bell, S. P. (2004). ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. *Molecular Cell* 16, 967-978.
- Chen, S. & Bell, S. P. (2011). CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes and Development* 25, 363-372.
- Chen, S.; De Vries, M. A. & Bell, S. P. (2007). Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. *Genes and Development* 21, 2897-2907.
- Claycomb, J. M. & Orr-Weaver, T. L. (2005). Developmental gene amplification: Insights into DNA replication and gene expression. *Trends in Genetics* 21, 149-162.
- Cocker, J. H.; Piatti, S.; Santocanale, C.; Nasmyth, K. & Diffley, J. F. X. (1996). An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. *Nature* 379, 180-182.
- Dahmann, C.; Diffley, J. F. & Nasmyth, K. A. (1995). S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state. *Curr. Biol.* 5, 1257-1269.
- Diffley, J. F. (2004). Regulation of early events in chromosome replication. *Curr. Biol.* 14, R778-R786.
- Donovan, S.; Harwood, J.; Drury, L. S. & Diffley, J. F. X. (1997). Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *Proceedings of the National Academy of Sciences of the United States of America* 94, 5611-5616.
- Drury, L. S.; Perkins, G. & Diffley, J. F. X. (1997). The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast. *EMBO J.* 16, 5966-5976.
- Drury, L. S.; Perkins, G. & Diffley, J. F. X. (2000). The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. *Current Biology* 10, 231-240.

- Du, Y.C. & Stillman, B. (2002) Yph1p, an ORC-interacting protein: potential links between cell proliferation control, DNA replication, and ribosome biogenesis. *Cell* 109, 835-848.
- Dulev, S.; De Renty, C.; Mehta, R.; Minkov, I.; Schwob, E. & Strunnikov, A. (2009). Essential global role of CDC14 in DNA synthesis revealed by chromosome underreplication unrecognized by checkpoints in *cdc14* mutants. *Proceedings of the National Academy of Sciences of the United States of America* 106, 14466-14471.
- Erzberger, J. P.; Mott, M. L. & Berger, J. M. (2006). Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nature Structural and Molecular Biology* 13, 676-683.
- Espinosa, M. C.; Rehman, M. A.; Chisamore-Robert, P.; Jeffery, D. & Yankulov, K. (2010). GCN5 is a positive regulator of origins of DNA replication in *Saccharomyces cerevisiae*. *PLoS ONE* 5.
- Evrin, C.; Clarke, P.; Zech, J.; Lurz, R.; Sun, J.; Uhle, S.; Li, H.; Stillman, B. & Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc Natl Acad Sci* 106, 20240-20245.
- Gambus, A.; Van Deursen, F.; Polychronopoulos, D.; Foltman, M.; Jones, R. C.; Edmondson, R. D.; Calzada, A. & Labib, K. (2009). A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase  $\alpha$  within the eukaryotic replisome. *EMBO Journal* 28, 2992-3004.
- Gambus, A. F.; Polychronopoulos, D.; Foltman, M.; Jones, R. C.; Edmondson, R. D.; Calzada, A. & Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nature Cell Biol.* 8, 358-366.
- Gibson, D.G., Bell, S.P. & Aparicio, O.M. (2006) Cell cycle execution point analysis of ORC function and characterization of the checkpoint response to ORC inactivation in *Saccharomyces cerevisiae*. *Genes Cells* 11, 557-573.
- Hanna, J. S.; Kroll, E. S.; Lundblad, V. & Spencer, F. A. (2001). *Saccharomyces cerevisiae* CTF18 and CTF4 are required for sister chromatid cohesion. *Molecular and Cellular Biology* 21, 3144-3158.
- Hennessy, K. M.; Lee, A.; Chen, E. & Botstein, D. (1991). A group of interacting yeast DNA replication genes. *Genes and Development* 5, 958-969.
- Hogan, E. & Koshland, D. (1992). Addition of extra origins of replication to a minichromosome suppresses its mitotic loss in *cdc6* and *cdc14* mutants of *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* 89, 3098-3102.
- Honey, S. & Futcher, B. (2007). Roles of the CDK phosphorylation sites of yeast Cdc6 in chromatin binding and rereplication. *Mol. Biol. Cell* 18, 1324-1336.
- Hook, S. S.; Lin, J. J. & Dutta, A. (2007). Mechanisms to control rereplication and implications for cancer. *Curr. Opin. Cell Biol.* 19, 663-671.
- Houchens, C.R., Perreault, A., Bachand, F. & Kelly T.J. (2008) *Schizosaccharomyces pombe* Noc3 is essential for ribosome biogenesis and cell division but not DNA replication. *Eukaryot. Cell* 7, 1433-1440.
- Iizuka, M. & Stillman, B. (1999). Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein. *Journal of Biological Chemistry* 274, 23027-23034.
- Ilves, I.; Petojevic, T.; Pesavento, J. J. & Botchan, M. R. (2010) Activation of the MCM2-7 Helicase by Association with Cdc45 and GINS Proteins. *Molecular Cell* 37, 247-258.
- Jaspersen, S. L.; Charles, J. F. & Morgan, D. O. (1999). Inhibitory phosphorylation of the APC regulator Hct1 is controlled by the kinase Cdc28 and the phosphatase Cdc14. *Curr. Biol.* 9, 227-236.

- Johmura, Y.; Osada, S.; Nishizuka, M. & Imagawa, M. (2008a). FAD24 acts in concert with histone acetyltransferase HBO1 to promote adipogenesis by controlling DNA replication. *Journal of Biological Chemistry* 283, 2265-2274.
- Johmura, Y.; Osada, S.; Nishizuka, M. & Imagawa, M. (2008b). FAD24, a regulator of adipogenesis, is required for the regulation of DNA replication in cell proliferation. *Biological and Pharmaceutical Bulletin* 31, 1092-1095.
- Jorgensen, P. & Tyers, M. (2004). How cells coordinate growth and division. *Current Biology* 14, R1014-R1027.
- Kamimura, Y.; Masumoto, H.; Sugino, A. & Araki, H. (1998). Sld2, which interacts with Dpb11 in *Saccharomyces cerevisiae*, is required for chromosomal DNA replication. *Molecular and Cellular Biology* 18, 6102-6109.
- Kamimura, Y.; Tak, Y. S.; Sugino, A. & Araki, H. (2001). Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in *saccharomyces cerevisiae*. *EMBO Journal* 20, 2097-2107.
- Kan, J.; Zou, L.; Zhang, J.; Wu, R.; Wang, Z. & Liang, C. (2008). Origin recognition complex (ORC) mediates histone 3 lysine 4 methylation through cooperation with Spp1 in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 283, 33803-33807.
- Kanemaki, M.; Sanchez-Diaz, A.; Gambus, A. & Labib, K. (2003). Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. *Nature* 423, 720-724.
- Kroll, E. S.; Hyland, K. M.; Hieter, P. & Li, J. J. (1996). Establishing genetic interactions by a synthetic dosage lethality phenotype. *Genetics* 143, 95-102.
- Kubota, Y.; Takase, Y.; Komori, Y.; Hashimoto, Y.; Arata, T.; Kamimura, Y.; Araki, H. & Takisawa, H. (2003). A novel ring-like complex of *Xenopus* proteins essential for the initiation of DNA replication. *Genes and Development* 17, 1141-1152.
- Labib, K. (2010). How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? *Genes and Development* 24, 1208-1219.
- Labib, K.; Diffley, J. F. & Kearsley, S. E. (1999). G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. *Nature Cell Biol.* 1, 415-422.
- Lee, C.; Liachko, I.; Bouten, R.; Kelman, Z. & Tye, B. K. (2010). Alternative mechanisms for coordinating polymerase and MCM helicase. *Molecular and Cellular Biology* 30, 423-435.
- Li, C. J. & DePamphilis, M. L. (2002). Mammalian Orc1 protein is selectively released from chromatin and ubiquitinated during the S-to-M transition in the cell division Cycle. *Molecular and Cellular Biology* 22, 105-116.
- Liang, C.; Weinreich, M. & Stillman, B. (1995). ORC and Cdc6p interact and determine the frequency of initiation of DNA replication in the genome. *Cell* 81, 667-676.
- Liku, M. E.; Nguyen, V. Q.; Rosales, A. W.; Irie, K. & Li, J. J. (2005). CDK Phosphorylation of a Novel NLS-NES Module Distributed between Two Subunits of the Mcm2-7 Complex Prevents Chromosomal Rereplication. *Mol. Biol. Cell* 16, 5026-5039.
- Lipford, J. R. & Bell, S. P. (2001). Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Molecular Cell* 7, 21-30.
- Loo, S.; Fox, C. A.; Rine, J.; Kobayashi, R.; Stillman, B. & Bell, S. (1995). The origin recognition complex in silencing, cell cycle progression, and DNA replication. *Mol. Biol. Cell* 6, 741-756.
- Méndez, J.; Zou-Yang, X. H.; Kim, S. Y.; Hidaka, M.; Tansey, W. P. & Stillman, B. (2002). Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. *Molecular Cell* 9, 481-491.

- Ma, L.; Zhai, Y.; Feng, D.; Chan, T. C.; Lu, Y.; Fu, X.; Wang, J.; Chen, Y.; Li, J.; Xu, K. et al. (2010). Identification of novel factors involved in or regulating initiation of DNA replication by a genome-wide phenotypic screen in *Saccharomyces cerevisiae*. *Cell Cycle* 9, 4399-4410.
- Machida, Y. J.; Hamlin, J. L. & Dutta, A. (2005). Right place, right time, and only once: Replication initiation in metazoans. *Cell* 123, 13-24.
- Maine, G. T.; Sinha, P. & Tye, B. W. (1984). Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics* 106, 365-385.
- Maiorano, D.; Moreau, J. & Méchali, M. (2000). XCDT1 is required for the assembly of pre-replicative complexes in *Xenopus laevis*. *Nature* 404, 622-625.
- Makise, M.; Takehara, M.; Kuniyasu, A.; Matsui, N.; Nakayama, H. & Mizushima, T. (2009). Linkage between phosphorylation of the origin recognition complex and its ATP binding activity in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 284, 3396-3407.
- Milkereit, P.; Gadal, O.; Podtelejnikov, A.; Trumtel, S.; Gas, N.; Petfalski, E.; Tollervey, D.; Mann, M.; Hurt, E. & Tschochner, H. (2001). Maturation and intranuclear transport of pre-ribosomes requires Noc proteins. *Cell* 105, 499-509.
- Mimura, S.; Seki, T.; Tanaka, S. & Diffley, J. F. X. (2004). Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control. *Nature* 431, 1118-1123.
- Miotto, B. & Struhl, K. (2010). HBO1 Histone Acetylase Activity Is Essential for DNA Replication Licensing and Inhibited by Geminin. *Molecular Cell* 37, 57-66.
- Muramatsu, S.; Hirai, K.; Tak, Y. S.; Kamimura, Y. & Araki, H. (2010). CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol  $\epsilon$ , and GINS in budding yeast. *Genes and Development* 24, 602-612.
- Nguyen, V. Q.; Co, C. & Li, J. J. (2001). Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* 411, 1068-1073.
- Nishitani, H.; Lygerou, Z.; Nishimoto, T. & Nurse, P. (2000). The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature* 404, 625-628.
- Noton, E. A. & Diffley, J. F. X. (2000). CDK inactivation is the only essential function of the APC/C and the mitotic exit network proteins for origin resetting during mitosis. *Mol. Cell* 5, 85-95.
- Perkins, G.; Drury, L. S. & Diffley, J. F. X. (2001). Separate SCFCDC4 recognition elements target Cdc6 for proteolysis in S phase and mitosis. *EMBO J.* 20, 4836-4845.
- Piatti, S.; Lengauer, C. & Nasmyth, K. (1995). Cdc6 is an unstable protein whose de novo synthesis in G1 is important for the onset of S phase and for preventing a 'reductional' anaphase in the budding yeast *Saccharomyces cerevisiae*. *EMBO J.* 14, 3788-3799.
- Raghuraman, M. K.; Winzeler, E. A.; Collingwood, D.; Hunt, S.; Wodicka, L.; Conway, A.; Lockhart, D. J.; Davis, R. W.; Brewer, B. J. & Fangman, W. L. (2001). Replication dynamics of the yeast genome. *Science* 294, 115-121.
- Randell, J. C. W.; Bowers, J. L.; Rodríguez, H. K. & Bell, S. P. (2006). Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Molecular Cell* 21, 29-39.
- Remus, D.; Beuron, F.; Tolun, G.; Griffith, J. D.; Morris, E. P. & Diffley, J. F. X. (2009). Concerted Loading of Mcm2-7 Double Hexamers around DNA during DNA Replication Origin Licensing. *Cell* 139, 719-730.
- Ricke, R. M. & Bielinsky, A. K. (2004). Mcm10 regulates the stability and chromatin association of DNA polymerase- $\alpha$ . *Molecular Cell* 16, 173-185.
- Rock, J. M. & Amon, A. (2009). The FEAR network. *Current Biology* 19, R1063-R1068.

- Sawyer, S. L.; Cheng, I. H.; Chai, W. & Tye, B. K. (2004). Mcm10 and Cdc45 cooperate in origin activation in *Saccharomyces cerevisiae*. *Journal of Molecular Biology* 340, 195-202.
- Sclafani, R. A. & Holzen, T. M. (2007). Cell cycle regulation of DNA replication. *Annu. Rev. Genet.* 41, 237-280.
- Seki, T. & Diffley, J. F. X. (2000). Stepwise assembly of initiation proteins at budding yeast replication origins in vitro. *Proceedings of the National Academy of Sciences of the United States of America* 97, 14115-14120.
- Sheu, Y. J. & Stillman, B. (2006). Cdc7-Dbf4 Phosphorylates MCM Proteins via a Docking Site-Mediated Mechanism to Promote S Phase Progression. *Molecular Cell* 24, 101-113.
- Sheu, Y. J. & Stillman, B. (2010). The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature* 463, 113-117.
- Shima, N.; Alcaraz, A.; Liachko, I.; Buske, T. R.; Andrews, C. A.; Munroe, R. J.; Hartford, S. A.; Tye, B. K. & Schimenti, J. C. (2007). A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nature Genetics* 39, 93-98.
- Speck, C. & Stillman, B. (2007). Cdc6 ATPase activity regulates ORCxCdc6 stability and the selection of specific DNA sequences as origins of DNA replication. *Journal of Biological Chemistry* 282, 11705-11714.
- Speck, C.; Chen, Z.; Li, H. & Stillman, B. (2005). ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nature Structural and Molecular Biology* 12, 965-971.
- Spencer, F.; Gerring, S. L.; Connelly, C. & Hieter, P. (1990). Mitotic chromosome transmission fidelity mutants in *Saccharomyces cerevisiae*. *Genetics* 124, 237-249.
- Stegmeier, F. & Amon, A. (2004). Closing mitosis: the functions of the Cdc14 phosphatase and its regulation. *Annu. Rev. Genet.* 38, 203-232.
- Stillman, B. (2005). Origin recognition and the chromosome cycle. *FEBS Letters* 579, 877-884.
- Sullivan, M. & Morgan, D. O. (2007). Finishing mitosis, one step at a time. *Nature Rev. Mol. Cell Biol.* 8, 894-903.
- Takahashi, T. S.; Wigley, D. B. & Walter, J. C. (2005). Pumps, paradoxes and ploughshares: Mechanism of the MCM2-7 DNA helicase. *Trends in Biochemical Sciences* 30, 437-444.
- Takayama, Y.; Kamimura, Y.; Okawa, M.; Muramatsu, S.; Sugino, A. & Araki, H. (2003). GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes and Development* 17, 1153-1165.
- Takeda, D. Y.; Parvin, J. D. & Dutta, A. (2005). Degradation of Cdt1 during S phase is Skp2-independent and is required for efficient progression of mammalian cells through S phase. *Journal of Biological Chemistry* 280, 23416-23423.
- Tanaka, H.; Katou, Y.; Yagura, M.; Saitoh, K.; Itoh, T.; Araki, H.; Bando, M. & Shirahige, K. (2009). Ctf4 coordinates the progression of helicase and DNA polymerase  $\alpha$ . *Genes to Cells* 14, 807-820.
- Tanaka, S. & Diffley, J. F. X. (2002). Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nature Cell Biology* 4, 198-207.
- Tanaka, S.; Umemori, T.; Hirai, K.; Muramatsu, S.; Kamimura, Y. & Araki, H. (2007). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* 445, 328-332.
- Tanaka, T.; Knapp, D. & Kim, N. (1997). Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs. *Cell* 90, 649-660.
- Tominaga, K.; Johmura, Y.; Nishizuka, M. & Imagawa, M. (2004). Fad24, a mammalian homolog of Noc3p, is a positive regulator in adipocyte differentiation. *Journal of Cell Science* 117, 6217-6226.

- Tye, B. K. (1999). Minichromosome maintenance as a genetic assay for defects in DNA replication. *Methods: A Companion to Methods in Enzymology* 18, 329-334.
- Ubersax, J. A.; Woodbury, E. L.; Quang, P. N.; Paraz, M.; Blethrow, J. D.; Shah, K.; Shokat, K. M. & Morgan, D. O. (2003). Targets of the cyclin-dependent kinase Cdk1. *Nature* 425, 859-864.
- Verma, R.; Annan, R. S.; Huddleston, M. J.; Carr, S. A.; Reynard, G. & Deshaies, R. J. (1997). Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S phase. *Science* 278, 455-460.
- Visintin, R.; Craig, K.; Hwang, E. S.; Prinz, S.; Tyers, M. & Amon, A. (1998). The phosphatase Cdc14 triggers mitotic exit by reversal of Cdk-dependent phosphorylation. *Mol. Cell* 2, 709-718.
- Wang, J.; Wu, R.; Lu, Y. & Liang, C. (2010). Ctf4p facilitates Mcm10p to promote DNA replication in budding yeast. *Biochemical and Biophysical Research Communications* 395, 336-341.
- Wei, Z.; Liu, C.; Wu, X.; Xu, N.; Zhou, B.; Liang, C. & Zhu, G. (2010). Characterization and structure determination of the cdt1 binding domain of human minichromosome maintenance (Mcm) 6. *Journal of Biological Chemistry* 285, 12469-12473.
- Whittaker, A. J.; Royzman, I. & Orr-Weaver, T. L. (2000). Drosophila double parked: A conserved, essential replication protein that colocalizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts. *Genes and Development* 14, 1765-1776.
- Wohlschlegel, J. A.; Dwyer, B. T.; Dhar, S. K.; Cvetcic, C.; Walter, J. C. & Dutta, A. (2000). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science* 290, 2309-2312.
- Wu, J. Q.; Guo, J. Y.; Tang, W.; Yang, C. S.; Freel, C. D.; Chen, C.; Nairn, A. C. & Kornbluth, S. (2009). PP1-mediated dephosphorylation of phosphoproteins at mitotic exit is controlled by inhibitor-1 and PP1 phosphorylation. *Nature Cell Biology* 11, 644-651.
- Yuste-Rojas, M. & Cross, F. R. (2000). Mutations in CDC14 results in high sensitivity to cyclin gene dosage in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 263, 60-72.
- Zegerman, P. & Diffley, J. F. X. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* 445, 281-285.
- Zhai, Y.; Yung, P. Y. K.; Huo, L. & Liang, C. (2010). Cdc14p resets the competency of replication licensing by dephosphorylating multiple initiation proteins during mitotic exit in budding yeast. *Journal of Cell Science* 123, 3933-3943.
- Zhang, Y.; Yu, Z.; Fu, X. & Liang, C. (2002). Noc3p, a bHLH protein, plays an integral role in the initiation of DNA replication in budding yeast. *Cell* 109, 849-860.
- Zhu, W.; Ukomadu, C.; Jha, S.; Senga, T.; Dhar, S. K.; Wohlschlegel, J. A.; Nutt, L. K.; Kornbluth, S. & Dutta, A. (2007). Mcm10 and And-1/CTF4 recruit DNA polymerase  $\alpha$  to chromatin for initiation of DNA replication. *Genes and Development* 21, 2288-2299.
- Zou, L. & Stillman, B. (1998). Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin. *Science* 280, 593-596.
- Zou, L.; Mitchell, J. & Stillman, B. (1997). CDC45, a novel yeast gene that functions with the origin recognition complex and Mcm proteins in initiation of DNA replication. *Molecular and Cellular Biology* 17, 553-563.

# Cdc6 Knockdown Renders p16<sup>INK4a</sup> Re-Activation, Leading to Senescence Human Breast Carcinoma Cells

Luo Feng, Jerry R. Barnhart, Lingtao Wu, Greg Shackleford,  
Sheng-he Huang and Ambrose Jong  
*Department of Hematology and Oncology  
Childrens Hospital Los Angeles  
University of Southern California, Los Angeles  
USA*

## 1. Introduction

DNA replication is initiated by assembling pre-replicative complex (pre-RC), and DNA replicative protein Cdc6 plays a central role in the initiation of DNA replication. As cells enter G1 from mitosis, Cdc6, together with Cdt1 are recruited to the replicating origins bound with six components of origin recognition complex (ORC) (Dutta and Bell, 1997; Méndez and Stillman, 2000). The pre-RC is formed by loading minichromosome maintenance proteins (MCMs) along with other replicating proteins onto origins, and Cdc6 controls the loading. The stability of Cdc6 during G1 is a key step to regulate pre-RC assembly. Cdc6 is targeted to ubiquitin-mediated proteolysis by the anaphase promoting complex (APC) in G1. The cyclin-dependent kinase (CDK), specifically, cyclin E-Cdk2 kinase is responsible for phosphorylation of human Cdc6, and Cdc6 is stabilized with its amino-terminal end phosphorylated (Mailand and Diffley, 2005). The Cdc6 phosphorylation by CDKs ensures that origin is fired at S phase, and DNA is replicated only once per S phase. Human Cdc6 is also involved in G2/M phase checkpoint as overexpression of Cdc6 in G2 blocks HeLa cell into mitosis (Clay-Farrace et al., 2003). More recently, human Cdc6 shows to play a role in late S phase and S-G2/M transition in controlling cell cycle progression (Lau et al., 2006). These studies suggest that human Cdc6 functions beyond its essential role in the initiation of DNA replication.

Senescence is an irreversible cell event with permanent G1 arrest. Senescent normal somatic cells show features such as the permanent arrest of proliferation, repression of telomerase reverse transcriptase (TERT) expression, and expression of senescence associated- $\beta$ -galactosidase (SA- $\beta$ -Gal). The lack of TERT activities has been observed in most human somatic cells and is linking to replicative senescence (Campisi, 2001). Replicative senescence occurs with functional p53 and p16<sup>INK4a</sup>-Rb pathway (Campisi, 2001; Beauséjour et al., 2003). Accumulated evidence shows that tumor cells experience p53 mutation, p16<sup>INK4a</sup> or Rb inactivation or deletion, accompanied with the re-expression of TERT. Evidence shows that senescence normal fibroblasts undergo the formation of senescence-associated heterochromatic foci (SAHF), and tri- or di-methylation of lysine 9 on histone H3 (H3K9me3/me2) has been found in the SAHF (Narita et al., 2003, 2006).

High level of Cdc6 has been associated with the oncogenic activities in human cancers, and there is a reciprocal association in protein levels between Cdc6 and p16<sup>INK4a</sup> (Gonzalez et al., 2006). Genetically, human Cdc6 suppresses p16<sup>INK4a</sup> functions. Cdc6 shows dual effects on the *INK4/ARF* locus: it associates with the replication origin on chromatin together with ORCs for loading MCM proteins, and it binds a region called regulatory domain (RD<sup>INK4/ARF</sup>) in the putative replication origin. This association results in transcriptional repression of the three genes encoding for p15<sup>INK4b</sup>, p16<sup>INK4a</sup>, and p14<sup>ARF</sup>, leading to heterochromatinization of the *INK4/ARF* locus (Gonzalez et al., 2006). The gene products from the *INK4/ARF* locus are thought as barriers to immortalization and oncogenesis. p16<sup>INK4a</sup> inhibits D-type cyclins-Cdk4/Cdk6 kinases, which activates p16<sup>INK4a</sup>-Rb pathway in controlling cell cycle progression.

## 2. Cdc6 is a decisive factor in promoting cell proliferation

Cdc6 is essential in the initiation of DNA replication. Human Cdc6 function is of importance in cell cycle progression and cell proliferation. In human cells, the levels of Cdc6 change markedly in response to various stresses. DNA re-replication has been found in those Cdc6 and Cdt1 overexpressing normal or tumor cells. Cdc6 could be more stable in tumor cells, due to p53 deficiency, and due to strong Cdc6 phosphorylation by activated cyclin E-Cdk2 kinase (Bartkova et al., 2005; Duursma and Agami, 2005; Gorgoulis et al., 2005; Mailand and Diffley, 2005).

### 2.1 Oncogenic Cdc6 is involved in replicative stress response

Recently, a report that p53 downregulates Cdc6 in response to DNA damage links Cdc6 directly to p53 mediated replication stress response (Duursma and Agami, 2005). The p53-mediated Cdc6 downregulation takes place without causing detectable cell cycle redistribution. The inhibition of endogenous Cdc6 via Cdc6 knockdown, however, results in an increased G1 content with decreased S fraction in cells with p53 deficiency (Duursma and Agami, 2005).

Human Cdc6 has been involved in cancer cell growth, while the depletion of Cdc6 induces apoptosis in several cancer cell lines, but not in normal cells. The siRNA or gene transduction vectors producing short hairpin RNA (shRNA) mediated Cdc6 depletion occurs specifically without disturbing other replicative proteins, such as Orc2, Cdt1, and MCM7 (Lau et al., 2006 ).

Here, we report the reactivation of p16<sup>INK4a</sup>-Rb pathway and senescence human breast carcinoma MCF7 cells with Cdc6 knockdown.

#### 2.1.1 Overexpression of Cdc6 in human cancer cell lines

We have extended our research to explore human Cdc6 functions. We first examined Cdc6 overexpression in several human cancer cell lines including neuroblastoma LA-N-2 and CHLA255 cells, MCF7 cells, and HeLa cells. Cells were transfected with tetracycline-inducible Cdc6 expressing vector, or infected with a retrovirus for expressing Cdc6. Our results showed that Cdc6 overexpression in some cancer cell lines could not induce significant Chk1 or Chk2 phosphorylation. Cdc6 overexpression resulted in, however, p53 upregulation. At the present moment, we haven't observed senescence these cancer cells by overexpressing Cdc6. Instead, DNA re-replication and DNA damages caused by Cdc6

overexpression makes DNA replication checkpoints compromised, which might be related to unlimited cancer cell proliferation.

### 2.1.2 Generating a human Immunodeficiency virus type 1 (HIV-1) derived lentivirus expressing *CDC6* short hairpin RNA (shRNA)

We have constructed an HIV-1 derivative vector (THTD) for producing *CDC6* shRNA. The details of the lentivirus vector are published. A control lentivirus for producing a scramble sequence (THTP) is also generated. Infection of THTD in various kinds of human cell lines causes the depletion of Cdc6. The inhibition of endogenous Cdc6 expression is over 98% (Figure 1., lane 3, 6, and 8). However, infection of THTP shows no change of Cdc6 expression (Figure 1., lane 2, 5, and 7). The recombinant lentivirus, therefore, becomes a powerful vector for us to study Cdc6 function.

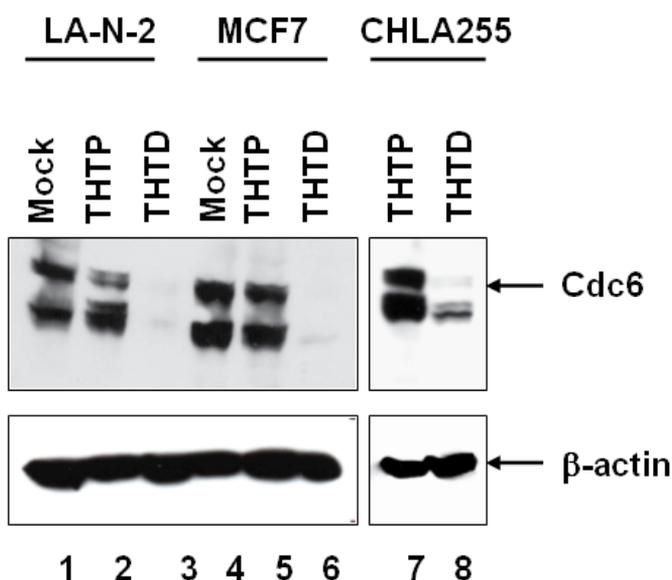


Fig. 1. Inhibition of Cdc6 expression via recombinant lentivirus producing *CDC6* shRNA .

LA-N-2 cells, CHLA-255 cells, and MCF7 cells were infected with THTD for Cdc6 knockdown (lane 3, 6 and 8), or infected with THTP (lane 2, 5 and 7), or mock infected (lane 1 and 4) at m.o.i. of 1.0. Two days postinfection, the cell extracts were prepared. About 20  $\mu$ g proteins in extracts from each sample were assayed in Western blots probed with indicated antibodies.

### 2.1.3 Cdc6 knockdown induces apoptosis of human neuroblastoma cells

Elevated levels of Cdc6 were found in the LA-N-2, CHLA255, and other human neuroblastoma cell lines that grew fast. Cdc6 knockdown via THTD infection caused the accumulation of sub G1 populations with declined S contents in the LA-N-2 and CHLA255 cells. The cells were stained positive with Annex V and 7-AAD by 31% (Figure 2.). Expression profile from the selected genes showed the reduction of cell cycle progression proteins such as cyclin E, cyclin A, and Cdc25C with a boosted increase of CDK inhibitor

p27<sup>Kip1</sup>, indicating the suppression of tumor cell proliferation. Further, Cdc6 knockdown caused the increase of pro-apoptotic Bax accompanied with the decrease of anti-apoptotic Bcl-2, resulting in massive cell death. Our study indicates that human Cdc6 functions in several pathways to control cell proliferation and the cell death.

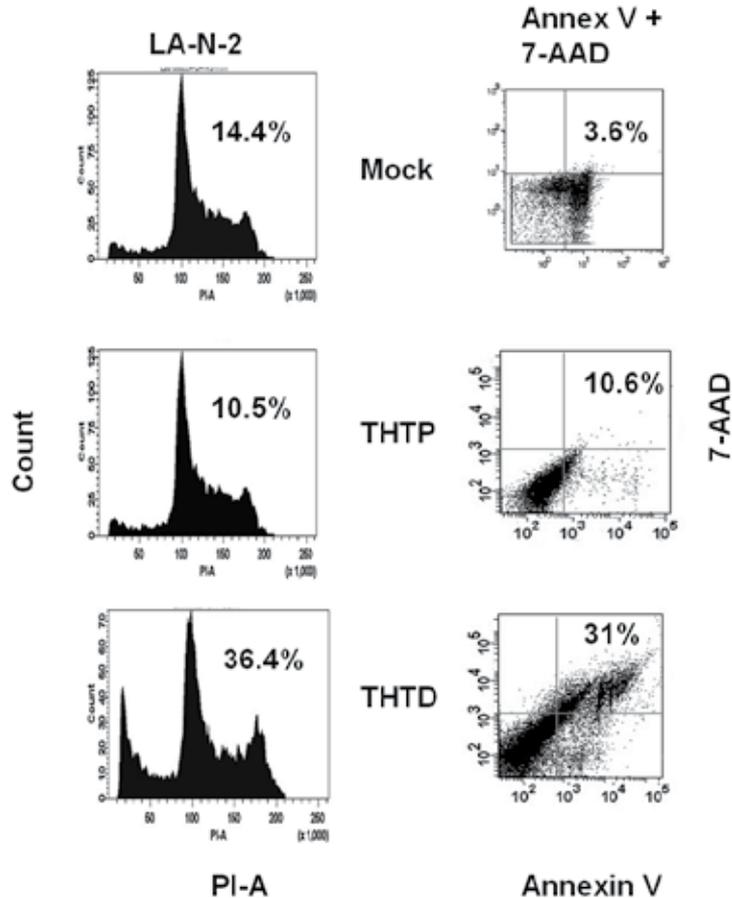


Fig. 2. Cdc6 knockdown induces apoptosis of human neuroblastoma cells.

LA-N-2 cells were infected with the *CDC6* shRNA lentivirus (THTD), or the control virus with scramble sequence (THTP), or mock infected. Left panel: About  $0.5 \times 10^6$  cells were harvested 48 hours post infection, and stained with propidium iodide (PI) for flow cytometry analysis. The data presented were the average values from 4 separated experiments with the standard deviation between 5-15%. Right panel: The cells were harvested 48 hours post infection, and about  $0.2 \times 10^6$  cells were used for Annexin V-PE and 7-ADD double staining. The flow cytometry analysis was performed for the stained cells. Fluorescence emission was detected at  $575 \pm 13$  for Annexin V-PE, and  $660 \pm 10$  for 7-AAD.

#### 2.1.4 Depletion of Cdc6 renders senescence MCF7 cancer cells

We have found apoptosis of human neuroblastoma LA-N-2 and CHLA255 cells with Cdc6 knockdown. However, Cdc6 knockdown could not confer the programmed cell death in

MCF7 cells. Instead, Cdc6 knockdown induced the cancer cell senescence. Cell cycle profile showed that Cdc6 knockdown caused G1 contents increased by 30% along with declined S populations. At the same condition analyzed, MCF7 cells infected with the control virus showed no decline of S populations, neither the increase of the G1 contents. Brd (U) incorporation was performed to determine the cell proliferation status. Fluorescent signals from incorporated Brd (U) in DNA in the nucleus could be observed in both mock and THTP infected MCF7 cells, but not in the Cdc6 knockdown cells infected with THTD, indicating no DNA replication in the Cdc6 depleted cells.

The Cdc6 depleted MCF7 cells were subjected to SA- $\beta$ -Gal staining 48 hours postinfection. While the mock and the control virus infected MCF7 showed a background staining, the Cdc6 depleted cancer cells displayed positive staining for SA- $\beta$ -Gal, a hall marker for senescence (Figure 3A.). Because senescence cells undergo heterochromatin formation, a

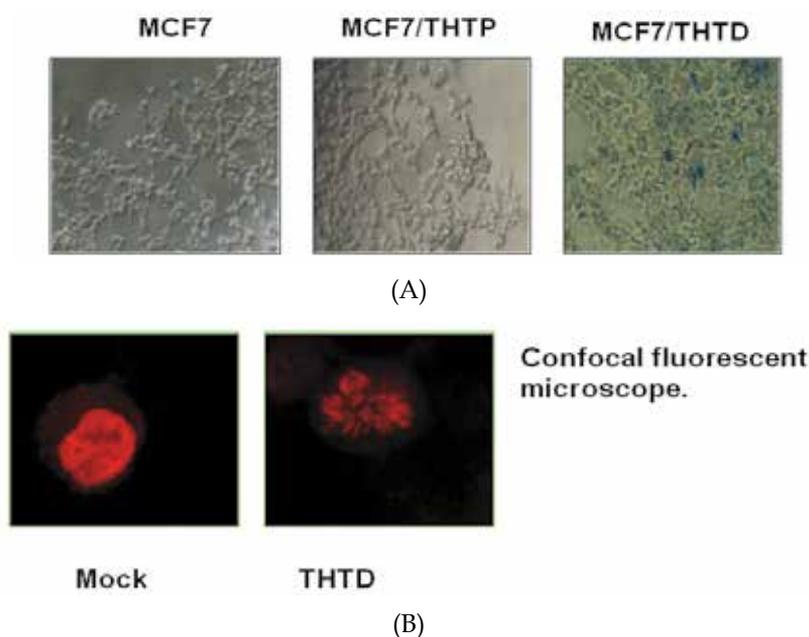


Fig. 3. Knockdown Cdc6 renders senescence human breast carcinoma MCF7 cells. (A). Senescence-associated  $\beta$ -galactosidase (SA- $\beta$ -Gal) staining: MCF7 cells ( $2.0 \times 10^4$ /well) were grown in a six-well plate, infected with THTD, THTP, and mock infected at an m.o.i. of 1.0. Two days postinfection, the cell was washed in 2 ml 1x PBS, fixed in 2 ml of 3% formaldehyde, washed, and incubated for 24 h at 37°C with a solution containing 1 mg/ml of 5-bromo-4-chloro-3-indolyl B-D-galactoside (X-Gal), 40 mM citric acid/sodium phosphate at pH 6.0, 5 mM potassium ferrocyanide, 5 mM potassium ferricyanide, 150 mM NaCl, and 2 mM  $MgCl_2$ . After washing, the cell staining was viewed with a fluorescent microscope, and photographed. (B). Anti-H3K9me2 indirect immunofluorescent staining. MCF7 cells were infected with THTD at m.o.i. of 1.0. Two days postinfection, the cell was fixed and stained with the anti-H3K9me2 antibody. The slides were then stained with a goat anti rabbit Ig G conjugated with rhodamine red, and then with DAPI. The fluorescent-stained cells were visualized with a fluorescent microscope and its image processing system. The RGB color model of the cell images was applied by using Adobe Photoshop program.

distinct chromatin structure known as SAHF with the accumulation of H3K9me3/me2 would be observed in the nuclei of senescent cells (Narita et al., 2003, 2006). To demonstrate this cellular event, the Cdc6 knockdown cells along with the mock infected MCF7 cells were stained with an antibody against methylated lysine 9 on histone H3 (H3K9me2). Indirect-immunofluorescent staining revealed an overall elevated positive staining for H3K9me2 in the nuclei of the Cdc6 knockdown MCF7 cells in comparison of those mock infected cells. By conducting confocal fluorescent microscope, a grainy-structured nucleus with the accumulated H3K9me2 was evident, a characteristic change for the complete proliferation arrest. This change was in contrasting to the mock infected cells in which a uniformly distributed H3K9me2 staining was observed in the nucleus of mock-infected cell (Figure 3B.).

## **2.2 Human Cdc6 stimulates D-type cyclins-Cdk4 kinases by suppressing p16<sup>INK4a</sup> Cdk inhibitory activity**

In MCF7 cells, Cdk4 kinase activity is high as hyper Rb-C phosphorylation is observed, even if elevated levels of p16<sup>INK4a</sup> is presented. One reason is that p16<sup>INK4a</sup> CDK inhibitory activity has been repressed. We have found that this repression is relieved when Cdc6 is knockdown.

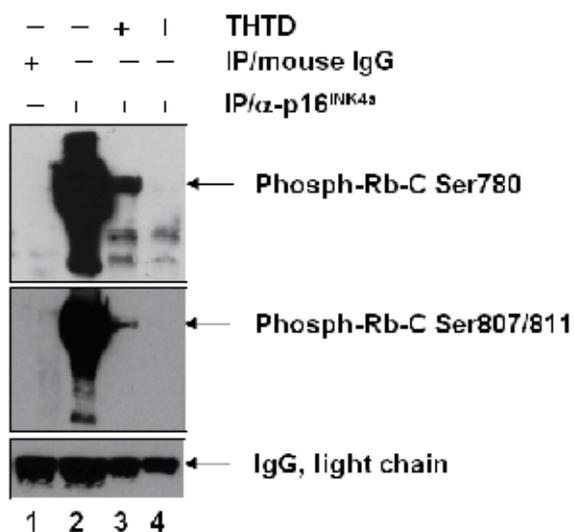
### **2.2.1 Cdc6 is involved in Rb-C phosphorylation**

The cell cycle progression in MCF7 cells was not inhibited even in the presence of p16<sup>INK4a</sup>. We reasoned that Cdc6 might play a role in inactivating p16<sup>INK4a</sup>. To test this speculation, a nonradioactive immunoprecipitation (IP)-kinase assay was carried out. Because p16<sup>INK4a</sup> associates with D-type cyclins-Cdk4/Cdk6 complexes, and because phosphorylation of Rb C-terminal domain is catalyzed by D-type cyclins-Cdk4/Cdk6 kinases (Harbour et al., 1999), the Rb phosphorylation was determined using anti-Phospho Rb S780 and anti-Phospho Rb S807/811 specific antibodies. p16<sup>INK4a</sup> IP from MCF7 cells phosphorylated the Rb-C in vitro (Figure. 4., lane 2), while IP with mouse IgG could not (Figure. 4., lane 1). When Cdc6 was depleted via *CDC6* shRNA lentivirus (THTD) infection, the Rb-C phosphorylation was inhibited (Figure. 4., lane 3), and no kinase activity was detected with transgene p53 expression in the THTD-infected cells (lane 4). We determined the Cdc6 inhibitory effect on p16<sup>INK4a</sup> activity using p16<sup>INK4a</sup> IP dilutions in the kinase reactions. We found that p16<sup>INK4a</sup> CDK inhibitory activities were at least suppressed 25 times in the presence of Cdc6. In other words, at least 25 times suppression of Rb phosphorylation was archived upon Cdc6 knockdown.

### **2.2.2 Cdc6 is found in Cdk4 complex and is required for Rb-C phosphorylation**

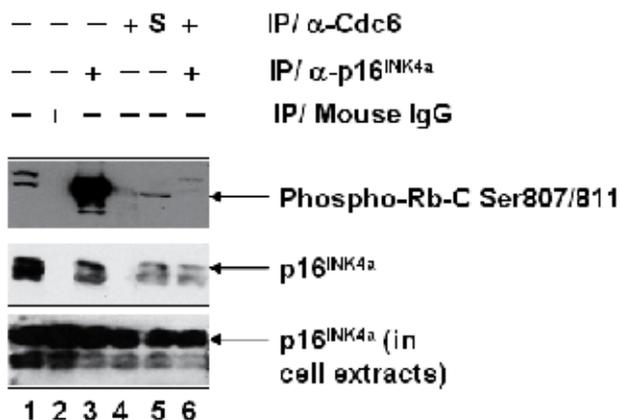
The result that p16<sup>INK4a</sup> IP from the Cdc6 depleted MCF7 cells showed very low levels of Rb-C phosphorylation (Figure 4.) indicates that Cdc6 might be required for Rb phosphorylation. To test this prediction, Cdc6 was first depleted from whole cell extracts prepared from MCF7 cells in vitro by using anti-Cdc6 antibody (Immunodepletion). The Cdc6 immunodepleted extracts were then subjected to kinase reactions.

Crude extracts turned Rb-C phosphorylated at very low level, while Cdc6 IP could not confer the phosphorylation, nor did the Cdc6 immunodepletion (Figure 5., lanes 4 and 5). In contrast to the direct p16<sup>INK4a</sup> immunoprecipitation, the Cdc6 depletion followed by p16<sup>INK4a</sup> IP showed no Rb-C phosphorylation (Figure 5., lane 6). Cdc6, therefore, was required for the Rb-C phosphorylation.



MCF7 cells were infected with THTD for Cdc6 knockdown, or mock infected at m.o.i. of 1.0. Two days postinfection, the cell extracts were prepared. The whole cell extracts (~1000  $\mu$ g of protein) were mixed with the anti-p16<sup>INK4a</sup> monoclonal antibody. Immunoprecipitation (IP) proceeded at 4°C overnight. In vitro kinase reactions were set up with one quarter of IP beads mixed with Rb-C substrate and ATP, and incubated at 37°C for 30 minutes according to the manufactory's instruction. The specific Rb-C phosphorylations were determined in Western blot with the anti-Phospho Ser780 antibody (top panel), or anti-Phospho Ser807/811 antibody (middle panel). The mouse IgG signals were presented as IP controls (bottom panel).

Fig. 4. Knockdown Cdc6 causes sharp decline of Rb-C phosphorylation.



MCF7 cell extracts were mixed with the anti-Cdc6 monoclonal antibody, and the IP proceeded at 4°C overnight. In Lane 1 are cell extracts only. Top panel: samples for Rb-C phosphorylation were assayed (S: the supernatants of Cdc6 IP). Middle panel: immunoblotting (IB) of p16<sup>INK4a</sup> with different IPs, or immunodepletion (S). Bottom panel: p16<sup>INK4a</sup> immunoblotting with extracts from each sample.

Fig. 5. Cdc6 is required for Rb-C phosphorylation.

It was interesting to note the presence of large amounts of Cdk4, but no p16<sup>INK4a</sup> in Cdc6 IP. Cdc6, therefore, forms complex with Cdk4. In the Cdc6 knockdown MCF7 cells, however, immunoblotting showed a sharp decline of Cdk4 contents in the Cdc6 IP in which p16<sup>INK4a</sup> could be detected.

### 2.2.3 Cdc6 actively represses p16<sup>INK4a</sup> CDK inhibitory functions

In our experimental system, Cdk-4 IP could not confer the Rb-C phosphorylation (Figure 6., lane 7) due to Cdk-4 antibody used had neutralization effect on Cdk-4 kinase activity. In order to study the effect of Cdc6 on stimulating D-type-cyclins-Cdk4 kinase activity, a hemagglutinin (HA)-tagged human Cdk4 (Cdk4-HA) was expressed in MCF7 cells. Ectopic Cdk4 expression induced high levels of p16<sup>INK4a</sup> whether Cdc6 was presented or depleted. The Cdk4-HA was immunoprecipitated with anti-HA-agarose. The presence of Cdk4-HA was confirmed in Western blot. In contrast to the Cdk4 IP prepared from the parental MCF7 cells, the anti-HA-agarose IP contained p16<sup>INK4a</sup>, indicating that Cdk4-HA was bound with the CDK inhibitor p16<sup>INK4a</sup> even in the presence of endogenous Cdc6. With p16<sup>INK4a</sup> associated, the Cdk4-HA IP, however, rendered phosphorylation of Rb-C, though the kinase activity was lower in comparison of that from the p16<sup>INK4a</sup> IP (Figure. 6., lane 2, 3 and 4).

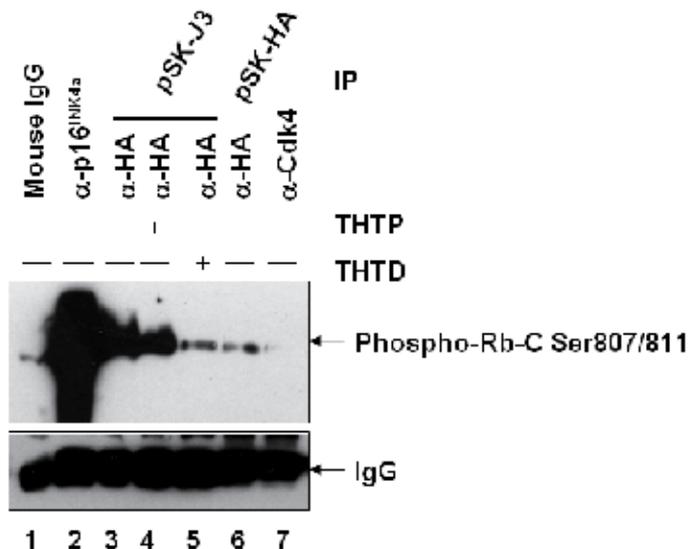


Fig. 6. Cdc6 suppresses p16<sup>INK4a</sup> CDK inhibitory activities even p16<sup>INK4a</sup> is bound with Cdk4.

MCF7 cells were infected with THTD or THTP. Two days postinfection, the cells were then transfected with pSK-J3 (for Cdk4-HA) or pSK-HA (HA-tag control). Cell extracts were prepared, and were subjected to IP with different antibodies. Top panel: each IP sample was added into the kinase reactions, and the in vitro Rb-C phosphorylation was performed. The phosphorylated Rb-C was examined in Western blot with the anti-Phospho Ser807/811 antibody. Bottom panel: mouse IgG was presented as the loading control.

The ectopic Cdk4-HA could contribute to the Rb-C phosphorylation without Cdc6 stimulation given the p16<sup>INK4a</sup> CDK inhibitory activity is bypassed upon Cdk4 overexpression. To test this possibility, Cdc6 was first knockdown with THTD lentivirus

producing *CDC6* shRNA in MCF7 cells. The cell was then transfected with the Cdk4-HA expressing plasmid pSK-J3. As *Cdc6* was knockdown, Rb-C phosphorylation by Cdk4-HA kinases precipitated with anti-HA-agarose, if any, was to a basal level (Figure. 6., lane 5), indicating *Cdc6* is required for stimulating the D-type cyclins-Cdk4 kinases. These results indicated that the effect of *Cdc6* was dominated over p16<sup>INK4a</sup> even if Cdk4 was associated with p16<sup>INK4a</sup>, and the CDK inhibitory activity had been suppressed, rendering the Rb hyperphosphorylation by D-type cyclins-Cdk4 kinases.

### 2.3 The significance of inhibiting p16<sup>INK4a</sup>-Rb pathway by *Cdc6*

Our studies address the mechanisms of cell proliferation and senescence. The approach that we have adopted and the results we have accomplished are novel in that we don't try to artificially restore the function of p16<sup>INK4a</sup>-Rb pathway by overexpressing either p16<sup>INK4a</sup>, or pRb protein. Rather, we biologically remove a specific blocker, *Cdc6*, from the cancer cell. We have found that *Cdc6* forms complex with Cdk4, and this association activates D-type cyclins-Cdk4 kinases in Rb-C phosphorylation. Moreover, *Cdc6* stimulates Rb-C phosphorylation by D-type cyclins-Cdk4 kinases even in the presence of high levels of p16<sup>INK4a</sup>. Further, *Cdc6* is required for Cdk4 kinase activity because ectopic Cdk4 expression can not augment phosphorylation of Rb-C when *Cdc6* is depleted (Figure 6.). The effect of *Cdc6* and Cdk4 interaction is overwhelm that a dramatic decline of association between p16<sup>INK4a</sup> and Cdk4 takes place, indicating that *Cdc6* could block or interfere p16<sup>INK4a</sup> associating with Cdk4. *Cdc6* has been shown in the interaction with cyclin B, which is involved in S phase entry (Dutta and Bell, 1997; Méndez and Stillman, 2000), or with cyclin E-Cdk2 complex (Mailand and Diffley, 2005), which controls the initiation of DNA replication. In all of these cases *Cdc6* is the target of CDK kinase reactions. Different from these interactions, the association with Cdk4 enables *Cdc6* actively suppressing p16<sup>INK4a</sup> CDK inhibitory activity, which is necessary for cell proliferation.

The most striking feature of oncogene-induced senescence is through p53 pathway (Hemann and Narita, 2007; Yaswen and Campisi, 2007). However, there is always exception, and p16<sup>INK4a</sup>-Rb network exists as secondary senescence pathway (Sedivy, 2007). p16<sup>INK4a</sup> is responsible for telomerase-independent senescence induced in various cellular stresses. p16<sup>INK4a</sup> inhibits D-type cyclins-Cdk4/Cdk6, which activates Rb pathway, and functional p16<sup>INK4a</sup>-Rb pathway is involved in cellular senescence (Beauséjour et al., 2003; Narita et al., 2003, 2006). The key issue is that the p16<sup>INK4a</sup>-Rb pathway has been inactivated in most human cancers. *Cdc6* has been shown to suppress p16<sup>INK4a</sup> expression via chromatin epigenetic modifications (Gonzalez et al., 2006). Our study demonstrates that human *Cdc6* suppresses p16<sup>INK4a</sup> CDK inhibitory activity, and p16<sup>INK4a</sup>-Rb pathway is inactivated in MCF7 cells and perhaps in many different kinds of human cancers. Elevated levels of *Cdc6* expression have been evident in many human cancers. *Cdc6* overexpression induces senescence of several human cancer cell lines with double-stranded DNA breaks (Bartcova et al., 2006). The *Cdc6*-induced senescence human cancer cells are associated with the ATM-Chk2 DNA damage checkpoint control, and the depletion of ATM kinase, or its substrate p53, suppresses the senescence (Bartcova et al., 2006). Unlike ATM-Chk2/p53 pathway, which can be spontaneously activated in response to replicative stresses (Bartkova et al., 2005, 2006), and which is potential for the emergence of p53-resistant tumors, the reactivation of p16<sup>INK4a</sup>-Rb pathway needs *Cdc6* depletion accompanied with the formation of SAHF inside the cancer cells.

Functional p16<sup>INK4a</sup>-Rb pathway is essential for the SAHF formation (Narita et al., 2003). The production of H3K9me3/me2 catalyzed by histone methyltransferase SUV39H1 is an important step of heterochromatinization, and is regulated by p16<sup>INK4a</sup>-Rb pathway. Recently, a non-histone protein, high mobility group A (HMGA), has been identified as an essential component of SAHF. Moreover, HMGA proteins promote complete proliferation arrest, and p16<sup>INK4a</sup> has been shown to coordinate with HMGA to induce senescence normal fibroblasts (Narita et al., 2006). HMGA protein functions by interacting with the minor groove of AT-rich DNA sequences. The association of HMGA protein with chromatin can be displaced by DNA-binding drug distamycin A, which binds the minor groove of AT-rich DNA (Narita et al., 2006).

In late M phase, accumulated human Cdc6 associates with chromatin for pre-RC assembly (Dutta and Bell, 1997; Méndez and Stillman, 2000). Cdc6 could block Rb in preventing heterochromatinization since Cdc6 knockdown induces hypophosphorylated Rb, which is the activated form of tumor suppressor, and which promotes the formation of SAHF (Narita et al., 2006) (Figure 3.). We have shown that budding yeast Cdc6 interacts with DNA non-specifically with preference to A/T-rich tracks. It is likely that human Cdc6 interacts with DNA directly since this protein shares identical amino acid residuals with the budding yeast *Saccharomyces cerevisiae* Cdc6 within highly conserved domains. The positive results could support one of this protein's important biological functions in promoting cell proliferation.

Future work will focus on demonstrating the competition of chromatin association between Cdc6 and HMGA proteins. By interacting with the minor groove of the AT-rich DNA sequences, Cdc6 would compete with HMGA proteins in associating with chromatin. It is essential that the association of HMGA proteins with DNA for the complete proliferation arrest, that is the formation of SAHF (Narita et al., 2006). In the parental MCF7 cells, Cdc6 is expected to prevent the heterochromatin formation by blocking and displacing HMGA proteins in binding to DNA. In contrast, in the Cdc6 depleted cells, no chromatin-bound Cdc6 will be expected. Moreover, the Cdc6 depletion reactivates p16<sup>INK4a</sup>-Rb pathway, this will stimulate the HMGA proteins binding to DNA constitutively and promotes SAHF formation.

### 3. Conclusion

We would draw some conclusions from these studies: Human Cdc6 is required for phosphorylation of cell-cycle controller and tumor suppressor Rb protein, which is of important for cell proliferation. Cdc6 functions to associate with D-type cyclins-Cdk4 and suppresses p16<sup>INK4a</sup>-Rb pathway, which otherwise controls cell cycle progression and prevents cell from unlimited proliferation. In many human cancers, p16<sup>INK4a</sup>-Rb pathway has been inactivated. Cdc6 plays a role in inactivating biological functions of p16<sup>INK4a</sup>-Rb pathway to promote cancer cell proliferation. Cdc6 depletion via gene transduction vector producing shRNA renders reactivation of p16<sup>INK4a</sup>-Rb pathway and leads the cancer cells to replicative senescence.

### 4. Acknowledgement

This work is partially supported by the T.J. Martell Foundation (AJ) and by Green Life Star, Inc., and by JAC RNTEin Biotech. Institute, Tokyo, Japan (FL). We would thank Dr. S. Kyo of Kanazawa University of Japan for providing *hTERT* promoter plasmid, Dr. S. A. Chow for

providing HIV-1 plasmid and 293T cells, and the Beckman Institute at the City of Hope for providing us with MCF7 cell line. We would also like to thank Ms. Chun-Hua Wu's excellent technical assistance.

## 5. References

- Bartkova, J.; Horejsi, Z.; Koed, K.; Kramer, A.; Tort, F.; Zieger, K.; Guldberg, P.; Sehested, M.; Nesland, J.M.; Lukas, C.; Orntoft, T.; Lukas, J. & Bartek, J. (2005). DNA Damage Response as a Candidate Anti-Cancer Barrier in Early Human Tumorigenesis. *Nature*, Vol.434, pp. 864-870
- Bartkova, J. et al. (2006). Oncogene-Induced Senescence is Part of the Tumorigenesis Barrier Imposed by DNA Damage Checkpoints. *Nature*, Vol.444, pp. 633-637
- Beauséjour, C.M.; Krtolica, A.; Galimi, F.; Narita, M.; Lowe, S.W.; Yaswen, P. & Campisi, J. (2003). Reversal of Human Cellular Senescence: Roles of the p53 and p16 Pathways. *EMBO J.*, Vol.22, No.16, (August), pp. 4212-4222
- Campisi, J. (2001). Cellular Senescence as a Tumor-Suppressor Mechanism. *Trends Cell Biol.*, Vol.11, Suppl. pp. 27-31
- Clay-Farrace, L.; Pelizon, C.; Pines, J. & Laskey, R.A. (2003). Human Replication Protein Cdc6 Prevents Mitosis through a Checkpoint Mechanism that Implicates Chk1. *EMBO J.*, Vol.22, pp. 704-712
- Duursma, A. & Agami, R. (2005). p53-Dependent Regulation of Cdc6 Protein Stability Controls Cellular Proliferation. *Mol. Cell. Biol.*, Vol.25, pp. 6937-6947
- Dutta, L. & Bell, S.P. (1997). Initiation of DNA Replication in Eukaryotic Cells. *Annu. Rev. Cell Dev. Biol.*, Vol.13, pp. 293-332
- Gonzalez, S.; Klatt, P.; Delgado, S.; Conde, E.; Lopez-Rios, F.; Sanchez-Cespedes, M.; Mendez, J.; Antequera, F. & Serrano, M. (2006). Oncogenic Activity of Cdc6 through Repression of the *INK4/ARF* Locus. *Nature*, Vol.440, pp. 702-706
- Gorgoulis, V.G.; Vassiliou, L.V.; Karakaidos, P.; Zacharatos, P.; Kotsinas, A.; Liloglou, T.; Venere, M.; Dittullo, R.A. Jr.; Kastrinakis, N.G.; Levy, B.; Kletsas, D.; Yoneta, A.; Herlyn, M.; Kittas, C. & Halazonetis, T.D. (2005). Activation of the DNA Damage Checkpoint and Genomic Instability in Human Precancerous Lesions. *Nature*, Vol.434, pp. 907-913
- Harbour, J.W.; Luo, R.X.; Dei Santi, A.; Postigo, A.A. & Dean, D.C. (1999). Cdk Phosphorylation Triggers Sequential Intramolecular Interactions that Progressively Block Rb Functions as Cells Move through G1. *Cell*, Vol.98, pp. 859-869
- Hemann, M.T. & Narita, M. (2007). Oncogenes and Senescence: Breaking down in the Fast Lane. *Genes & Dev.*, Vol.21, pp. 1-5
- Lau, E.; Zhu, C.J.; Abraham, R.T. & Jiang, W. (2006). The Functional Role of Cdc6 in S-G2/M in Mammalian Cells. *EMBO Rep.*, Vpl.4, pp. 425-430
- Mailand, N. & Diffley, J.F.X. (2005). CDKs Promote DNA Replication Origin Licensing in Human Cells by Protecting Cdc6 from APC/C-Dependent Proteolysis. *Cell*, Vol.122, pp. 915-926
- Mendez, J. & Stillman, B. (2000). Chromatin Association of Human Origin Recognition Complex, Cdc6, and Minichromosome Maintenance Proteins during the Cell Cycle: Assembly of Prereplication Complexes in Late Mitosis. *Mol. Cell. Biol.*, Vol.20, pp. 8602-8612

- Narita, M.; Nuñez, S.; Heard, E.; Narita, M.; Lin, A.W.; Hearn, S.A.; Spector, D.L.; Hannon, G.J. & Lowe, S.W. (2003). Rb-Mediated Heterochromatin Formation and Silencing of E2F Target Genes during Cellular Senescence. *Cell*, Vol.113, pp. 703-716
- Narita, M.; Narita, M.; Krizhanovsky, V.; Nuñez, S.; Chicas, A.; Hearn, S.A.; Myers, M.P. & Lowe, S.W. (2006). A Novel Role for High-Mobility Group A Proteins in Cellular Senescence and Heterochromatin Formation. *Cell*, Vpl.126, pp. 503-514
- Sedivy, J.M. (2007). Telomeres Limit Cancer Growth by Inducing Senescence: Long-Sought in vivo Evidence Obtained. *Cancer Cell*, Vol.11, pp. 389-391
- Yaswen, P. & Campisi, J. (2007). Oncogene-Induced Senescence Pathways Weave an Intricate Tapestry. *Cell*, Vol. 128, pp. 233-234

# Visualize Dynamics of Chromosome Structure Formation and DNA Repair/Recombination Coupled With DNA Replication: Tight Coupled Role of DNA Replication in Chromosome Compaction and DNA Recombination

Eisuke Gotoh<sup>1,2</sup>

<sup>1</sup>*Division of Genetic Resources, National Institute of Infectious Diseases Japan, Tokyo*

<sup>2</sup>*Department of Radiology, Jikei University School of Medicine, Tokyo Japan*

## 1. Introduction

The genetic information of living organisms is carried and stored in a gene, which is a huge complex molecule involving DNA and proteins. In eukaryote cell, gene is usually higher hierarchical ordered packaged in a large complex structure called "chromosome". The genetic information must be strictly conserved throughout life cycle, otherwise lost or false DNA information may cause malfunction of cell, cell death, impaired development or cancer prone. The genetic information must also be principally correctly transmitted to the next generations for stabilize the species. Three major cellular functions ensure maintain the complete gene information and transmit to daughter cells or next generations: (1) DNA replication; duplicates complete DNA sequence information to transmit for next cell generation, (2) DNA repair; recover structural damages of DNA and corrects false information anything else introduced in DNA molecule by intra- or extra-cellular factors and (3) Chromosome condensation; compaction of DNA molecule to architect a higher-organized packaged and compacted architecture known as "chromosomes" to assure protect and stabilize the integrity of labile DNA molecules and segregate the identical chromosome sets to daughter cells during mitosis. In addition, DNA recombination/rearrangement mechanism also plays very crucial role for various cellular function. For example, exchange genetic information between paternal and maternal homologous chromosome gene during recombination process in fertilization will promote evolution of species and prevent the gene homogeneity. Recombination of Immuno-globulin gene during lymphocyte maturation process promises the antibody diversity against vast numbers of foreign antigens. Recently accumulated evidences have strongly suggested that these mechanisms are mutually tight-coupled functions rather than they independently work. Molecular genetic studies have also provided supporting evidence for the idea that mutation in genes (HIRA/Tuple1, XCDT1, cdt1, Orc2, Orc3, Orc5, MCM2, MCM4, MCM10, RECQL4)(Maiorano et al. 2000; Nishitani et al. 2000; Pflumm and Botchan 2001; Christensen and Tye 2003; D'Antoni et al. 2004; Prasanth et al. 2004; Tatsumi et al. 2006) required for

DNA replication or DNA repair (SCC2, SMC1, SMC3, ESCO2) (Gillespie and Hirano 2004; Bekker-Jensen et al. 2006; Gandhi et al. 2006; Yoshizawa-Sugata and Masai 2009) showed abnormal phenotype in chromosome structure, inherited diseases, genomic instability or prone to cancer, or aberrant replication timing causes abnormal chromosome condensation. Cytogenetical observation of chromosome is certainly a most direct approach for elucidates how dynamics of chromosome structure formation and chromosome recombination are coupled with DNA replication. Numbers of supporting evidences have reported the relationship of DNA replication, chromosome conformation/recombination and DNA repair mechanisms. However, the results shown in these reports were mainly yielded from the observation in interphase nuclei, thus the resolution and the resulting information were still limited because chromosomes are only visible in the mitotic stage of the cell-cycle and are invisible as de-condensed in interphase. Premature chromosome condensation (PCC) technique is a unique and useful technique that allows the interphase nuclei to be visualized like as condensed form of mitotic chromosome. Conventional PCC has been carried out by cell fusion using either fusogenic viruses or polyethylene glycol (cell fusion-mediated PCC, (Johnson and Rao 1970; Pantelias and Maillie 1983). But cell fusion mediated PCC protocols are usually technically demanding and keenly depend on activity of the virus or PEG. More recently, a much easier and more rapid PCC technique using calyculin A or okadaic acid, specific inhibitors of protein phosphatases (drug-induced PCC, Gotoh et al. 1995; Gotoh and Asakawa 1996). Drug-induced PCC is now becoming much more popular and has been used in a wide range of cytogenetic studies (Durante et al. 1998; Gotoh et al. 1999; Ito et al. 2002; El Achkar et al. 2005; Gotoh and Tanno 2005; Srebniak et al. 2005; Deckbar et al. 2007; Gotoh 2007).

In this chapter, we first show the visualizing of the dynamics of chromosome structure formation and chromosome repair/recombination coupled with DNA replication in interphase nuclei, in particular by use of drug-induced PCC. We next show the spatial and temporal rearrangement of chromosome fragment after  $\gamma$ -ray irradiation, and this movement is tightly coupled with DNA replication. Possible hypothetical model for chromosome condensation/compaction involving the role of DNA replication and model for chromosome recombination coupled with DNA-replication will be suggested and discussed.

## **2. Visualizing of the dynamics of chromosome structure formation coupled with DNA replication**

### **2.1 Background**

Chromosomes condense during mitosis under a higher ordered stringent mechanism in mitotic phase, which is further divided into several sub-phases; (preprophase), prophase, prometaphase, metaphase, anaphase, telophase (these phases are called as mitosis) and followed by cytokinesis (Alberts et al. 1989). In the course of mitotic phase, number of drastic sequential transactions proceed; (1) chromatin condense to visible structure under a microscope "chromosomes" and mitotic spindle begins form (2) nuclear envelope breakdown into membrane vesicles, centriole and mitotic spindle formation followed by spindle attaches to centromere (3) kinetochore microtubules align the chromosomes at metaphase plate (4) chromosome separates and segregates to spindle poles (5) separated daughter chromatids reached at the poles, nuclear envelope re-forms (6) formation of contractile ring and cleavage furrows which constrict the cell center, cytokinesis and cell divides two daughter cells and chromosome de-condensation in divided cells and re-enter the cells in G1 phase (Alberts et al. 1989). The detailed whole mechanism is still almost

unclear, however number of molecules which involved in the mitotic events have been identified such as SMC proteins include condensin (chromosome condensation) and cohesin (chromosome cohesion of replicated chromosomes) (Swedlow and Hirano 2003), NuMA protein for spindle pole formation (Chang et al. 2009; Haren et al. 2009; Silk et al. 2009; Torres et al. 2010), nuclear lamins (Moir et al. 2000), aurora kinases in centromere function (Tanno et al. 2006; Meyer et al. 2010; Tanno et al. 2010), shugoshin and protein phosphatase 2A in chromosome cohesion (Kitajima et al. 2006; Tanno et al. 2010), cdk1 in chromosome condensation, chromosome bi-orientation (Tsukahara et al. 2010), cyclin B, cdc2, cdc25 in chromosome condensation (Masui 1974; Draetta and Beach 1988; Dunphy et al. 1988; Kumagai and Dunphy 1992), Polo and Rho in cytokinesis (Burkard et al. 2009; Wolfe et al. 2009; Li et al. 2010), and many other molecules. Dynamics in mitotic phase involves such as various elements, and numerous studies for visualize these dynamics have been reported, because these events are relatively easy observable in mitosis under microscopes. However, visualizing approaches in chromosome dynamics coupled with DNA replication is still limited, because it should be required to observe chromosomes in S-phase but chromosomes are usually invisible in S-phase as they de-condense (Gotoh et al. 1995; Gotoh and Durante 2006).

In this section, we just focus on visualize the chromosome dynamics coupled with DNA replication during S-phase progression use with drug-induced premature chromosome condensation method and we show how replicating DNA is folded to higher order chromosomes.

## **2.2 Materials and methods**

### **2.2.1 Cell, chemicals and antibodies**

Human normal karyotype fibroblast cell line GM05389 was from Coriell Cell Repositories (NJ, USA). Calyculin A was from Wako Chemicals (Osaka, Japan), dissolved in 100% DMSO, 100  $\mu$ M of stock solution was stored at  $-20^{\circ}\text{C}$ . Cy3-dUTP was from Amersham (Uppsala, Sweden). Sizes of 425-600  $\mu\text{m}$  diameter acid-washed glass beads were from SIGMA (MO, USA).

### **2.2.2 Cy3-dUTP labeling by bead-loading method**

Cy3-dUTP was loaded in the cells using beads loading procedure as described previously (McNeil and Warder 1987; Manders et al. 1999; Ito et al. 2002; Gotoh 2007). The procedure facilitate the Cy3-dUTP to be incorporated in the cell nucleus in very short time whereby transiently permeabilizes the cell membranes (McNeil and Warder 1987; Manders et al. 1999). GM05389 cells, seeded on a 35 mm glass-base culture dish (Iwaki, Japan) which is designed for observation under inverted microscopes, were grown exponentially and asynchronous in MEM supplemented with 15% foetal calf serum at  $37^{\circ}\text{C}$  in 5%  $\text{CO}_2$  atmosphere with 95% humidification. For replication DNA labeling, culture medium was replaced with  $37^{\circ}\text{C}$  prewarmed 50  $\mu\text{l}$  of 10  $\mu\text{M}$  Cy3-dUTP dissolved in MEM medium, 425-600  $\mu\text{m}$  diameter glass beads were sprinkled onto the cells, tapping and rocking the dish several times, then the beads were rinsed off with PBS. It usually took a couple of minutes for complete the loading procedure. Pre-warmed ( $37^{\circ}\text{C}$ ) MEM was added to the culture and incubated until 10 minutes after starting of Cy3-dUTP loading. Then cells were subjected to be condensed prematurely using calyculin A, or to be immunostained.

### **2.2.3 Premature chromosome condensation**

Immediately after Cy3-dUTP loading procedure, chromosomes were condensed prematurely using 50 nM of calyculin A. Premature chromosome condensation (PCC) using

calyculin A was done as described elsewhere (Gotoh et al. 1995; Asakawa and Gotoh 1997; Johnson et al. 1999; Ito et al. 2002; Gotoh 2007; Gotoh 2009) except that the incubation time was shortened as possible to 10 minutes to obtain a high spatial and temporal resolution, otherwise replication foci can be merged together resulting in a less spatial/temporal resolution. After incubation with calyculin A, cells were harvested, swollen in 0.075M KCl for 10 minutes at 37°C, fixed 3 times with carnoy's fixative (methanol: acetic acid = 3 vol.: 1 vol.), dropped on a glass slide and air-dried. Finally, DNA was counterstained with 200 ng/ml of DAPI, mounted with PBS and covered with a cover slip.

#### **2.2.4 Laser scanning microscopy**

Blue, green, or red confocal images were collected using a Confocal Laser Scanning microscopy system Zeiss LSM510 (Jena, Germany) equipped on an inverted epifluorescence microscope Zeiss Axiovert100M (Jena, Germany). A water immersion objective lens Zeiss C-Apochromat 63x/1.2 w corr. (Jena, Germany) was used. UV-, Ar- or He-Ne-laser was tuned at 366, 488, 568 nm to excite DAPI, FITC, Cy-3, respectively. The fluorescence signals from each fluorochrome were recorded separately in multiple scan to minimize optical cross talk. Digital images were manipulated using Zeiss LSM510 software (Jena, Germany).

#### **2.3 Dynamics of chromosome structure formation coupled with DNA replication**

Much study has achieved to visualize the dynamics of chromosome condensation during in interphase nuclei, particularly in S-phase. However, the visualizing study on the relationship between chromosome condensation and DNA replication has still limited. Several studies tried and defined fairly well the replication foci distribution in interphase nuclei (Nakamura et al. 1986), however little is still known about how replicating DNA is folded to higher order chromosomes, because chromosomes are invisible in interphase as they de-condense.

To visualize the chromosome compaction dynamics coupled with DNA replication more precisely in S-phase nucleus, we took the advantage of the drug-induced PCC method (Gotoh et al. 1995; Asakawa and Gotoh 1997; Johnson et al. 1999; Ito et al. 2002). Individual substage of S-phase can be easily identified by typical diagnostic appearances seen in different phases of S-PCCs (Mullinger and Johnson 1983; Gollin et al. 1984; Hameister and Sperling 1984; Savage et al. 1984; Gotoh et al. 1995; Gotoh and Durante 2006). A drastic conformational change of chromosome structure formation along with the proceed of DNA replication, as shown in Figure 1 (reproduced from *Chromosoma*. 2007; 116(5): 453-462, Gotoh 2007), clearly revealed in PCCs following Cy3-dUTP labeling (Gotoh 2007; Gotoh 2011). Cy3-dUTP loading procedure was finished within 10 minutes followed by 10 minutes of PCC induction and fixation (for the detail of the experiment procedures, see the Materials and Methods section in *Chromosoma*. 2007; 116(5): 453-462, Gotoh 2007). Accordingly, only DNA replicated in this short lapse fluoresces. Thus, the observed S-PCCs in the present study reflected the replication stages at most 20 minutes (see Materials and Methods described in Gotoh 2007) before the cell fixation. (i) In early S-phase, PCCs showed a cloudy spreading mass of thin fibers like a "nebula". Numerous fine granular foci distributed homogeneously on overall the fibers (Figure 1i), which shows "beads on a string" appearance that are similar to the same named structures observed under an electron microscope (Olins and Olins 1974; Thoma et al. 1979). (ii) In middle S-phase, typical "pulverized" PCCs structure was recognized, and the size of foci increased with the number of foci decreased and distributed unevenly on chromosomes. The foci became much brighter (Figure 1j). (iii) In late S-phase, chromosomes showed mostly

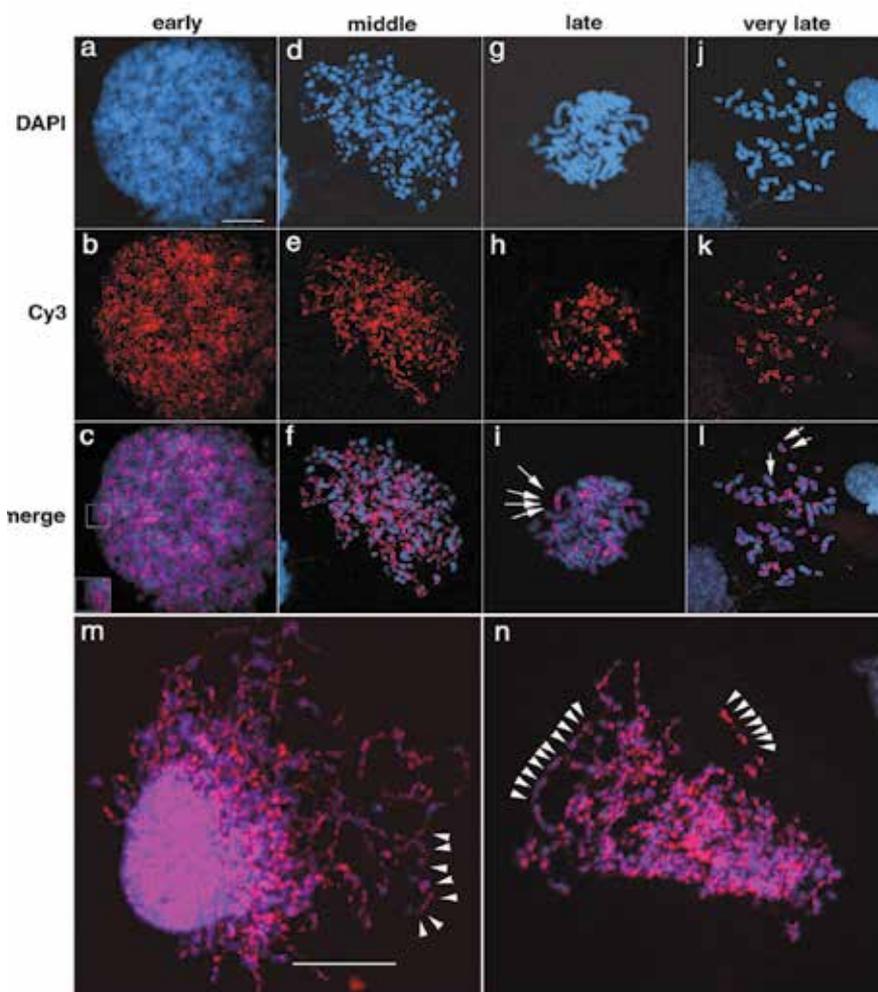


Fig. 1. (1) DNA replication regions on prematurely condensed chromosomes (PCCs) of different substages of S-phase. Ten minutes after Cy3-dUTP loading, cells were condensed prematurely using 50 nM of calyculin A (Gotoh et al. 1995). From left to right column, (a,b,c) early S-phase PCCs, (d,e,f) middle S-PCCs, (g,h,i) late S-PCCs and (j,k,l) very late S-PCCs. (a,d,g,j) DAPI counterstained DNA, (b,e,h,k) Cy3-dUTP labelled DNA replication region and (c,f,i,l) Merged image of DAPI and Cy3. Centromeric region (arrow) or telomeric region (arrowhead) replicates in very-late S-phase are indicated. (i,l) Late S- and very late S-PCCs already condensed like as mitotic chromosomes, but these PCCs were actually S-phase chromosomes because they incorporated Cy3-dUTP. G2/M chromosomes are easily distinguished from late or very late S chromosomes as G2/M chromosomes do not incorporate Cy3-dUTP (data not shown). Inset in (c) is higher magnification of the boxed portion. Scale bar, 10  $\mu$ m. (2) DNA replication regions seen on prominent fibre of PCCs. (m) early-S-phase (918 foci scored) and (n) middle S-phase (707 foci scored). Replication foci are clearly seen as 'beads on a string' structure, some of these are indicated by arrowhead. Scale bar, 10  $\mu$ m. (reproduced from *Chromosoma*. 2007; 116(5): 453-462, Gotoh 2007)

condensed like mitotic chromosomes. Cy3-dUTP incorporated regions were recognized as band arrays inserted in the condensed chromosome (Figure 1k, indicated by arrows). The similar appearance of replication foci aligned longitudinally on chromosomes were previously reported on metaphase of kangaroo-rat kidney PtK1 cells (Ma et al. 1998). The size of foci increased and the number decreased. (iv) In very late S-phase, the number of foci further reduced and predominantly localized at centromeric or telomeric regions (Figure 1l, indicated by arrows). These regions are actually known as satellite heterochromatic DNA regions where DNA replicates at very late S (O'Keefe et al. 1992).

#### 2.4 The number, size and spacing of replication foci in different subphase of S

Figure 2 shows the typical “beads on a string” structure of replication foci seen in Cy3-dUTP fluoresced early S-phase PCC. The “beads on a string” structure is more evidently seen on the prominent fiber as shown in Figure 1m,n. The distance between foci was easily measured than those seen in interphase nuclei. Table 1 (reproduced from Chromosoma. 2007; 116(5): 453-462, Gotoh 2007) summarizes the number, size and the spacing of the foci of different S-phase stages. The number of foci scored maximum ~1400 in early S-phase that was much precisely than ever reported (Manders et al. 1996; Jackson and Pombo 1998; Ma et al. 1999), then decreased to ~100 in very late S-phase (Gotoh 2007). The size changing of foci during S-phase was also shown clearly (Gotoh 2007). The spacing between replication foci clusters was ranged from 85.3 kbp (0.64  $\mu\text{m}$ ) to 536 kbp (4.02  $\mu\text{m}$ ) (average, 208 kbp, 1.56  $\mu\text{m}$ , Gotoh 2007) based on the assumption that 0.75  $\mu\text{m}$  length of 30 nm chromatin fiber is

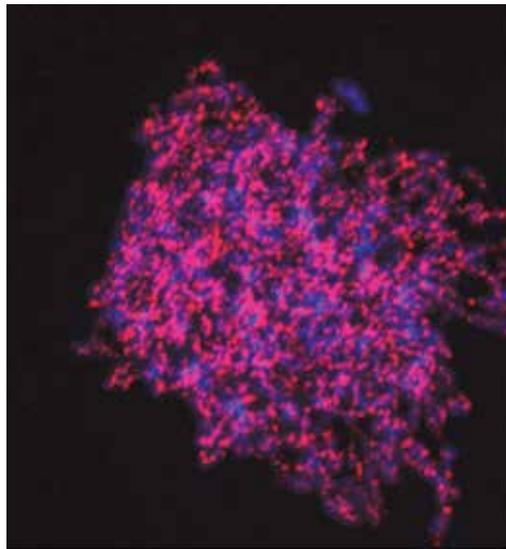


Fig. 2. Cy3 fluoresced DNA replication regions (replication foci) of early S-phase prematurely condensed chromosomes (PCCs). DNA replication regions were directly labeled with Cy3-dUTP by a bead-loading method followed by calyculin-A induced premature chromosome condensation. DNA was counterstained with DAPI. Chromosome spread shows a “cloudy nebula” appearance and, active replication regions (red) along with DNA fiber (faint blue) were seen as “beads on a string” structure. Condensed chromosome regions after finishing of DNA replication were seen as thick blue.

equivalent to 100 kbp (linear packing ratio of  $\sim 40:1$ ) (Berezney et al. 2000). Fiber autoradiography of DNA labeled with [ $^3\text{H}$ ]thymidine revealed the synchronized synthesizing replicon of spacing distance ranging from 50-300kbp (Fakan and Hancock 1974; Edenberg and Huberman 1975; Hand 1978). Therefore the spacing of replication foci seems to vary more than previously reported values. The measure of the spacing of foci were also reported on stretched DNA fiber with detergent and calculated the average distance of foci as 144 kbp (Jackson and Pombo 1998) that is fairly good agreement with the present and the previous studies, although the intact structure of foci was no more retained in stretched fiber after lysis treatment.

Subphase of S	early	middle	late	very late
Number of foci (max~min) <sup>a</sup>	1046 $\pm$ 187 (1396~758)	678 $\pm$ 83 (887~601)	450 $\pm$ 132 (607~237)	153 $\pm$ 42 (245~98)
Size of foci (max~min) ( $\mu\text{m}$ radius) <sup>b,c</sup>	0.35 $\pm$ 0.01 (0.08 <sup>d</sup> ~0.62)	0.52 $\pm$ 0.06 (0.1 <sup>d</sup> ~0.62)	0.98 $\pm$ 0.23 (0.24~1.44)	1.1 $\pm$ 0.44 (0.32~4.1)
Spacing between foci (max~min) ( $\mu\text{m}$ ) <sup>b,c</sup>	1.56 $\pm$ 0.68 (0.64~4.02)	1.01 $\pm$ 0.11 (0.33~1.86)	N.D. <sup>e</sup>	N.D. <sup>e</sup>

a: For each subphase, at least 12 spreads were scored., except for very late S (10 spreads to be scored were available)

b: Measured using the Zeiss LSM510 software

c: Randomly selected 20 points were measured, and the average and the error were calculated.

d: These data were beyond the resolution of optical microscope, and measured on digitized images.

e: Not done the measure because PCCs at these stages mostly condense as mitotic chromosomes thus measure the spacing seems less meaningful.

Table 1. The number, size and spacing of replication foci of different subphase of S (reproduced from Chromosoma. 2007; 116(5):453-462, Gotoh 2007)

## 2.5 A hypothetical chromosome compaction model coupled with DNA replication

Numbers of model for eukaryote chromosome architecture have been proposed (Marsden and Laemmli 1979; Woodcock et al. 1984; Woodcock and Dimitrov 2001; Swedlow and Hirano 2003; Kireeva et al. 2004), but they are controversial and many of things remain unclear. In addition, these models do not take account the involvement of DNA replication/transcription in chromosome packaging. DNA/RNA polymerase are known to be tightly immobilized to the replication/transcription factories (Cook 1999; Frouin et al. 2003). In the proposed model, DNA polymerase is thought to "reel in template DNA and extrude replicated DNA" (Hozak et al. 1996; Cook 1999), rather than the enzyme track along DNA template proposed in many conventional model. In Cook's model, some kinds of mechanical tension force should be generated in the template DNA along with DNA replication goes on because the factory does not freely suspended in the nucleus but attached to nucleoskelton, consequently this force may pull and aggregate the replication foci of both side as to release the tension in DNA strand, which may resulting in formation of chromosomes as seen in mitosis. Based on the above mechanism and the observed findings obtained from chromosome structure dynamics coupled with DNA replication, Figure 3 shows a hypothetical model for the relationship of DNA replication and

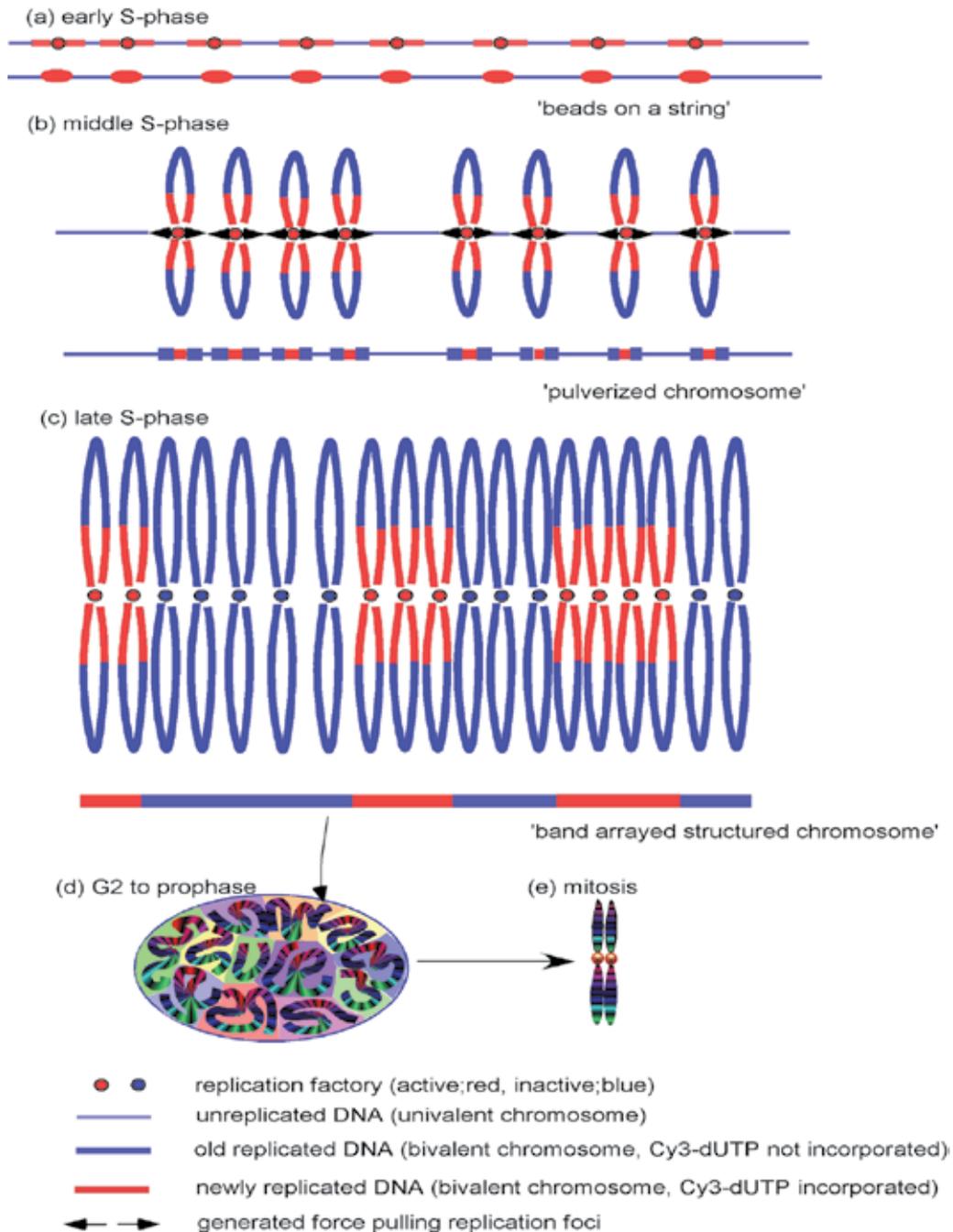


Fig. 3. A hypothetical two-dimensional model for chromosome conformational change involving DNA replication based on the models proposed by Cook (Cook 1995) or Pflumm (Pflumm 2002). (a) Early S-phase. DNA replication starts at multiple origins and proceeds bi-directionally. Early S-PCCs are seen as 'beads on a string' appearance. (b) Middle S-phase. As DNA replication proceeds, replicated DNA pass through replication factory and some tension

are generated. The generated tension may pull back the replication factories close together so as to release the tension. Replication factories may in turn fuse together and chromosomes compact. Middle S-PCCs are seen as well known 'pulverized chromosomes' appearance. (c) Late S-phase. Most of DNA finished replication and conformation was changed. Late S-PCCs are seen as 'tandem band arrayed structured chromosomes' like as mitotic chromosomes. (d) G2 to prophase. After finishing of DNA replication, chromosome conformation changed like as mitotic chromosomes, but still so elastic that packed in nucleus. Before fixation, each chromosome occupies individual chromosome territory (CT) in interphase nucleus, thus observed as compartment regions (colorized). (e) Mitosis. After prophase, chromosomes further shortening in longitudinal axis of chromosomes, consequently a straight rod shaped recognizable chromosome formed as usually seen by cytologists. For simplicity, the model is shown as two-dimensional and the scaling is arbitrary. The model intends not to depict actual events of chromosome conformation change but to help imagine how DNA replication is involved in chromosomal conformation. As the real chromosomes condense as three-dimensionally, other elements such as coiling and helical winding should be considered together to construct a stereoscopic hierarchical structure of eukaryote chromosomes (Woodcock and Dimitrov 2001; Swedlow and Hirano 2003). (reproduced from *Chromosoma* 2007; 116(5): 453-462, Gotoh 2007)

chromosome conformational changes, that shows how interphase chromatin is constructed to chromosomes observed in mitosis (Figure reproduced from *Chromosoma* 2007; 116(5): 453-462, Gotoh 2007). During S-phase, chromosomal conformation changes as DNA replicates, and the chromosome formation would be mostly completed at the finish of DNA replication (Figure 3a,b,c). From G2 to prophase, chromosomes are still more elastic, less condensed, folded only several times and pre-aligned in interphase nuclei (Manders et al. 1999). Chromosomes, at these phases, would be observed as chromosome territories (Cremer et al. 1993; Berezney et al. 2000) (Fig. 3d). Entering in mitosis, chromosomes would condense even more as shortening the longitudinal axis to form solid and rod shape appearance of recognizable mitotic chromosomes (Manders et al. 1999) (Fig. 3e).

### **3. Visualizing the DNA replication coupled spatial and temporal rearrangement of $\gamma$ -ray cleaved chromosome fragment**

#### **3.1 Background**

Chromosomes are easily damaged by many kinds of clastogenic agents such as ionizing irradiation (IR), ultra violet light (UV), chemicals (i.e. alkylating agents) or biological resources (i.e. some kinds of viruses) (Therman 1980). Mostly damaged chromosomes can be repaired before the cell division, some damages might be mis-repaired or remain unrepaired, resulting in chromosome aberrations (Savage 1991). These aberrations may cause cell death or cancer prone. Although the numbers of knowledge in molecular level approaches have been accumulated, little is still known about how, when and where the ends of individual chromosome fragments close and associate to form rearranged chromosomes (Savage 2000; Aten et al. 2004). It is clearly required that the cleaved fragments must contact physically to form rearranged chromosomes at some stages of the repair process (Savage 2000). Two major hypotheses are still controversial: Whether these fragments move to contact after damage (the "breakage first" hypothesis), or whether association of fragments happen only where close contact already exists at time of irradiation (the "contact first" hypothesis) (Savage 2000). As DNA is fragmented in

interphase and subsequent repair and rejoining finish until cell division, it should be required to elucidate the process occurs in interphase nuclei. Number of studies has challenged to analyze the dynamics of repair and recombination in interphase.

In this section, to elucidate how and when chromosome fragments close and association to form rearranged chromosomes in interphase, we visualized the dynamics of spatial and temporal occupation of chromosome 4 domains in  $\gamma$ -irradiated human peripheral blood lymphocytes in interphase nuclei and in drug-induced PCCs by means of chromosome painting method.

## **3.2 Materials and methods**

### **3.2.1 Cell culture, $\gamma$ -irradiation and chromosome preparation**

Human peripheral blood lymphocytes (PBLs) were separated from 10 ml of whole blood from healthy donor using LymphoPrep (Becton and Dickinson, Franklin Lakes, NJ) as per supplier's protocol. PBLs were then suspended in RPMI 1640 medium (GIBCO, Grand Island, NY) supplemented with 20% foetal calf serum (Biocell Laboratory, Ramincho-Domingues, CA). To induce chromosomal breaks, aliquot of PBLs were irradiated with 10 Gy of  $\gamma$ -rays ( $^{137}\text{Cs}$ , 0.662 MeV, 0.9 Gy/min) using a GammaCell 40 (Atomic Energy of Canada LTD, Canada). Irradiated PBLs were then stimulated immediately with PHA-P and rIL-2 as described previously (Gotoh and Asakawa 1996), cultured at 37 °C in 5%  $\text{CO}_2$  atmosphere with 95% humidification for different incubation time (from 24 hours to 120 hours). One hour before the each harvesting time, cellular DNA were labeled with 10  $\mu\text{M}$  of BrdU (10 mM of stock solution). The same irradiation experiment of PBLs except for without PHA-P stimulation or DNA-synthesis inhibition with 2  $\mu\text{M}$  aphidicholine were performed simultaneously. After harvesting, cells were swollen in 0.075M KCl hypotonic buffer and incubated for 20 minutes at 37°C. Then the swollen cells were fixed with cold Carnoy's fixative (methanol: acetic acid = 3 parts: 1part), fixed three times with the fixative and dropped on microscopic glass slides and air dried. For obtaining chromosome spreads, 50 nM calyculin A was added to the duplicated PBLs culture to induce PCC 30 minutes before harvesting the cells at each incubation time (Gotoh and Tanno 2005; Gotoh 2009). For control studies, aliquot of non-irradiated PBLs cells were harvested at either 48 hours or 72 hours after stimulation. After harvesting, cells were swollen in 0.075 M KCl for 20 minutes at 37 °C, fixed with methanol: glacial acetic acid (3: 1 vol./vol.), washed and fixed three times in the same fixative, dropped on a glass slide. Usually, the slides were stored in a desiccated box for one or two weeks for sufficient drying until chromosome painting was performed.

### **3.2.2 Chromosome painting**

Chromosome painting was done use with human chromosome 4 painting probe (GIBCO, Grand Island, NY) per se the supplier's protocol, except for additional washing with 0.1x SSC at 60°C for 5 minutes as described previously (Gotoh and Asakawa 1996). Briefly chromosomes were denatured in 70% formamide in 2xSSC at 70°C for 3 minutes followed by dehydration in ethanol series (70, 90 and 99.5%). A 5  $\mu\text{l}$  aliquot of hybridization mixture (50% formamide, 10% dextran sulphate and 100  $\mu\text{g}/\text{ml}$  carrier DNA in 2xSSC) containing probes (2 ng/ml) was denatured at 70°C for 2 minutes and loaded on a sample, covered with a cover slip ( $\phi$  15 mm) and sealed with a silicon rubber (Exabite normal type; GC Corp., Japan) to prevent dry up the hybridization buffer during hybridization. Hybridization was

carried out at 37°C overnight, then non-specific hybridized probes were washed out in 2xSSC at 37°C 5 minutes three times and in 0.1xSSC at 60°C 5 minutes. Cell nuclei were then counterstained with 200 ng/ml of propidium iodide (PI; Sigma, St Louis, MO) for SpectrumGreen™ labeled probe or with 200 ng/ml of diamidino-2-phenylindole (DAPI; Sigma, St Louis, MO) for SpectrumOrange™ labeled probe. The slides were mounted with 50% glycerol in PBS supplemented with 0.1% antifade (1,4-phenylene-diamine, Sigma, St Louis, MO) and observed under an epifluorescence microscope (Nikon, Japan), using B-2A or G-2 excitation-emission filter (Nikon, Japan) to visualize SpectrumGreen™ or SpectrumOrange™ fluorochrome, respectively.

### 3.2.3 BrdU detection for cell cycle analysis

Detection of incorporated BrdU in replicating DNA was performed as described previously (Gotoh et al. 1995; Asakawa and Gotoh 1997). After chromosome painting study, slides were immersed in 2N HCl at room temperature for 30 minutes to denature DNA and then immersed in 0.1M Na<sub>2</sub>B<sub>4</sub>H (pH 8.5) to neutralize the acid for a few minutes at room temperature. Slides were briefly washed twice with 0.5% Tween 20/PBS twice. Twenty µl of FITC conjugated anti-BrdU monoclonal antibody (Becton Dickinson) was diluted with 50 µl of 0.5% Tween 20/PBS and added on a glass slide, covered with a cut Parafilm sheet and incubated at room temperature for 30 minutes in a humidified chamber. Slides were washed twice with PBS and counterstained with 40 ng/ml of PI. Slides were observed under an epifluorescence microscope using a B-2A filter and images were taken using a CCD camera.

## 3.3 Spatial and temporal rearrangement of chromosome fragment coupled with DNA replication following $\gamma$ -irradiation

### 3.3.1 Determination of irradiation dose for effective detection of the chromosome fragments in interphase nuclei

Human peripheral blood lymphocytes (PBLs) is an excellent system for study the irradiation induced chromosome aberrations, because *in vivo* PBLs arrest in G<sub>0</sub> phase (quiescent stage) and start cell cycle well synchronized until first cell division after stimulation by foreign antigens or non-specific mitogens such as LPS (lipopolysaccharides) or PHA (phytohemagglutinine). It is well known that the frequency of chromosome breakage increasing as the increase of irradiation dose (Gotoh and Asakawa 1996; Gotoh et al. 2005). Accordingly it is expected that increasing the irradiation dose will make easier to observe the fragmented chromosome domain because the number of fragmented domains will increase seen in interphase nuclei. It is, however, difficult or even impossible to obtain sufficient number of chromosomes from cells exposed high dose of ionizing irradiation, because these cells arrest in S or G<sub>2</sub> and do not enter in mitosis. As drug-induced PCC allows us to obtain chromosomes in cells even irradiated 40 Gy dose of  $\gamma$ -rays (Gotoh and Asakawa 1996; Gotoh and Tanno 2005; Gotoh et al. 2005; Wang et al. 2009; Balakrishnan et al. 2010; Lindholm et al. 2010), we used to calyculin A induced PCC method to visualize the chromosome aberration in G<sub>2</sub> cells after  $\gamma$ -irradiation, but we irradiated the cells with more milder dose of 10 Gy of  $\gamma$ -rays as not give severe damages to the cells. Figure 4 (reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press) shows the chromosome painting of human #4 chromosomes against calyculin A induced prematurely condensed chromosomes (PCCs)

of PBLs at 48 hours after 10 Gy of  $\gamma$ -rays irradiation. At this time point, number of mitotic chromosomes was quite very low (lower than 0.01% as shown in Table 2), which indicates that the observed chromosomes is G2-PCC of cells arrested in G2 rather than mitosis. G2-chromosomes (G2-PCC) were moderately damaged by 10 Gy of  $\gamma$ -irradiation, and more than 10 substantially recognizable fragments of chromosome 4 were inserted in rearranged chromosomes (indicated by arrowheads), as previously reported (Gotoh and Asakawa 1996; Gotoh and Tanno 2005; Gotoh et al. 2005). This also suggests that the cells retain repair ability of double strand breaks (DSBs) by  $\gamma$ -irradiation. We expected that more than 10 fragmented pieces of chromosome 4 domains can be observed in an interphase nucleus after 10 Gy of  $\gamma$ -rays irradiation.

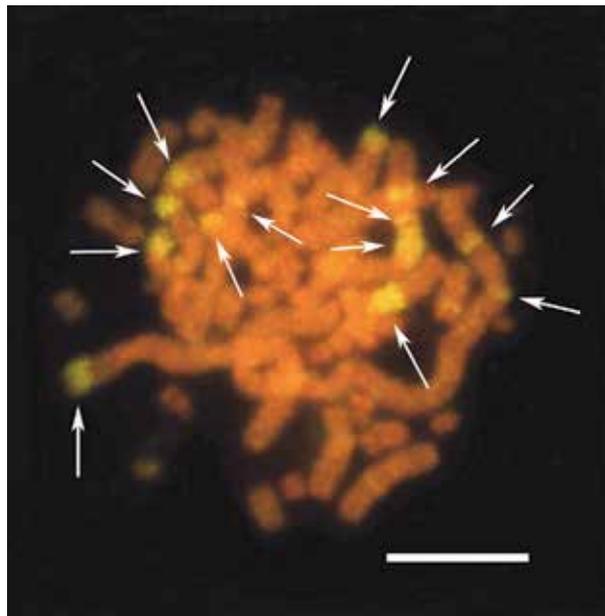


Fig. 4. Chromosome painting on a calyculin A induced PCCs at 48 hours after irradiation of 10 Gy of  $\gamma$ -rays. SpectrumGreen™ labeled specific probe for human chromosome 4 was used for detection of chromosomal fragments derived from chromosome 4. Chromosomes were counterstained with propidium iodide (PI). Moderate damaged chromosomes including long and short size of ones are clearly seen. Arrows indicate many fragmented pieces of chromosomes 4 which inserted in rearranged chromosomes. Scale bar : 10  $\mu$ m. (reproduced from Molecular Biology International, Special Issue on “DNA in 3R: Repair, Replication, and Recombination”, Gotoh and Asakawa 2011, in press)

### 3.3.2 Cell cycle analysis of control peripheral blood lymphocytes (PBLs) after mitogen stimulation

As a control study to identify the cell cycle stage of PBLs after PHA-P stimulation, we performed chromosome painting and BrdU incorporation analysis on interphase nuclei of peripheral blood lymphocytes (PBLs) without irradiation, which are shown in Figure 5 (left column, chromosome painting; right column, BrdU incorporation, reproduced from Molecular Biology International, Special Issue on “DNA in 3R: Repair, Replication, and

Sampling Time (hours)	Dose (Gy)	Treatment	S PCC (% cell)	G2/Mphase (% cell) <sup>†</sup>	Interphase (% cell)
0	0 (control)	Calyculin A	0.0	0.0	100.0
		Colcemid	- **	0.0	100.0
	10	Calyculin A	0.0	0.0	100.0
		Colcemid	- **	0.0	100.0
24	0 (control)	Calyculin A	17.6	0.0	82.4
		Colcemid	- **	0.0	100.0
	10	Calyculin A	14.8	0.0	85.2
		Colcemid	- **	0.0	100.0
48	0 (control)	Calyculin A	16.3	11.5	72.2
		Colcemid	- **	1.4	98.6
	10	Calyculin A	12.1	17.4	70.5
		Colcemid	- **	0.01	99.99
72	0 (control)	Calyculin A	11.5	9.7	78.8
		Colcemid	- **	2.4	97.6
	10	Calyculin A	9.3	16.1	74.6
		Colcemid	- **	0.2	99.8

Table 2. Percentage population of PCC or metaphase obtained with calyculin A treatment (PCC) or colcemid treatment (metaphase) at different sampling time point after 10Gy of  $\gamma$ -irradiation. N=200 G2/M-PCC or Mitotic cells are observed in calyculin A or colcemid treated cells, respectively. \*\* In case of colcemid treatment, S-PCC is not observable. (reproduced from Molecular Biology International, Special Issue on „DNA in 3R: Repair, Replication, and Recombination“, Gotoh and Asakawa 2011, in press)

Recombination“, Gotoh and Asakawa 2011, in press). Figure 5a and 5c show two chromosome 4 domains in interphase nuclei and Figure 5b and 5d are incorporation of BrdU in the cells of the same fields at 48 hours after PHA-P stimulation. At 48 hours after PHA-P stimulation, many PBL cells entered in S-phase, which was clearly shown by much amount of BrdU incorporation as shown by the previous study using Cy3-dUTP labeling cell (Gotoh 2007). The BrdU positive cells (proliferating cells) increased nuclear size compared with the BrdU negative cells (non-proliferating cells), which suggested the increase of DNA amount as DNA replication progressed. Two intact chromosome 4 domains were clearly recognized either in BrdU positive larger sized nuclei (enter in S-phase) or BrdU negative small nuclei (still in G0/G1-phase). At 72 hours after PHA-P stimulation, further enlarged BrdU positive nuclei were seen (Figure 5e-5h), which suggested that DNA replication progressed much more. Some nuclei enlarged much more and BrdU was less incorporated, which suggests that these cells were already in very late S phase as almost completion of DNA replication. One pair of chromosomes 4 stained with painting probe is indicated by arrows (Figure 5i). BrdU was not incorporated in the mitotic chromosomes (Figure 5j). The incorporation patterns of BrdU in replicating DNA at each cell cycle stage was the similar as described previously either using BrdU labeling (Nakamura et al. 1986; O'Keefe et al. 1992) or Cy3-dUTP labeling (Manders et al. 1999; Gotoh 2007). Two chromosomes 4 domains in a nucleus were not fully diffused over the nucleus but they occupy separated spaces each other in an interphase nucleus throughout cell cycling (Lengauer et al. 1991; Cremer et al. 1993; Savage 2000; Aten and Kanaar 2006).

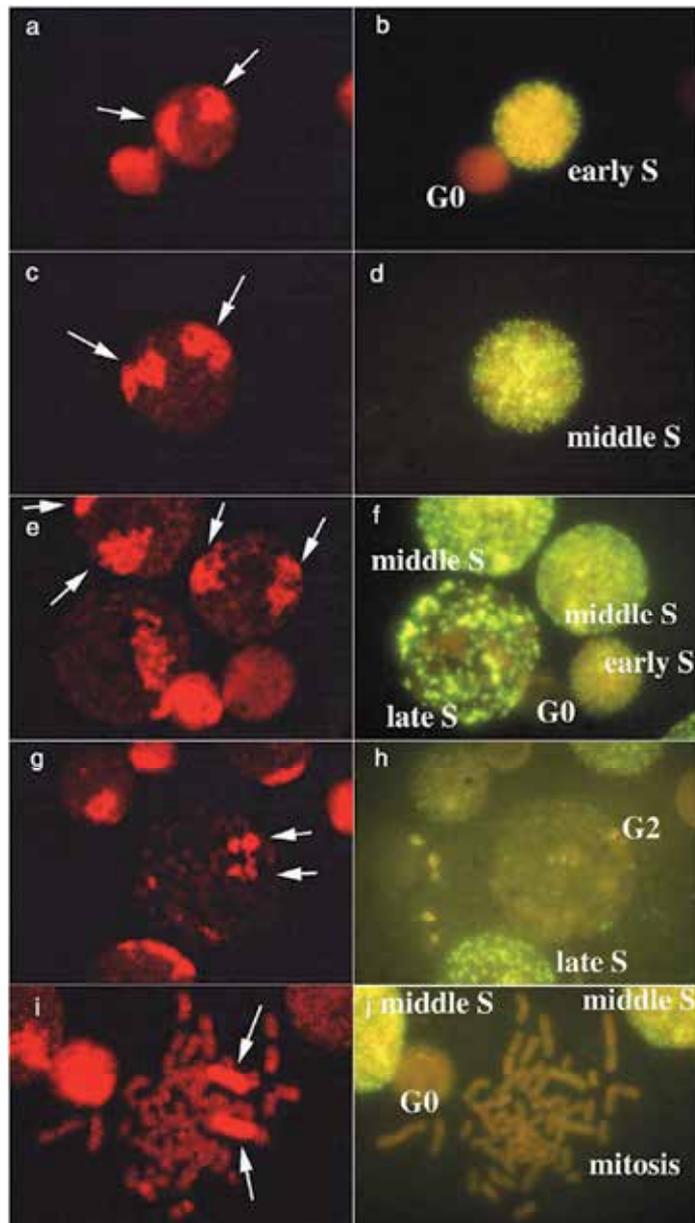


Fig. 5. Chromosome painting and incorporation of BrdU in various cell cycle stages of human peripheral blood lymphocytes stimulated with PHA-P. Left column (a, c, e, g and i): Chromosome painting using SpectrumOrange™ labeled specific probe for human chromosome 4. Arrows indicate two chromosome 4 (i) or two chromosome 4 domains in interphase nuclei (a, c, e and g). Right column (b, d, f, h and j): Incorporation of BrdU in cells in the same field. Incorporated BrdU was detected by FITC conjugated anti-BrdU antibody. Cell cycle stage of individual cell is indicated in figures. Scale bar: 10  $\mu$ m. (reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press)

### 3.3.3 Visualizing spatial and temporal dynamics of $\gamma$ -ray cleaved chromosome fragments in interphase nuclei

We next studied the spatial and temporal distribution of chromosomes 4 domains after  $\gamma$ -rays irradiation. Immediately after 10 Gy of  $\gamma$ -irradiation, PBLs were stimulated with T-cell specific mitogen PHA-P and cells were harvested at different culture time as control study (every 24 hour up to 72 hours; usually 48 hours incubation time is enough to obtain sufficient number of G2-PCCs even in cells  $\gamma$ -irradiated of 10 Gy). Table 2 (reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press) shows the percentage population of PCC or mitotic cell observed in calyculin A or colcemid treated cell samples. At 48 or 72 hours after  $\gamma$ -ray irradiation, more than 10 % of G2/M PCCs were obtained and chromosomes were highly damaged as shown in Figure 4, but number of mitotic cells was very low. As shown in Table 2, control un-irradiated human PBLs started cell cycle from G0 and at 48 hours after stimulation, 1.4% of cells reached in mitosis but others still in G2 or earlier stages even the case of un-irradiated cells. Therefore, the observed chromosome aberration seen in G2-PCCs is before first mitosis after irradiation. Figure 6 (reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press) shows the dynamics of chromosomal domain transition through the cell cycling. At 24 hours after irradiation, mostly cells did not enlarge their size and seemed to be still in G1 stage (population of stimulated or un-stimulated cell is shown in Table 2). At this stage, the cells showed two intact (un-fragmented) chromosome domains (Table 3, Figure 5a), that is quite the same as observed in G0/G1 nuclei of control un-irradiated PBLs cells. At 48 hours after irradiation, some cells slight enlarged their size, which suggests that these cells entered in S-phase. As cell enlarged, two chromosomal domains began slightly enlarged (Figure 5b), but still kept in intact form as a control S-phase PBLs nuclei as shown in Figure 4. However, in a small population (6.3%) of cells, chromosomal domains seemed somehow coarse appearance or rather seemed fragmented (Figure 5c), which suggested that chromosome domains fragments started separation and dispersion. Such phenomenon was not totally observed in control PBLs. At 72 hours after  $\gamma$ -irradiation (Figure 6d-6f), the number of enlarged cell became much more as 60.3%. The degree of separation and dispersion of chromosomal fragments was much more clearly in 58.5% of cells as the increase of the cell nuclear size.

Sampling Time(hours)	Non-proliferating cells (small sized cell)		proliferating (blastocytes) cells (enlarged sized cell)	
	fragmentation (-)	fragmentation (+)	fragmentation (-)	fragmentation (+)
0	100	0	0	0
24	95.4	0	4.6	0
48	42.6	0	51.1	6.3
72	39.7	0	1.8	58.5

Table 3. Population of non-proliferating or proliferating (blastocytes) peripheral blood lymphocyte stimulated by PHA-P, with or without fragmentation of chromosome domain at different sampling time after irradiation of 10 Gy of  $\gamma$ -rays were scored. N=200. (reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press)

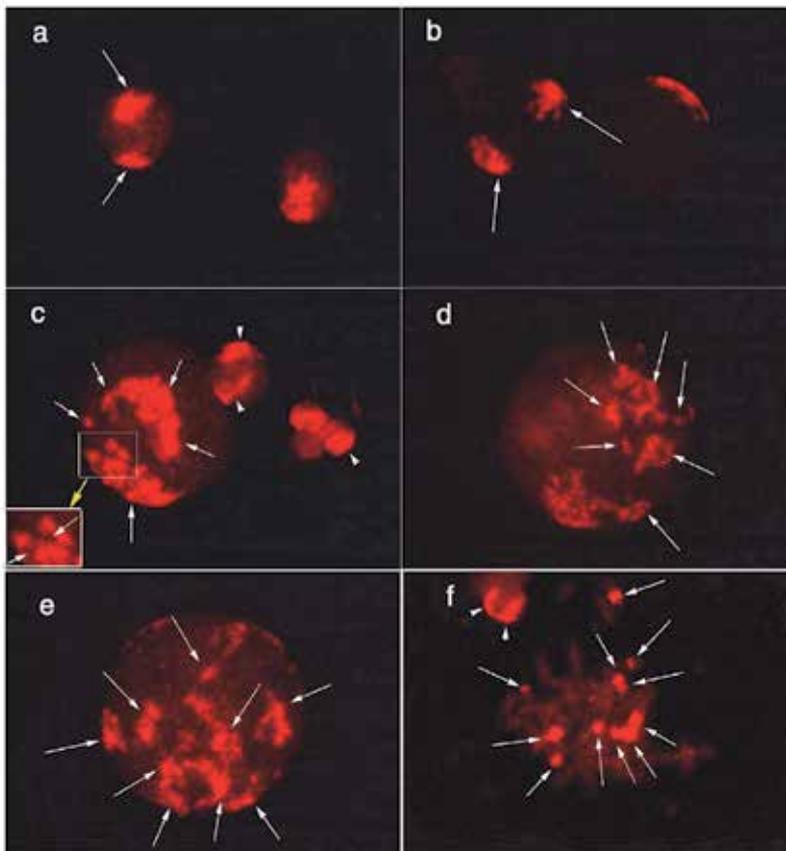


Fig. 6. Chromosome 4 painting on nuclei of peripheral blood lymphocytes irradiated with 10 Gy of  $\gamma$ -rays. Chromosome painting with SpectrumOrange™ labeled probe for chromosome 4. (a and b) Arrows indicate chromosome 4 domains. (c) Arrows indicate some of the fragments of chromosome 4. Note that the cells not enlarge their size (i.e. G0 arrested cells) have two un-fragmented chromosome domains (indicated by arrowheads). (Inset in c) Magnifying the white line boxed region, arrows indicate gaps between fragmented domains. (d-f) Arrowheads indicate some of the chromosome 4 fragments. Arrowheads in Figure 3f indicate two un-fragmented chromosome domains in the cells which not enlarge their size (i.e. G0 arrested cells). Scale bar: 10  $\mu$ m. (reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press)

According to the nuclear size, the degree of condensation of chromosome domain (also described by Mukherjee et al. 1992) and the degree of dispersion of the fragments, we could order the temporal dynamic series of chromosome domain diffusion as from Figure 6d to 6f. Figure 6d shows the further progress moving, dispersion and separation of the fragmented chromosome. Furthermore, as shown in Figure 6e, individual small fragmented chromosome domain completely separated over the nucleus. At 48 hours after irradiation, some cells already reached in G2 stage (Gotoh and Tanno 2005). Figure 6f shows an example of calyculin induced G2-PCCs, which clearly shows severely damaged chromosomes. Many

pieces of fragmented chromosome 4 were inserted (shown by arrowheads) in highly aberrant chromosomes.

### **3.3.4 Spatial and temporal transition of chromosome fragment is coupled with DNA replication**

We noticed that about 40% of cells did not enlarge their size, which suggested that these cells were not proliferated and arrested in G<sub>0</sub>-stage. These cells might be B-cells, monocytes or natural killer cells, which are consist of mononuclear cells in peripheral blood. Because we stimulated the MNCs with T-cell specific mitogen PHA-P, the cells except for T-cells were not stimulated and thought to arrest in G<sub>0</sub> stage. Stimulation failed T-cells also involved in this small sized cell population. Surprisingly and unexpectedly, two chromosome domains were kept as two un-fragmented (see like as two intact domains) in these non-proliferated cells and did not diffuse over the nucleus (indicated by arrowheads in Figure 6c and 6f), although these cells must be received the same 10 Gy dose of  $\gamma$ -rays as proliferating cells. We did not see any un-stimulated cells having fragmented chromosome domain. We therefore hypothesized that spatial and temporal redistribution of  $\gamma$ -irradiated chromosome fragments might be coupled with cell proliferation.

To confirm this hypothesis, we  $\gamma$ -irradiated PBLs as same as the above experiment (i.e. 10 Gy of irradiation dose) but stimulation with PHA-P was omitted. At 48 or 72 hours after irradiation, more than half of cells were lost presumably because of interphase death caused by lack of mitogen stimulation, however some cells were still alive. Figure 7a (reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press) shows chromosome painting against these un-stimulated cells. The size of cell and nucleus did not increase and the chromosome 4 domains retained as intact. This finding also seems to support that spatial redistribution of fragmented chromosomes in interphase nuclei is tightly coupled with cell cycle.

We next questioned whether the spatial redistribution of fragmented chromosomes in S-phase is coupled with DNA replication. We irradiated PBLs with 10 Gy of  $\gamma$ -rays, stimulated with PHA, but cellular DNA replication was inhibited by 2  $\mu$ M of aphidicholin. As the same as observed in the irradiation experiment of PBLs without PHA-P stimulation, chromosomal domains were kept as un-fragmented, in spite of received 10 Gy dose of  $\gamma$ -ray exposure as PHA stimulated cells (Figure 7b). According to these results, we concluded that nuclear movement that causes spatial redistribution of the fragments was tightly coupled with cell cycling and in particular with DNA replication.

### **3.4 Hypothetical model of chromosome repair/recombination coupled with DNA replication**

Two possible models have been proposed for explain how chromosome fragments close and associate to form radiation-induced rearranged chromosomes. Obviously, radiation-induced chromosome fragments must "touch" at some stage of the repair process (Savage 2000). One model is that fragmented chromosomes move to contact with another apart from fragment after breakage by irradiation ("breakage first" hypothesis or post-dispersal rejoining, Figure 8 (1) a and Figure 8 (2) a, figures reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press). Another model is that fragment association occurs only where close contact already exists. Fragments adjacent to another fragment have rejoined each other and moved away ("contact first" hypothesis or pre-dispersal rejoining, Figure 8 (1) b and Figure 8 (2) b,

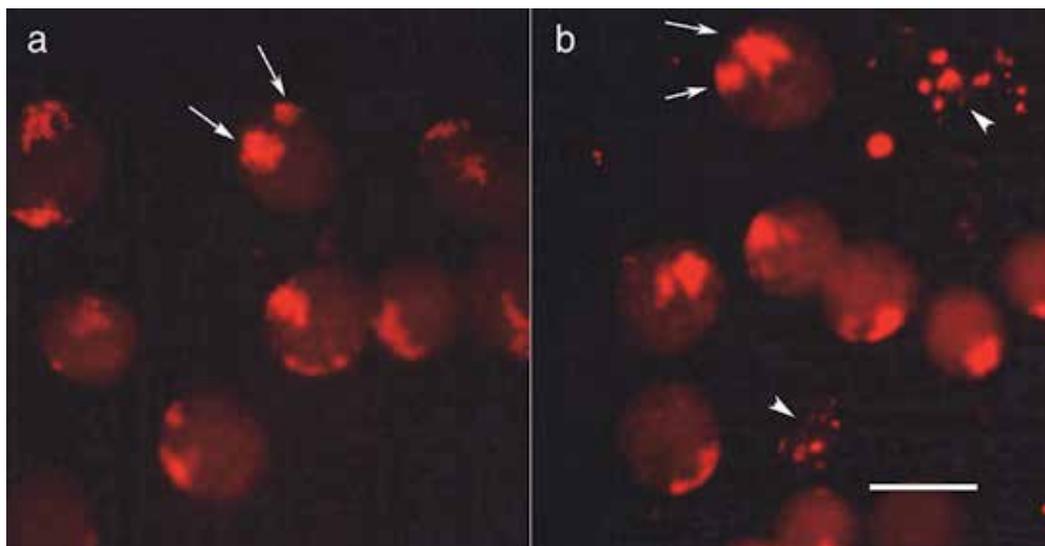


Fig. 7. (a) Chromosome painting against 10 Gy dose of  $\gamma$ -irradiated nuclei without stimulation by PHA-P, therefore cells yet arrested in G0 at 48 hours after irradiation. Two un-fragmented chromosome 4 domains are clearly visible (indicated by arrows). (b) Chromosome painting against 10 Gy dose of  $\gamma$ -irradiated cell nuclei, but DNA synthesis was inhibited by 2  $\mu$ M of aphidicholin. Arrows indicate one pair of un-fragmented chromosome 4 domain. Arrowheads indicate apoptotic fragmented nuclei, which is possible caused by DNA synthesis inhibition by aphidicholin. Scale bar: 10  $\mu$ m. (reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press)

figures reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press). In the case of the latter model, rejoining might occur limited to among those fragments adjacent to each other at G0/G1 phase. Recently accumulated evidences indicate that the radiation-induced exchanges cannot be possible between very far apart DNA strand breaks in chromosomes (Savage 2000). In the present report, quiescent G0 lymphocytes were  $\gamma$ -irradiated with 10 Gy of  $\gamma$ -rays, and chromosome type aberrations (i.e. acentric fragments, di-centric chromosome or translocations) are exclusively observed in G2-PCC at either 48 hours or 72 hours after irradiation. The reason for formation of chromosome type aberrations in irradiated quiescent lymphocytes is that damages were introduced in G0/G1 cell, repair or mis-rejoining of univalent chromosomes finishes before duplication, then DNA duplicates and result in bivalent chromosome type aberrations (Savage 1975). In this context, contact and association of cleaved chromosome ends should be finished before the cells enter in S-phase for the case of chromosome type aberrations formation. As the chromosome domain occupies the fixed space apart away in nucleus until S-phase starts as shown in Figure 5 and Figure 6, "breakage-first" hypothesis seems less likely to responsible in chromosome type aberration formation. Because chromosome fragment must move before begin of S-phase, but chromosome fragment move actually after S-phase starts. Thus the results reported here seems strongly support the "contact first" hypothesis in chromosome type aberration

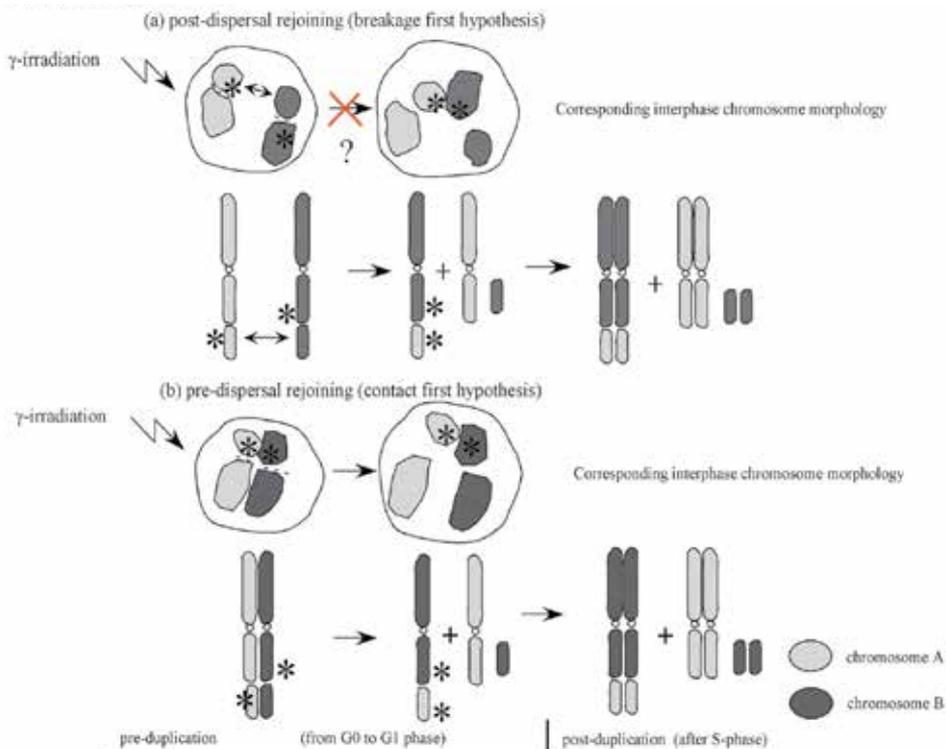


Fig. 8. Two models for chromosome cleavage and re-union involving nuclear movement progress in interphase nuclei  
 (1) Chromosome type aberration. Chromosome type aberration is introduced by DSBs in un-replicated univalent chromatid (Savage 1975). Therefore, chromosome domain is fragmented by DSBs. (a) Chromosome-end rejoining occurs after dispersion of chromosome fragments (post-dispersal rejoining or dynamic “breakage-first” model). In this case, chromosomes were cleaved by  $\gamma$ -irradiation and the fragments moved away. Then the fragments moved to touch and rejoin with fragmented chromosomes. However, breakage-first model does not seem likely to happen, because association should be finished before chromosome duplication. However, chromosome fragments may not move before DNA replication starts. (b) Rejoining occurs before chromosomal fragments dispersed (pre-dispersal rejoining or static “contact-first” model). In this case, chromosome rejoining occurs among the fragments that contact physically at the irradiated time (i.e. G0/G1 phase). Then recombined chromosomes moved away.

formation in  $\gamma$ -irradiated lymphocytes. Nikiforova et al. reported the inversion aberration of radiation-induced thyroid cancer (Nikiforova et al. 2000). They concluded that radiation-induced interchange between RET and H4 gene is presumably due to positional dependent and thus supports “contact first” hypothesis. Recently, Durante et al. showed that the exchange process was highly dependent on chromosome localization in nucleus or chromosome aberration was highly dependent on cell position and irradiation geometry (Durante et al. 2010). Therefore, their result also seems to support the “contact first” hypothesis. Contrary, clustering of fragmented domains after irradiation exposure was visualized using  $\gamma$ H2-AX, which suggests the “breakage-first” hypothesis is dominant in

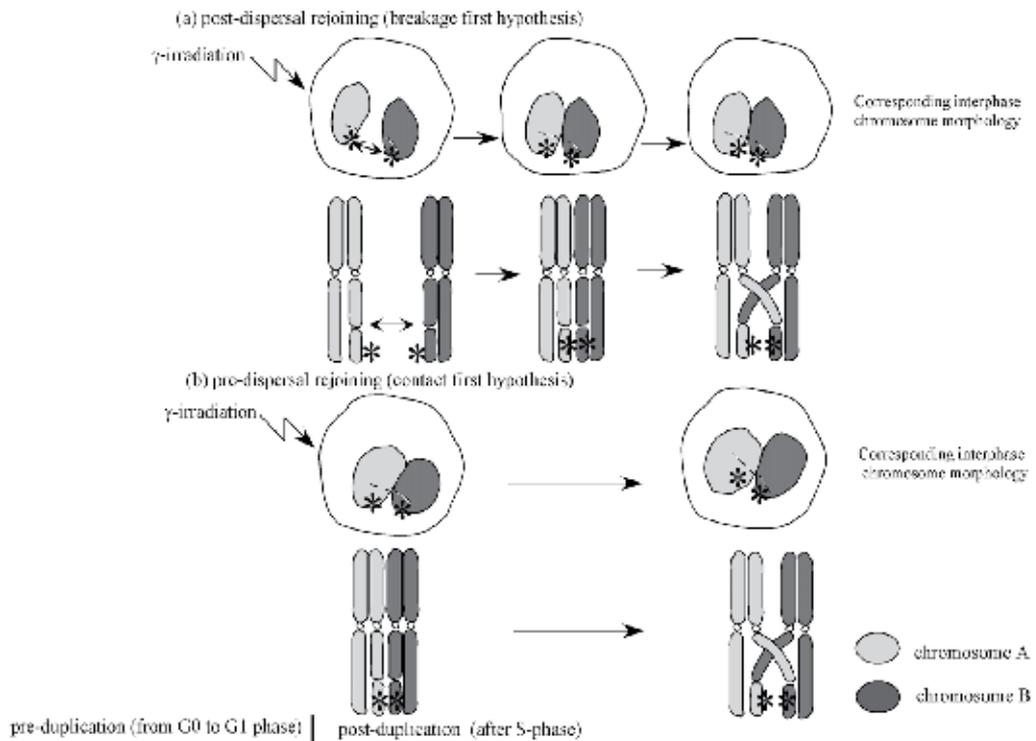


Fig. 8. Two models for chromosome cleavage and re-union involving nuclear movement progress in interphase nuclei

(2) Chromatid type aberration. Chromatid type aberration is introduced in the cells after S-phase by the DSBs only one of the sister chromatid of bivalent chromosome (Savage 1975). Chromatid does not apart away, because the cleaved chromatid coheres with opposite sister chromatid. "Breakage-first" model (a) and "contact-first" (b) model may both responsible in chromatid type aberration formation, because the nuclear movement already exists in the cells after DNA replication starts. Asterisks (\*) indicate the chromosome ends, which will later contact and rejoin. Following points are for the simplicity: only one DSBs per chromosome are illustrated and indicated by dotted line. Many types of chromosome or chromatid type aberrations are known to be produced after ionizing irradiation exposure, but only typical type of aberrations are illustrated. Illustration is shown as a 2-D image for simplification, however dispersal and rejoining of chromosome proceed in 3-D, of course. (reproduced from Molecular Biology International, Special Issue on "DNA in 3R: Repair, Replication, and Recombination", Gotoh and Asakawa 2011, in press)

exchange process (Aten et al. 2004). Which mechanism works principle in chromosome ends association may be depend on the existence of DNA replication coupled nuclear movement. If replication coupled nuclear movement exists (i.e. S-phase or later), both "breakage first" and "contact first" hypothesis may responsible. Aten et al.'s report seems to support this idea because they observed exchange formation in  $\alpha$ -particle irradiated cycling HeLa cells (Aten et al. 2004). On contrast, "contact first" hypothesis only may contribute for chromosome ends joining if not nuclear movement occurred (i.e. before S-phase). It is, however still controversy, which mechanism is dominantly work or both are simultaneously occur, and should be elucidated in the future.

The finding also implies other aspects as follows: 1) Generally it is believed that cycling cells are more sensitive to ionizing radiation or other agents that causes DNA strand break. Existence of S-phase coupled kinetics that drives fragmented chromosomes diffuse over nuclei may be one explanation for such phenomenon. After receiving damages, cell must repair the damages. If nuclear movement begin before the damage repaired completely, unrepaired fragments may be separated away from each other, which gives the cells irreversible chromosome aberrations which in turn resulting in cell death or prone cancer. 2) By analysis of unstable type chromosomal aberration (such as di-centric chromosome or acentric fragment) induced by irradiation, it is known that quiescent PBLs survive more than 10 years in vivo (Buckton et al. 1967). PBLs that are not stimulated by foreign antigens are thought to be in resting state in vivo for long time. Cleaved DNA fragments are not separated and dispersion did not occur until cells enter in S-phase. So the PBL cells at resting state are stable and may survive for a long period in vivo.

#### **4. Summary and conclusion**

To visualize the dynamics of chromosome structure formation coupled with DNA replication, Cy3-dUTP direct-labeled active replicating DNA was observed in prematurely condensed chromosomes (PCCs) utilized with drug-induced premature chromosome condensation (PCC) technique that facilitates the visualization of interphase chromatin as well as condensed chromosome form. S-phase PCCs revealed clearly the drastic dynamic transition of chromosome formation during S-phase along with the progress of DNA replication; from a "cloudy nebula" structure in early S-phase to numerous number of "beads on a string" in middle S-phase and finally to "striped arrays of banding structured chromosome" in late S-phase like as usual observed mitotic chromosomes. Drug-induced PCC clearly provided the new insight that eukaryote DNA replication tightly coupled with the dynamics of chromosome condensation/compaction for construction of eukaryote higher ordered chromosome structure.

To elucidate how and when individual chromosome fragments close and association to form rearranged chromosomes in interphase, we next visualized the dynamics of spatial and temporal occupation of chromosome 4 domains in  $\gamma$ -irradiated human peripheral blood lymphocytes in interphase nuclei and in drug-induced PCCs by means of chromosome painting method. After  $\gamma$ -irradiation exposure, breakage and dispersion of chromosome 4 domains starts when the cells entered in S-phase. This S-phase dependent dynamics was however not seen in un-stimulated cells or in the cells of which DNA replication was inhibited by aphidicholin. Thus the results seem to support "contact first" hypothesis for association of chromosome ends in chromosome type aberration formation. Spatial and temporal dynamics of chromosome fragments may tightly coupled with DNA replication, and this dynamic might be drive chromosome rearrangement.

Therefore, these findings presented in this chapter strongly suggested that DNA replication, rejoining/recombination and chromosome structure formation are tightly coupled mechanism in eukaryote cells.

#### **5. Condolences**

During preparation of this chapter, an enormously big Earthquake and Tsunami disaster attacked Japan 11<sup>th</sup> March 2011, by which a lot of people have been killed or lost. I lamentably dedicate this article to all of the victims.

## 6. Acknowledgement

I wish to express my thank to Prof. Aleksandar Lazinica and Prof. Viktorija Zgela for giving me an opportunity to write this manuscript.

## 7. References

- Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. (1989). *Molecular Biology of The Cell*. New York: Garland Publishing Inc.
- Asakawa Y, Gotoh E (1997) A method for detecting sister chromatid exchanges using prematurely condensed chromosomes and immunogold-silver staining. *Mutagenesis* 12: 175-177
- Aten JA, Stap J, Krawczyk PM, van Oven CH, Hoebe RA, Essers J, Kanaar R (2004) Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science* 303: 92-95.
- Aten JA, Kanaar R (2006) Chromosomal organization: mingling with the neighbors. *PLoS Biol* 4: e155
- Balakrishnan S, Shirsath K, Bhat N, Anjaria K (2010) Biodosimetry for high dose accidental exposures by drug induced premature chromosome condensation (PCC) assay. *Mutat Res* 699: 11-16
- Bekker-Jensen S, Lukas C, Kitagawa R, Melander F, Kastan MB, Bartek J, Lukas J (2006) Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *J Cell Biol* 173: 195-206
- Berezney R, Dubey DD, Huberman JA (2000) Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. *Chromosoma* 108: 471-484
- Buckton KE, Brown WM, Smith PG (1967) Lymphocyte survival in men treated with x-rays for ankylosing spondylitis. *Nature* 214: 470-473
- Burkard ME, Maciejowski J, Rodriguez-Bravo V, Repka M, Lowery DM, Clauser KR, Zhang C, Shokat KM, Carr SA, Yaffe MB, Jallepalli PV (2009) Plk1 self-organization and priming phosphorylation of HsCYK-4 at the spindle midzone regulate the onset of division in human cells. *PLoS Biol* 7: e1000111
- Chang P, Coughlin M, Mitchison TJ (2009) Interaction between Poly(ADP-ribose) and NuMA contributes to mitotic spindle pole assembly. *Mol Biol Cell* 20: 4575-4585
- Christensen TW, Tye BK (2003) *Drosophila* MCM10 interacts with members of the prereplication complex and is required for proper chromosome condensation. *Mol Biol Cell* 14: 2206-2215
- Cook PR (1995) A chromomeric model for nuclear and chromosome structure. *J Cell Sci* 108: 2927-2935
- Cook PR (1999) The organization of replication and transcription. *Science* 284: 1790-1795
- Cremer T, Kurz A, Zirbel R, Dietzel S, Rinke B, Schrock E, Speicher MR, Mathieu U, Jauch A, Emmertich P, Scherthan H, Ried T, Cremer C, Lichter P (1993) The role of chromosome territories in the functional compartmentalization of the cell nucleus. *Cold Spring Harbor Symp Quant Biol* 58: 777-792
- D'Antoni S, Mattina T, Di Mare P, Federico C, Motta S, Saccone S (2004) Altered replication timing of the HIRA/Tuple1 locus in the DiGeorge and Velocardiofacial syndromes. *Gene* 333: 111-119

- Deckbar D, Birraux J, Krempler A, Tchouandong L, Beucher A, Walker S, Stiff T, Jeggo P, Lobrich M (2007) Chromosome breakage after G2 checkpoint release. *J Cell Biol* 2007 176(6): 749-755.
- Draetta G, Beach D (1988) Activation of cdc2 protein kinase during mitosis in human cells: cell cycle-dependent phosphorylation and subunit rearrangement. *Cell* 54: 17-26
- Dunphy WG, Brizuela L, Beach D, Newport J (1988) The *Xenopus* cdc2 protein is a component of MPF, a cytoplasmic regulator of mitosis. *Cell* 54: 423-431
- Durante M, Furusawa Y, Gotoh E (1998) A simple method for simultaneous interphase-metaphase chromosome analysis in biodosimetry. *Int J Radiat Biol* 74: 457-462
- Durante M, Pignalosa D, Jansen JA, Walboomers XF, Ritter S (2010) Influence of nuclear geometry on the formation of genetic rearrangements in human cells. *Radiat Res* 174: 20-26.
- Edenberg HJ, Huberman JA (1975) Eukaryotic chromosome replication. *Annu Rev Genet* 9: 245-284
- El Achkar E, Gerbault-Seureau M, Muleris M, Dutrillaux B, Debatisse M (2005) Premature condensation induces breaks at the interface of early and late replicating chromosome bands bearing common fragile sites. *Proc Natl Acad Sci USA* 102: 18069-1807
- Fakan S, Hancock R (1974) Localization of newly-synthesized DNA in a mammalian cell as visualized by high resolution autoradiography. *Exp Cell Res* 83: 95-102
- Frouin I, Montecucco A, Spadari S, Maga G (2003) DNA replication: a complex matter. *EMBO Rep* 4: 666-670
- Gandhi R, Gillespie PJ, Hirano T (2006) Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Curr Biol* 16: 2406-2417
- Gillespie PJ, Hirano T (2004) Scc2 couples replication licensing to sister chromatid cohesion in *Xenopus* egg extracts. *Curr Biol* 14: 1598-1603
- Gollin SM, Wray W, Hanks SK, Hittelman WN, Rao PN (1984) The ultrastructural organization of prematurely condensed chromosomes. *J Cell Sci Suppl* 1: 203-221
- Gotoh E, Asakawa Y, Kosaka H (1995) Inhibition of protein serine/threonine phosphatases directly induces premature chromosome condensation in mammalian somatic cells. *Biomedical Research (Tokyo)* 16: 63-68
- Gotoh E, Asakawa Y (1996) Detection and evaluation of chromosomal aberrations induced by high doses of gamma-irradiation using immunogold-silver painting of prematurely condensed chromosomes. *Int J Radiat Biol* 70: 517-520
- Gotoh E, Kawata T, Durante M (1999) Chromatid break rejoining and exchange aberration formation following gamma-ray exposure: analysis in G2 human fibroblasts by chemically induced premature chromosome condensation. *Int J Radiat Biol* 75: 1129-1135
- Gotoh E, Tanno Y (2005) Simple biodosimetry method for cases of high-dose radiation exposure using the ratio of the longest/shortest length of Giemsa-stained drug-induced prematurely condensed chromosomes (PCC). *Int J Radiat Biol* 81: 379-385
- Gotoh E, Tanno Y, Takakura K (2005) Simple biodosimetry method for use in cases of high-dose radiation exposure that scores the chromosome number of Giemsa-stained drug-induced prematurely condensed chromosomes (PCC). *Int J Radiat Biol* 81: 33-40
- Gotoh E, Durante M (2006) Chromosome condensation outside of mitosis: Mechanisms and new tools. *J Cell Physiol* 209: 297-304 and Cover page

- Gotoh E (2007) Visualizing the dynamics of chromosome structure formation coupled with DNA replication. *Chromosoma* 116(5): 453-462
- Gotoh E (2009) Drug-Induced Premature Chromosome Condensation (PCC) Protocols: Cytogenetic Approaches in Mitotic Chromosome and Interphase Chromatin. In: Chellappan SP (ed) *Chromatin Protocols: Second Edition*. Humana Press, pp. 83-92
- Gotoh E (2011) Visualize Dynamic Chromosome. Deep Insight for the Atlas of Genetics and Cytogenetics in Oncology and Haematology (on line publication journal) <http://atlasgeneticsoncology.org/Deep/VisuDynChID20093.html> January 2011
- Gotoh E, Asakawa Y (2011) DNA replication coupled nuclear movement drives spatial and temporal rearrangement of gamma-irradiated cleaved chromosome fragments during cell cycling. *Molecular Biology International* (in press)
- Hameister H, Sperling K (1984) Description of a chromosome replication unit in individual prematurely condensed human S-phase chromosomes. *Chromosoma* 90: 389-393
- Hand R (1978) Eucaryotic DNA: organization of the genome for replication. *Cell* 15: 317-325
- Haren L, Gnadt N, Wright M, Merdes A (2009) NuMA is required for proper spindle assembly and chromosome alignment in prometaphase. *BMC Res Notes* 2: 64
- Hozak P, Jackson DA, Cook PR (1996) The role of nuclear structure in DNA replication. In: Blow JJ (ed) *Eukaryotic DNA Replication: Frontiers in Molecular Biology*. Oxford University Press, Oxford, pp. 124-142
- Ito S, Gotoh E, Ozawa S, Yanagi K (2002) Epstein-Barr virus nuclear antigen-1 is highly colocalized with interphase chromatin and its newly replicated regions in particular. *J Gen Virol* 83: 2377-2383
- Jackson DA, Pombo A (1998) Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *J Cell Biol* 140: 1285-1295
- Johnson RT, Rao PN (1970) Mammalian cell fusion: induction of premature chromosome condensation in interphase nuclei. *Nature* 226: 717-722
- Johnson RT, Gotoh E, Mullinger AM, Ryan AJ, Shiloh Y, Ziv Y, Squires S (1999) Targeting double-strand breaks to replicating DNA identifies a subpathway of DSB repair that is defective in ataxia-telangiectasia cells. *Biochem Biophys Res Commun* 261: 317-325
- Kireeva N, Lakonishok M, Kireev I, Hirano T, Belmont AS (2004) Visualization of early chromosome condensation: a hierarchical folding, axial glue model of chromosome structure. *J Cell Biol* 166: 775-785
- Kitajima TS, Sakuno T, Ishiguro K, Iemura S, Natsume T, Kawashima SA, Watanabe Y (2006) Shugoshin collaborates with protein phosphatase 2A to protect cohesin. *Nature* 441: 46-52. Epub 2006 Mar 2015.
- Kumagai A, Dunphy WG (1992) Regulation of the cdc25 protein during the cell cycle in *Xenopus* extracts. *Cell* 70: 139-151
- Lengauer C, Eckelt A, Weith A, Endlich N, Ponelies N, Lichter P, Greulich KO, Cremer T (1991) Painting of defined chromosomal regions by in situ suppression hybridization of libraries from laser-microdissected chromosomes. *Cytogenet Cell Genet* 56: 27-30
- Li J, Wang J, Jiao H, Liao J, Xu X (2010) Cytokinesis and cancer: Polo loves ROCK'n' Rho(A). *J Genet Genomics* 37: 159-172
- Lindholm C, Stricklin D, Jaworska A, Koivistoinen A, Paile W, Arvidsson E, Deperas-Standylo J, Wojcik A (2010) Premature chromosome condensation (PCC) assay for dose assessment in mass casualty accidents. *Radiat Res* 173: 71-78.

- Ma H, Samarabandu J, Devdhar RS, Acharya R, Cheng PC, Meng C, Berezney R (1998) Spatial and temporal dynamics of DNA replication sites in mammalian cells. *J Cell Biol* 143: 1415-1425
- Ma H, Siegel AJ, Berezney R (1999) Association of chromosome territories with the nuclear matrix. Disruption of human chromosome territories correlates with the release of a subset of nuclear matrix proteins. *J Cell Biol* 146: 531-542
- Maiorano D, Moreau J, Mechali M (2000) XCDT1 is required for the assembly of pre-replicative complexes in *Xenopus laevis*. *Nature* 404: 622-625
- Manders EM, Stap J, Strackee J, van Driel R, Aten JA (1996) Dynamic behavior of DNA replication domains. *Exp Cell Res* 226: 328-335
- Manders EM, Kimura H, Cook PR (1999) Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. *J Cell Biol* 144: 813-821
- Marsden MP, Laemmli UK (1979) Metaphase chromosome structure: evidence for a radial loop model. *Cell* 17: 849-858
- Masui Y (1974) A cytostatic factor in amphibian oocytes: its extraction and partial characterization. *J Exp Zool* 187: 141-147
- McNeil PL, Warder E (1987) Glass beads load macromolecules into living cells. *J Cell Sci* 88: 669-678
- Meyer H, Drozdowska A, Dobrynin G (2010) A role for Cdc48/p97 and Aurora B in controlling chromatin condensation during exit from mitosis. *Biochem Cell Biol* 88: 23-28
- Moir RD, Yoon M, Khuon S, Goldman RD (2000) Nuclear lamins A and B1: different pathways of assembly during nuclear envelope formation in living cells. *J Cell Biol* 151: 1155-1168
- Mukherjee AB, Murty VV, Chaganti RS (1992) Detection of cell-cycle stage by fluorescence in situ hybridization: its application in human interphase cytogenetics. *Cytogenet Cell Genet* 61: 91-94
- Mullinger AM, Johnson RT (1983) Units of chromosome replication and packing. *J Cell Sci* 64: 179-193
- Nakamura H, Morita T, Sato C (1986) Structural organizations of replicon domains during DNA synthetic phase in the mammalian nucleus. *Exp Cell Res* 165: 291-297
- Nikiforova MN, Stringer JR, Blough R, Medvedovic M, Fagin JA, Nikiforov YE (2000) Proximity of chromosomal loci that participate in radiation-induced rearrangements in human cells. *Science* 290: 138-141.
- Nishitani H, Lygerou Z, Nishimoto T, Nurse P (2000) The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature* 404: 625-628
- O'Keefe RT, Henderson SC, Spector DL (1992) Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences. *J Cell Biol* 116: 1095-1110
- Olins AL, Olins DE (1974) Spheroid chromatin units (v bodies). *Science* 183: 330-332
- Pantelias GE, Maillie HD (1983) A simple method for premature chromosome condensation induction in primary human and rodent cells using polyethylene glycol. *Somatic Cell Genet* 9: 533-547
- Pflumm MF, Botchan MR (2001) Orc mutants arrest in metaphase with abnormally condensed chromosomes. *Development* 128: 1697-1707
- Pflumm MF (2002) The role of DNA replication in chromosome condensation. *Bioessays* 24: 411-418

- Prasanth SG, Prasanth KV, Siddiqui K, Spector DL, Stillman B (2004) Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. *Embo J* 23: 2651-2663
- Savage JR (1975) Classification and relationships of induced chromosomal structural changes. *J Med Genet* 12: 103-122.
- Savage JRK (1991) Testing the participation of chromosomes in structural aberrations. In: Sobti RB and Obe G (ed) *Eukaryotic chromosomes Structural and functional aspects*. Springer Verlag, Berlin, Heidelberg, New York,
- Savage JR, Prasad R, Papworth DG (1984) Subdivision of S-phase and its use for comparative purposes in cultured human cells. *J Theor Biol* 111: 355-367
- Savage JR (2000) Cancer. Proximity matters. *Science* 290: 62-63
- Silk AD, Holland AJ, Cleveland DW (2009) Requirements for NuMA in maintenance and establishment of mammalian spindle poles. *J Cell Biol* 184: 677-690
- Srebniak MI, Trapp GG, Wawrzekiewicz AK, Kazmierczak W, Wiczkowski AK (2005) The usefulness of calyculin A for cytogenetic prenatal diagnosis. *J Histochem Cytochem* 53: 391-394
- Swedlow JR, Hirano T (2003) The making of the mitotic chromosome: modern insights into classical questions. *Mol Cell* 11: 557-569
- Tanno Y, Kobayashi K, Tatsuka M, Gotoh E, Takakura K (2006) Mitotic arrest caused by an X-ray microbeam in a single cell expressing EGFP-aurora kinase B. *Radiat Prot Dosimetry* 122: 301-306
- Tanno Y, Kitajima TS, Honda T, Ando Y, Ishiguro K, Watanabe Y (2010) Phosphorylation of mammalian Sgo2 by Aurora B recruits PP2A and MCAK to centromeres. *Genes* 24: 2169-2179.
- Tatsumi Y, Sugimoto N, Yugawa T, Narisawa-Saito M, Kiyono T, Fujita M (2006) Deregulation of Cdt1 induces chromosomal damage without rereplication and leads to chromosomal instability. *J Cell Sci* 119: 3128-3140
- Therman E. (1980). *Human chromosomes*. New York: Springer Verlag.
- Thoma F, Koller T, Klug A (1979) Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol* 83: 403-427
- Torres JZ, Ban KH, Jackson PK (2010) A specific form of phospho protein phosphatase 2 regulates anaphase-promoting complex/cyclosome association with spindle poles. *Mol Biol Cell* 21: 897-904
- Tsukahara T, Tanno Y, Watanabe Y (2010) Phosphorylation of the CPC by Cdk1 promotes chromosome bi-orientation. *Nature* 467: 719-723. Epub 2010 Aug 2025.
- Wang ZZ, Li WJ, Zhi DJ, Gao QX, Qu Y, Wang BQ (2009) Prematurely condensed chromosome fragments in human lymphocytes induced by high doses of high-linear-energy-transfer irradiation. *Mutat Res* 679: 9-12
- Wolfe BA, Takaki T, Petronczki M, Glotzer M (2009) Polo-like kinase 1 directs assembly of the HsCyk-4 RhoGAP/Ect2 RhoGEF complex to initiate cleavage furrow formation. *PLoS Biol* 7: e1000110
- Woodcock CL, Frado LL, Rattner JB (1984) The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J Cell Biol* 99: 42-52
- Woodcock CL, Dimitrov S (2001) Higher-order structure of chromatin and chromosomes. *Curr Opin Genet Dev* 11: 130-135
- Yoshizawa-Sugata N, Masai H (2009) Roles of human AND-1 in chromosome transactions in S phase. *J Biol Chem* 284: 20718-20720

# Sequence-Directed DNA Curvature in Replication Origins Segments

Adriana Fiorini, Fabrícia Gimenes, Quirino Alves de Lima Neto,  
Fábio Rogério Rosado and Maria Aparecida Fernandez  
*Universidade Estadual de Maringá – UEM, Maringá, Paraná  
Brazil*

## 1. Introduction

DNA replication is an essential cellular process for the propagation of life. Prokaryotic cells are examples of organisms that display a fast and precise replication process (Robinson & Bell, 2005; Wang & Sugden, 2005). Although the consensus nucleotide sequence in the bacterial replication origin exhibits variation in number and size between different bacterial species (DoriC, a database of the oriC region in bacterial genomes; Gimenes et al., 2008a), it is capable of directing the association of specific proteins in the initiator chromosomal site and thereby the success of the replicative process in eubacteria (reviewed in Mott & Berger, 2007). However, the identification of hundreds to thousands of replication initiation sites in eukaryotes is more complex than locating clearly defined sites with stretches of consensus sequences, as observed in prokaryotes. While complexity exists, some parallels can be drawn between prokaryotic and eukaryotic systems. Less derived eukaryotes, such as the unicellular yeast *Saccharomyces cerevisiae*, exhibit similar replicative processes to prokaryotes. Replication initiation sites have approximately 120-200 base pairs (bp), termed ARS (Autonomously Replicating Sequences). Although there is sequence diversity among the ARS sites, these sites present a short consensus sequence of 10-11 bp that is rich in A/T residues and called the ACS (ARS Consensus Sequence) and a more divergent motif known as domain B (review in Bell & Dutta, 2002). Although not all of ARS sites are efficiently used, it has been shown that they can be initiate replication if necessary (Sharma et al., 2001). However, *S. cerevisiae* and closely related species appear to be the only eukaryotic organisms that present consensus elements in their replication initiation sites. *Schizosaccharomyces pombe* has initiation replication sequences from 600 to 800 bp, which display numerous A/T-rich regions, but these replication initiation sites in *S. pombe* have a low level of similarity to the ARS of *S. cerevisiae*, and we have not been able to identify a consensus sequence within this region (Dubey et al., 1996; Zhu et al., 1994).

The presence of conserved sequences in *S. cerevisiae* was important for the isolation of the origin recognition complex (ORC; Bell & Stillman, 1992). The existence of this complex revealed that eukaryotes depend on the binding of specific proteins to identify and activate specific genomic regions to initiate replication. The ORC complex, which is composed of six closely related proteins, Orc1-6, is conserved through evolution and has been identified in all analyzed eukaryotic cells (reviewed in Bell, 2002). Before replication begins, numerous other

proteins form the pre-replicative complex (pre-RC), which signals permission for the initiation of replication (reviewed in Stillman, 2005). Recently, it was reported that the pre-RC complex is detected in initiation zones with low nucleosome occupancy (Lubelsky et al., 2010). Although the ORC composition is similar in eukaryotes, replication initiation in metazoans is a complex process, and the structure and mechanism of determining the selection of these initiation sites are not fully understood. The nature of the information required for activation of replication origins in multicellular organisms, where multiple initiation sites of replication occur in cells with differential gene expression, has made it difficult to associate function to a specific replication origin sequence. Analyses of replication intermediates using alternative methods indicated that the replication origin lacked a nucleotide consensus sequence (Biamonti et al., 2003; DePamphilis, 1999; Toledo et al., 1998; Tower, 2004). The activity of the initiation site for replication in metazoans, which lacks a consensus sequence, can be influenced by chromatin structure and/or selected by epigenetic factors (Anglana et al., 2003; Balani et al., 2010; Courbet et al., 2008; Debatisse et al., 2004; Fiorini et al., 2006a; Gimenes et al., 2009; Stehle et al., 2003). The presence of foci of replication (replication factories) that bring together numerous sites of replication initiation and the association of the nuclear matrix (S/MAR) demonstrates the importance of chromosome structure in DNA replication process (Anachkova et al., 2005; Courbet et al., 2008; Jackson, 2003; Newport & Yan, 1996). The analysis of DNA replication using microarrays of the complete genome has increased our understanding of this process and has led to the formulation of new questions about selection and activation of replication origins events in metazoans (MacAlpine & Bell, 2005). Thus, although a large amount of experimental data have been obtained in recent years, the question of whether there are sequences and/or structures conserved in the metazoan replication initiation regions has not been answered yet. The existence of important topological features in the replication initiation sites and/or in the segments that flank these sites could be essential for the selection and activation of accurate DNA replication in both prokaryotic and eukaryotic systems. The bacterial DnaA protein interacts with the sequences called DnaA boxes, and although this interaction is highly conserved among different species, the numbers and sequences of these sites are not conserved (Gimenes et al., 2008a; Mackiewicz et al., 2004). How then is DnaA capable of interacting with these sites, even within different regions of the DnaA box? The same question can be formulated for the initiation sites in yeast; although ARS regions are conserved, the elements A, B1 and B2 exhibit alterations in their sequences. Thus, how then is the specificity of the protein components of the ORC complex maintained in the ARS sites? The analysis of the molecular structure of DnaA and the ORC complex in *Saccharomyces cerevisiae* and *Drosophila melanogaster* revealed that these proteins belong to the superfamily of ATPases with AAA<sup>+</sup> domains and exhibit increased affinity when exposed to negatively supercoiled DNA (Clarey et al., 2006; Erzberger et al., 2006; Remus et al., 2004; Speck et al., 2005). The structure of negatively supercoiled DNA in replication initiation regions implies that repression of replication activity at these sites could be due to the loss of this supercoiling. For bacteria, the protein SEQA, which has repressor activity in DNA replication, has been reported to promote changes in the helical structure of DNA, causing positive supercoiling of the segment (Kjellesvik & Skarstad, 2004). In eukaryotes, the implication of the supercoiling changes in DNA replication seems to be more complex. The protein geminin, which regulates DNA replication by interacting with the Cdt1 protein, inhibits replication re-initiation in the same cell cycle (Pitulescu et al., 2005), but these proteins do not show DNA binding domains. The DEK protein, derived from a proto-oncogene,

was reported to modify the topology of replication initiation sites, but has not shown an affinity for DNA without the presence of histones (Alexiadis et al., 2000). Recently, studies have used genome-wide approaches to analyze the features associated with the choice of the replication origins and the nucleosome occupancy within these regions (review in Meisch & Prioleau, 2011).

The correlation between specific chromosomal regions and the topological features of DNA is an exciting area of research. Our laboratory has been devoted to mapping the presence of intrinsically bent DNA sites in regions of replication initiation in prokaryotes (Gimenes et al., 2008a), replication origins and promoter regions in gene amplified domains of eukaryotes (Balani et al., 2010; Fiorini et al., 2001, 2006a; Gimenes et al., 2009; Gouveia et al., 2008) and in recombination regions from the eukaryotes genome (Barbosa et al., 2008). Here we discuss the sequence and topological features of intrinsically bent DNA sites as well the methodology used to explore these sites in replication origin segments.

## 2. Determinants of DNA bending

The DNA molecule is composed of two nucleic acid polymer chains that are wound around one another to form a regular right-handed helix, which is similar to the canonical DNA conformation, called B-DNA (Marko & Cocco, 2003). There are two other well-known forms of DNA, A-DNA and Z-DNA, which are a right-handed helix and a left-handed helix, respectively. It has been found alternative genomic DNA conformations that are different from the canonical B-DNA helix (Potaman & Sinden, 2005; Ohshima, 2005). These alternative DNA structures are very important for certain biological functions, such as the DNA-protein binding involved in replication, gene expression and recombination. Alternative conformations may also participate in the formation of nucleosomes and other supramolecular structures involving DNA (Lu et al., 2003; Meisch & Prioleau, 2011; Richmond & Davey, 2003).

The DNA double helix is a highly dynamic structure; consequently, the curvature of DNA cannot be construed as a static and rigid alteration (Hagerman, 1990). The flexibility of DNA can be described by the alterations in the twist angles between adjacent base pairs, called torsional flexibility, and by deviations along the axis of the double helix from a straight trajectory, characterizing the bending flexibility (Travers, 2004). This bending flexibility depends primarily on the physico-chemical properties of individual base steps and secondarily on the DNA sequence context. The number of hydrogen bonds in a base pair, the stacking energy of a given base step and the occupation of both the major and minor grooves by nitrogenous bases are the major determinants of DNA (Calladine et al., 2004; Travers, 2004).

The bending of DNA is one of the most important deformations of the DNA structure, which is universal in biology, for both storage and retrieval of information encoded in the base-pair sequence (Gimenes et al., 2008b). Many base sequences can conform to a systematic curvature in DNA; however, a series of A<sub>5-6</sub> (A-tracts) repeats spaced in phase with the DNA helical repeat (10-11 bp intervals) are the most important causative sequences for the presence of intrinsically bent DNA sites. This discovery was made by Trifonov & Sussman (1980), who demonstrated the periodicity of A-tracts in genomic DNA and suggested that this observation was due to the DNA packing in chromatin. Particularly, they suggested that the AA/TT dimer has an intrinsic wedge-like shape, which would introduce intrinsic bending in DNA when repeated periodically. The model was called a

wedge model and highlighted that a series of small alterations in the roll or tilt angles between adjacent base-pair planes could be generating the bending of DNA (Figure 1). Alternatively, Wu & Crothers (1984) introduced their junction model, which assumed that A-tracts adopt an alternative non-B-DNA structure, where the cause of curvature is the deflection of the global axis of the A-tract structure from that of the adjoining B-DNA region (Figure 1). The main difference between the wedge and junction models is that the former is a nearest-neighbor dimeric model, assuming that the curvature of any dimeric step is independent of the other adjacent A-tracts, whereas the latter suggests cooperative interactions along the DNA chain. However, other DNA sequences can display an anomalous gel mobility that is characteristic of curved DNA, even in absence of A-tracts. It has been observed that the GGGCCC segment may cause curvature toward the major groove of the DNA helix, with the presence of a positive roll angle (Brukner et al., 1993). Thus, we have observed that DNA bending follows a simple rule: the purine-pyrimidine (RY) and AA/TT dimers bend predominantly into the minor groove, whereas the pyrimidine-purine and GG/CC dimers bend more frequently toward the major groove (Ohyama, 2005). From gel electrophoresis experiments, it was possible to establish three important features of curved DNA (Diekmann, 1986; Hagerman, 1985; Koo & Crothers, 1987): periodically phased A-tracts are very important for “strong” DNA curvature; the orientation of the A-tract is important because the  $A_4T_4$ -induced bending differs from that of  $T_4A_4$ ; and the magnitude of DNA curvature is influenced little by the flanking sequences.

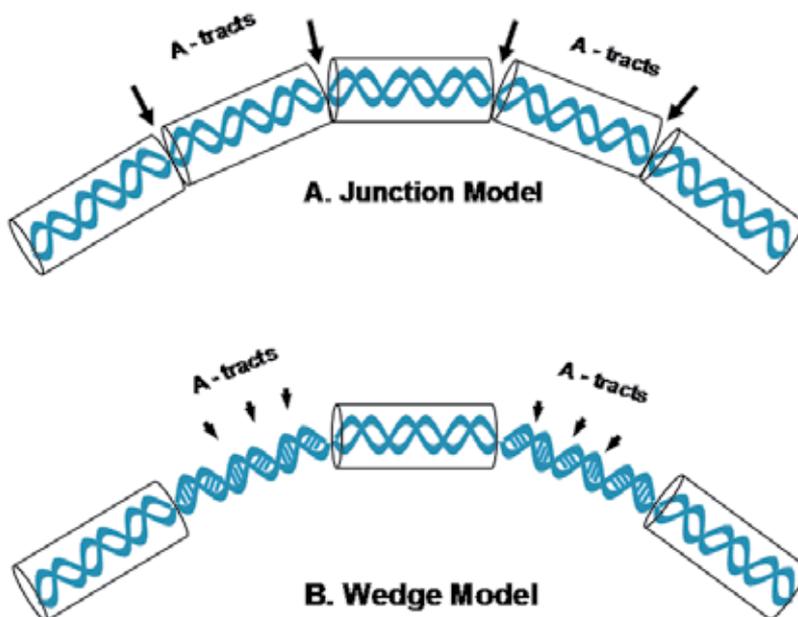


Fig. 1. Junction and Wedge models of an intrinsic DNA curvature.

### 3. Methods of analyzing intrinsically bent DNA

Many methods can be applied to localize intrinsically bent DNA sites, for example, circular permutation (Gimenes et al., 2008b; Wu & Crothers, 1984), computational analysis (Marilley

& Pasero, 1996; Pasero et al., 1993), and atomic force microscopy (Marilley et al., 2007). Here we present methods for analyzing bending DNA regions by computational simulation using algorithms for curvature and gel electrophoresis mobility assays, including the circular permutation strategy.

### 3.1 Curvature parameters calculations

The helical parameters reflect the structural characteristics of the DNA double helix, which was described initially as an ideal isotropic rod, with elastic properties independent of sequence. However, once it became clear that the conformation of DNA is dynamic, the initial isotropic representation was questioned, and the concept of anisotropic DNA was introduced. Therefore, there was a need to standardize the definitions and nomenclature of nucleic acid structure parameters. From the EMBO Workshop on DNA Curvature and Bending, the definitions of parameters used to describe the geometry of nucleic acid chains and helices and a common nomenclature for these parameters were introduced (Dickerson, 1989).

The values of helical parameters such as the ENDS ratio parameter and roll, tilt and twist rotational angles can be obtained using the computational program Map15a (Marilley & Pasero, 1996; Pasero et al., 1993), which was performed employing the algorithm of Eckdahl & Anderson (1987) and the helical parameters described by Bolshoy et al. (1991). The ENDS ratio parameter (the ratio of the axis outline of the helix to the smallest distance between the extremities of the fragment) reflects the probability of finding a bend at a determined site. Figure 2A shows an example of this application for a hypothetical segment of 800 bp, using a 120-bp window width and a 10-bp step. Values greater than or equal to 1.10 is indicative of the presence of intrinsically bent DNA sites (Milot et al., 1992).

The twist angle corresponds to a rotation around the local twist axis that runs vertically through, or near, the centers of any two neighboring base pairs. In the canonical B-DNA conformation, this angle is estimated to be approximately 32° and is shown in Figure 2B (Calladine et al., 2004; Dickerson, 1989). The roll angle (Figure 2B), described the rolling open of base pairs along their long axes. This movement compresses the major (positive roll) or minor (negative roll) grooves of the DNA helix (Calladine et al., 2004; Crothers et al., 1990). The tilt angle (Figure 2B) which is in the direction of hydrogen bonding (opening or closing towards the phosphate backbone) is the bending across the grooves, in which the rotational displacement causes a deviation of the DNA double helix (Hagerman, 1990; Ohyama, 2005; Travers, 2004).

### 3.2 2D Modeling

The 2D projections of the 3D trajectories from specific fragments can be obtained from computational programs such as 3D15m1 (Marilley & Pasero, 1996; Pasero et al., 1993). As described for the Map15a program, the algorithm for calculating a 3D trajectory from a nucleotide sequence was also developed by Eckdahl & Anderson (1987), and the three-dimensional coordinates of the helical axis are obtained using the parameters of the wedge model from Bolshoy et al. (1991). While using the program 3D15m1, it is possible to rotate the molecule in space in real time, and the user can obtain a 2D projection of the analyzed fragment. We can also use the server model.it, which creates 3D models of canonical or bent DNA starting from sequence data and presents the results in the form of a standard PDB file that can be viewed directly using programs such as Swiss-PDBviewer or RasMol

(Vlahovicek et al., 2003). Example 3D models of hypothetical sequences created by the model.it server are shown in Figure 3.

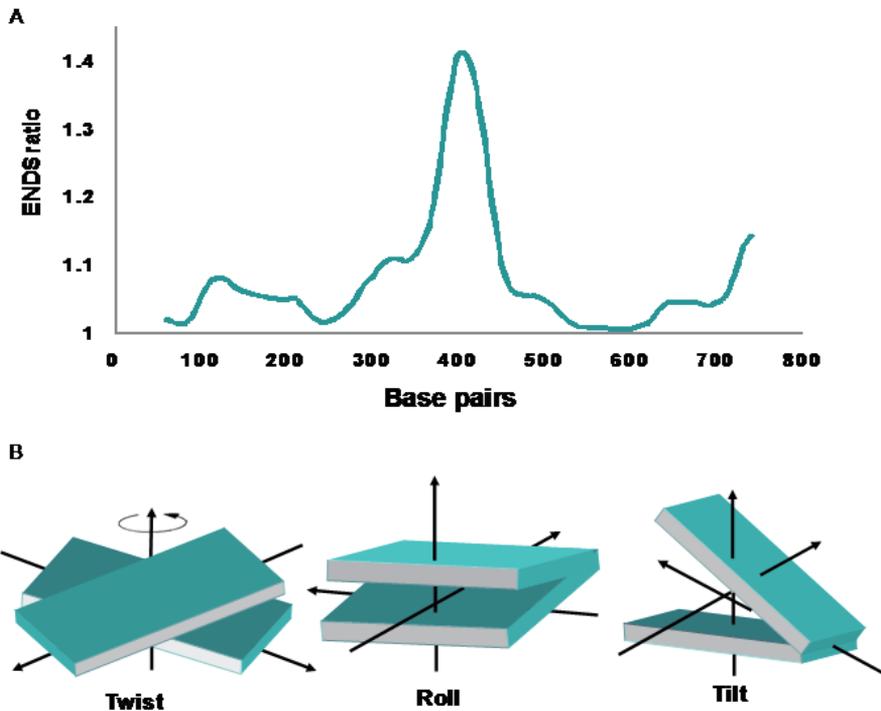


Fig. 2. Helical curvature parameters. **A.** ENDS ratio calculation graphic using the Map15a software. **B.** Rotational angles of DNA double helix, adapted from Dickerson (1989).

### 3.3 Experimental approaches

#### 3.3.1 Isolation and cloning of intrinsically bent fragments

After mapping intrinsically bent DNA sites by computational analysis, regions containing bent DNA can be isolated using polymerase chain reaction (PCR). Nucleotide sequences of approximately 100 bp containing DNA bent sites are amplified from a DNA sample (approximately 50 ng/ $\mu$ l), and the primers used to amplify the DNA bent sites are designed with the restriction sites of *Xba*I and *Sal*I at the ends for posterior cloning into a circular permutation vector such as pBendBlue (Sperbeck & Wistow, 1998). The PCR amplified fragments can be cloned into any PCR cloning vector such as the pGEM-T Easy Vector System (Promega), pMOSBLUE (GE Healthcare), pTZ57R/T (Balani et al., 2010; Gimenes et al., 2009; Hägg et al., 2004; Rodríguez-Lecompte et al., 2001) or the TOPO® PCR Cloning vector (Invitrogen). The recombinant plasmids may be sent for sequencing to confirm the identity of the insert. Helpful aids for primer design include commercially available primer design software such as Oligo<sup>®</sup> (National Biosciences, Plymouth, NC) and FastPCR© (4.0.27 versus) and online search web tools such as BLAST (NCBI, [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)), IDT SciTools software OligoAnalyzer 3.0 and PrimerQuest.

SEQUENCE	NUCLEOTIDE SEQUENCE	SOURCE
A	AGGGCCCTAGAGGGGCCCTAG	Brukner et al.; 1993
B	CGGAGGGGCCCTAGAGGGGCCCTAGAGGGC CCCCAAAAACCCCCAAAAACCCCC	Brukner et al.; 1993
C	TCTCTAAAAAATATATAAAAA	Bolshoy et al.; 1991
D	AAAAAAAAAAAAAAAAAAAAA	Bolshoy et al.; 1991
E	CATGTCACCGACGCATCACCG	Bolshoy, et al.; 1991
F	ATACACAGAAAATAGAAATGTCCTTAAAT TTTTATATTTTTCACACTTTAAAGT	Balani et al.; 2010

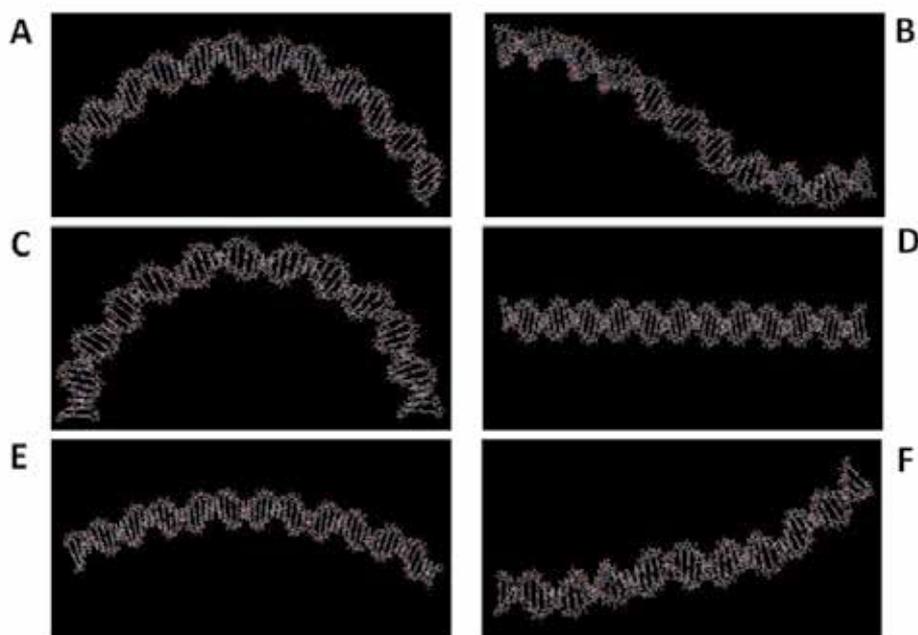


Fig. 3. A- F, 2D projection of the 3D trajectories of the DNA nucleotides sequences showed in the upper table.

### 3.3.2 Analysis of bending DNA fragments by gel electrophoresis

DNA fragments containing an intrinsic bent sequence can be analyzed by electrophoresis using a non-denaturing polyacrylamide gel run at a low temperature. This method contributes to the detection of alterations in the mobility of bent fragments.

Fragments isolated for restriction digestion of cloning plasmids or amplified by PCR are subjected to electrophoresis in a polyacrylamide gel. The gel concentration and the electrophoresis conditions vary depending on the fragment size. In general, for fragments of approximately 100 bp, a 12% polyacrylamide gel with 1X TBE running buffer (45 mM Tris-borate and 1 mM EDTA, pH 8.0) and a constant voltage of 5-10 V/cm at 4°C can be used. After running, the gel is stained with 1 µg/ml of ethidium bromide or another DNA intercalating agent and photographed under UV light. After running, the mobility pattern of each fragment can be obtained through the calculation of the R-value (ratio of apparent/real

fragment size) (de Souza & Ornstein, 1998; Fiorini et al., 2006b). The R-values indicate the degree of curvature for a given fragment. An R-value between 0.90 and 1.09 indicates no alterations in mobility, whereas those higher than or equal to 1.10 indicate reduced mobility, and those lower than 0.90 indicate fast mobility (Fiorini et al., 2006b; Milot et al., 1992). Fragments presenting significant R-values are then analyzed by circular permutation to determine the region responsible for the curvature.

### 3.3.3 Circular permutation assay

The circular permutation assay is used to determine whether bent DNA sites promote curvature in a distinct sequence context and to determine the bending-center sites on permuted fragments (Wu & Crothers, 1984).

To perform this assay, fragments containing curvature were chosen, isolated and cloned into the PCR cloning plasmids and were subcloned into the center of the duplicated polylinker of the pBendBlue vector (Sperbeck & Wistow, 1998) or other pBend plasmid such as pBend2 and pBend3 (Kim et al., 1989; Zwieb & Adhya, 1994). The pBendBlue vector contains two identical but inverted DNA segments with 17 repeated restriction sites spanning a central region containing *SalI* and *XbaI* cloning sites (Figure 4). Briefly, the insert is removed from the PCR cloning plasmid by digestion with *SalI* and *XbaI* restriction enzymes and subsequently subcloned between the *SalI* and *XbaI* sites of the pBendBlue plasmid. The generated recombinant pBendBlue plasmids are then introduced into host bacteria by means of DNA transformation and selected by a direct assay. The plasmid DNA is isolated by minipreparation and digested in independent reactions with suitable restriction enzymes, which recognize the duplicated pBendBlue polylinker, yielding a set of circularly permuted DNA fragments of identical size, but with a permuted bent position (Fiorini et al., 2006a; Gimenes et al., 2008b; Gimenes et al., 2009; Sperbeck & Wistow, 1998).



Fig. 4. pBendBlue polylinker.

### 3.3.4 Permuted fragment analysis by electrophoresis

Wu & Crothers (1984) have designed an elegant mobility gel electrophoresis assay to localize the bending locus of an intrinsically curved DNA fragment. De Santis et al. (1988; 1990) and Zuccheri et al. (2001) have proposed a theoretical model for DNA curvature and have shown that curvature dispersion is linearly correlated with gel electrophoretic retardation. The model has been experimentally verified and has been applied in analyzing several systems (Nair, 1998; Nair et al., 1994).

The permuted fragments have conformations that differ in the position of the bend relative to the molecular ends. Because the mobility of the fragments is a function of their intrinsic curvature, a bend near the middle of the molecule should encounter more difficulty traversing through the pores of a polyacrylamide gel than a more linear fragment. Curvature causes a slower migration compared with non-curved sequences of the same length. However, fragments containing intrinsic DNA structures near the ends migrate

faster in polyacrylamide gels (Drak & Crothers, 1991; Fiorini et al., 2001; Schroth et al., 1992) (Figure 5A).

In an experimental approach, the resulting circular permuted fragments are analyzed by electrophoresis mobility shift assay and resolved in a non-denaturing polyacrylamide gel following the electrophoresis conditions above described. After running, the relative mobility of each fragment is calculated by the ratio of the expected mobility according to the molecular weight of the fragment and the apparent mobility observed on the gel. The relative mobility of each permuted fragment is plotted against the distance (in base pairs) from the 5' end of the used restriction sites to the middle of the cloned fragment. Quantitative measurement of the relative gel mobility of a set of circularly permuted fragments allows extrapolating the position that would yield the maximum gel mobility and was therefore located at the center of the molecular bend (Wu & Crothers, 1984). The position at which maximal gel retardation should have occurred is extrapolated from the graph and is indicated by broken lines (Figure 5B).

The permuted fragments can be further analyzed theoretically to further understand their mobility behavior with the 2D path projection software, as described previously.

Transient interactions between intrinsically bent DNA sites and nucleic acids are widespread in nature. Almost all functions performed by DNA and RNA in the cellular context depend on the involvement of several associated proteins. Generally, interactions are divided into three main categories: structural, regulatory and enzymatic. Trials of DNA-protein interactions can be used to study specific regions of nucleic acids, such as curved regions (bent DNA) and linear, potentially flexible regions.

#### 4. Possible roles of intrinsically bent DNA in replication

Statically curved DNA elements are known to be present at many replication origins, and it has been proposed that these sequence elements are important in initiating DNA replication. An analysis of the secondary structure of the replication origins of the prokaryotic organism *Xylella fastidiosa* 9a5c was previously described by our group and indicated that these replication origins display intrinsically bent DNA sites that induce a curvature in this segment (Gimenes et al., 2008a). In the same work, the *in silico* analysis of the replication origin segments from *X. fastidiosa* Temecula, *Bacillus subtilis* and *Escherichia coli* showed that all of the replication sites, with some variability in their helical parameters, displayed curved segments. In eukaryotes cells, our work was associated with gene amplified domains in developmental systems (Fiorini et al., 2001; 2006a; Gimenes et al., 2009) and induced gene amplified segments in mammalian culture cells (Balani et al., 2010). Using the developmental amplified models, we analyzed amplicons from the gene *BhC4-1* from *Bradysia hygida* (Fiorini et al., 2001), an amplified segment of the C3-22 gene from *Rhynchosciara americana* and in the segment DAFC-66D from chromosome 3 of *Drosophila melanogaster*, which contains the amplification control element ACE3 and the replication origin ori- $\beta$  (Gimenes et al., 2009). In the dihydrofolate reductase (DHFR) amplicon, Altman & Fanning (2004) reported four elements that are required to initiate DNA replication at ori- $\beta$ . One of them is an intrinsically bent DNA site, which provides the capacity to replicate the DNA locus ectopically. Taken together, these results indicate a relationship between secondary DNA structure and replication origins. Functional experiments and the relationship between proteins in the curved structure could be enabling us to establish a strong relationship between the DNA structure and the initiation of DNA replication in eukaryotes cells.

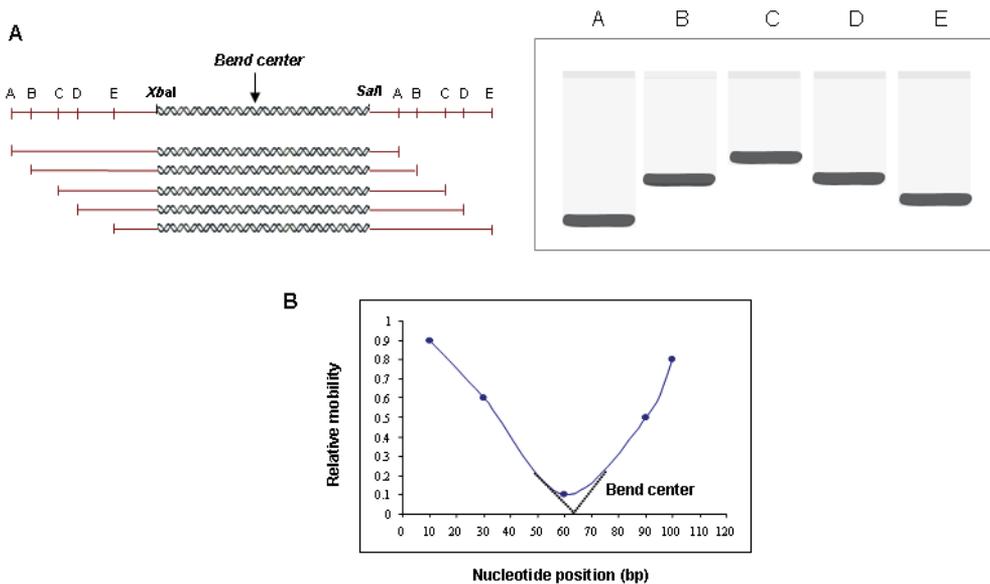


Fig. 5. **A.** Permuted fragments cloned into the duplicated polylinker of a pBend plasmid (on the left) and a schematic analysis of the electrophoresis mobility shift assay of the permuted fragments (on the right). **B.** Relative mobility versus nucleotide position for bent position determination.

## 5. Conclusion

Here we discussed reports in the literature concerning the presence of intrinsically bent DNA sites in regions of replication origins and the possible involvement of these sites in replication function in both prokaryotic and eukaryotic organisms. We presented *in silico* and *in vitro* structural analysis by specific software and electrophoresis mobility assays to detect and analyze intrinsically bent DNA sites in regions of replication initiation. Recent technologies for analyzing the relationship between and the binding of DNA and proteins and nucleosome occupancy in the replication domains could be used in the future in establishing the relationship among DNA secondary structure and sites of replication initiation in specialized cells.

## 6. Acknowledgments

We thank Valmir Peron and Marli Licero Schuete Silva for their dedicated technical assistance. This work was supported by Fundação Araucária, CAPES, CNPq, the Academy of Sciences for the Developing World-TWAS and the Universidade Estadual de Maringá facilities (COMCAP laboratories).

## 7. References

Alexiadis, V.; Waldmann, T; Andersen, J; Mann, M.; Knippers, R. & Gruss, C. (2000). The protein encoded by the proto-oncogene DEK changes the replication in a

- chromatin-specific manner topology of chromatin and reduces the efficiency of DNA. *Genes & Development*, Vol.14, No.11, (June 2000), pp. 1308-1312. ISSN 0890-9369.
- Altman, A. L. & Fanning, E. (2004). Defined sequence modules and an architectural element cooperate to promote initiation at an ectopic mammalian chromosomal replication origin. *Molecular and Cellular Biology*, Vol.24, No.10, (May 2004), pp. 4138-4150.
- Anachkova, B.; Djeliova, V. & Russev, G. (2005). Nuclear matrix support of DNA replication. *Journal of Cellular Biochemistry*, Vol.96, No.5, (December 2003), pp. 951-961. ISSN 0730-2312.
- Anglana, M.; Apiou, F.; Bensimon, A. & Debatisse, M. (2003). Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. *Cell*, Vol.114, No.3, (August 2003), pp. 385-394, ISSN 0092-8674.
- Balani, V. A.; de Lima Neto, Q. A.; Takeda, K. I.; Gimenes, F.; Fiorini, A.; Debatisse, M. & Fernandez, M. A. (2010) Replication origins oriGNAI3 and oriB of the mammalian AMPD2 locus nested in a region of straight DNA flanked by intrinsically bent DNA sites. *Biochemistry and Molecular Biology Reports*, Vol.43, No.11, (November 2010), pp. 744-749. ISSN 1976-6696.
- Barbosa, J. F.; Bravo, J. P.; Takeda, K. I.; Zanatta, D. B.; Silva, J. L. C.; Balani, V. A.; Fiorini, A. & Fernandez, M. A. (2008). Intrinsic bent DNA colocalizes with the sequence involved in the *Nd-s<sup>D</sup>* mutation in the *Bombyx mori* fibroin light chain gene. *Biochemistry and Molecular Biology Reports*, Vol.41, No.5, (May 2008), pp. 394-399. ISSN 1976-6696.
- Bell, S. P. & Dutta, A. (2002). DNA replication in eukaryotic cells. *Annual Review Biochemistry*, Vol.71, (November 2002), pp. 333-374, ISSN 0066-4154.
- Bell, S. P. & Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature*, Vol.357, No.6374, (May 1992), pp. 128-134, ISSN 0028-0836.
- Bell, S. P. (2002). The origin recognition complex: from simple origins to complex functions. *Genes Development*, Vol.16, No.6, (March 2002), pp. 659-672, ISSN 0890-9369.
- Biamonti, G.; Paixao, S.; Montecucco, A.; Peverali, F. A.; Riva, S. & Falaschi, A. (2003). Is DNA sequence sufficient to specify DNA replication origins in metazoan cells? *Chromosome Research*, Vol.11, No.5, (July 2003), pp. 403-412, ISSN 0967-3849.
- Bolshoy, A.; Mcnamara, P.; Harrington, R. E. & Trifonov, E. N. (1991). Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. *Proceedings of the National Academy of Sciences USA*, Vol.88, No.6, (March 1991), pp. 2312-2316. ISSN 0027-8424.
- Brukner, I.; Dlakic, M.; Savic, A.; Susic, S.; Pongor, S. & Suck, D. (1993). Evidence for opposite groove-directed curvature of GGGCC and AAAAA sequence elements. *Nucleic Acids Research*, Vol.21, No. 4, (February 1993), pp. 1025-1029. ISSN 0305-1048.
- Calladine, C. R.; Drew, H. R.; Luisi, B. F. & Travers, A. A. (2004). *Understanding DNA: The molecule and how it works* (3rd ed.) Elsevier Academic Press, ISBN 0-12-1550893. London, UK.
- Clarey, M. G.; Erzberger, J. P.; Grob, P.; Leschziner, A. E.; Berger, J. M.; Eva, N. & Botchan, M. (2006). Nucleotide-dependent conformational changes in the DnaA-like core of

- the origin recognition complex. *Nature Structural & Molecular Biology*, Vol.13, No.8, (August 2006), pp. 684-690. ISSN 1545-9993.
- Courbet, S.; Gay, S.; Arnoult, N.; Wronka, G.; Anglana, M.; Brison, O. & Debatisse, M. (2008). Replication fork movement sets chromatin loop size and origin choice in mammalian cells. *Nature*, Vol.455, No.7212, (September 25), pp. 557-560. ISSN 0028-0836.
- Crothers, D. M.; Haran, T. E. & Nadeau, J. G. (1990). Intrinsically bent DNA. *The Journal of Biological Chemistry*, Vol.265, No.13, (May 1990), pp. 7093-7096. ISSN 0021-9258.
- De Santis, P.; Palleschi, A.; Savino, M. & Scipioni, A. (1988). A theoretical model of DNA curvature. *Biophysical Chemistry*, Vol.32, No.2-3, (December 1988), pp. 305-317. ISSN 0301-4622.
- De Santis, P.; Palleschi, A.; Savino, M. & Scipioni, A. (1990). Validity of the nearest-neighbor approximation in the evaluation of the electrophoretic manifestations of DNA curvature. *Biochemistry*, Vol.29, No.39, (October 1990), pp. 9269-9273. ISSN 0006-2960.
- de Souza, O. N. & Ornstein, R. L. (1998). Inherent DNA curvature and flexibility correlate with TATA box functionality. *Biopolymers*, Vol.46, No.6, (November 1998), pp. 403-415. ISSN 0006-3525.
- Debatisse, M.; Toledo, F. & Anglana, M. (2004). Replication initiation in mammalian cells: changing preferences. *Cell Cycle*, Vol.3, No.1, (January 2004), pp. 19-21, ISSN 1538-4101.
- DePamphilis, M. L. (1999). Replication origins in metazoan chromosomes: fact or fiction? *Bioessays*, Vol.21, No.1, (January 1999), pp. 5-16, ISSN 0265-9247.
- Dickerson, R. E. (1989). Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Research*, Vol.17, No.5, (March 1989), pp. 1797-1803. ISSN 0305-1048.
- Diekmann, S. (1986). Sequence specificity of curved DNA. *FEBS Letters*, Vol.195, No.1-2, (January 1986), pp. 53-56. ISSN 0014-5793.
- Drak, J. & Crothers, D. M. (1991). Helical repeat and chirality effects on DNA gel electrophoretic mobility. *Proceedings of the National Academy of Sciences*, Vol.88, No.8, (April 1991), pp. 3074-3078. ISSN 0027-8424.
- DoriC: a database of oriC region in bacterial genomes. <http://tubic.tju.edu.cn/doriC/>. Accessed March, 2011.
- Dubey, D. D.; Kim, S. M.; Todorov, I. T. & Huberman, J. A. (1996). Large, complex modular structure of a fission yeast DNA replication origin. *Current Biology*, Vol.6, No.4, (April 1996), pp. 467-473, ISSN 0960-9822.
- Eckdahl, T. T. & Anderson, J. N. (1987). Computer modelling of DNA structures involved in chromosome maintenance. *Nucleic Acids Research*, Vol.15, No.20, (October 1987), pp. 8531-8545. ISSN 0305-1048.
- Erzberger, J. P.; Mott, M. I. & Berger, J. M. (2006). Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nature Structural & Molecular Biology*, Vol.13, No.8, (August 2005), pp. 676-683. ISSN 1545-9993.
- Fiorini, A.; Basso Junior, L. R.; Paço-Larson, M. L. & Fernandez, M. A. (2001). Mapping of intrinsic bent DNA sites in the upstream region of DNA puff *BhC4-1* amplified gene. *Journal of Cellular Biochemistry*, Vol.83, No.1, (June 2001), pp. 1-13, ISSN 0733-1959.

- Fiorini, A.; Gouveia, F. S. Soares, M. A. M.; Stocker, A. J.; Ciferri, R. R. & Fernandez, M. A. (2006a). DNA bending in the replication zone of the C3 DNA puff amplicon of *Rhynchosciara americana* (Diptera: Sciaridae). *Molecular Biology Reports*, Vol.33, No.1, (March 2006), pp. 71-82, ISSN 0301-4851.
- Fiorini, A.; Gouveia, F. S. & Fernandez, M. A. (2006b). Scaffold/Matrix Attachment Regions and Intrinsic DNA Curvature. *Biochemistry (Moscow)*, Vol.71, No.5, (May 2006), pp. 481-488, ISSN 0006-2979.
- Gimenes, F.; Fiorini A. & Fernandez, M. A. (2008a). Bent DNA sites in the *Xylella fastidiosa* chromosomal replication origin. *Brazilian Journal of Medical and Biological Research*, Vol.41, No.4, (April 2008), pp. 295-304, ISSN 0100-879X.
- Gimenes, F.; Takeda, K. I.; Gouveia, F. S.; Fiorini A. & Fernandez, M. A. (2008b). Intrinsically bent DNA in replication origins and gene promoters. *Genetics and Molecular Research*, Vol.7, No.2, (June 2008), pp. 549-558, ISSN 1676-5680.
- Gimenes, F.; Assis, M. A.; Fiorini, A.; Mareze, V. A.; Monesi, N. & Fernandez, M. A. (2009). Intrinsically bent DNA sites in the *Drosophila melanogaster* third chromosome amplified domain. *Molecular Genetics and Genomics*, Vol.281, No.5, (May 2009), pp. 539-549. ISSN 1617-4615.
- Gouveia, F. de S.; Gimenes, F.; Fiorini, A. & Fernandez, M. A. (2008). Intrinsic bent DNA sites in the developmentally amplified C3-22 gene promoter of *Rhynchosciara americana* (Diptera: Sciaridae). *Bioscience, Biotechnology, and Biochemistry*, Vol.72, No.5, (May 2008), pp. 1190-1198.
- Hagerman, P. J. (1985). Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. *Biochemistry*, Vol.24, No.25, (December 1985), pp. 7033-7037. ISSN 0006-2960.
- Hagerman, P. J. Sequence-directed curvature of DNA. (1990). *The Annual Review of Biochemistry*, Vol.59, pp. 755-781. ISSN 0066-4154.
- Hägg, P.; de Pohl, J. W.; Abdulkarim, F. & Isaksson L. A. (2004). A host/plasmid system that is not dependent on antibiotics and antibiotic resistance genes for stable plasmid maintenance in *Escherichia coli*. *Journal of Biotechnology*, Vol.111, No.1, (July 2004), pp. 17-30. ISSN 0168-1656.
- Jackson, D. A. (2003). The principles of nuclear structure. *Chromosome Research*, Vol.11, No.5, pp.387-401.
- Kim, J.; Zwieb, C.; Wu, C. & Adhya, S. (1989). Bending of DNA by gene-regulatory proteins: construction and use of a DNA bending vector. *Gene*, Vol.85, No.1, (December 1989), pp. 15-23. ISSN 0378-1119.
- Kjellesvik, H. & Skarstad, K. (2004). Positive supercoiling is generated in the presence of *Escherichia coli* SeqA protein. *Molecular Microbiology*, Vol.54, No.1, (October 2004), pp. 123-131. ISSN 0950-382X.
- Koo, H. S. & Crothers, D. M. (1987). Chemical determinants of DNA bending at adenine-thymine tracts. *Biochemistry*, Vol.26, No.12, (June 1987), pp. 3745-3748. ISSN 0006-2960.
- Lu, Y.; Weers, B. D.; Stellwagen, N. C. (2003). Analysis of the intrinsic bend in the M13 origin of replication by atomic force microscopy. *Biophysical Journal*, Vol.85, No.1, (July 2003), pp. 409-415. ISSN 0006-3495.
- Lubelsky, Y.; Sasaki, T.; Kuipers, M. A.; Lucas, I.; Le Beau, M. M.; Carignon, S.; Debatisse, M.; Prinz, J. A.; Dennis, J. H. & Gilbert, D.M. (2010). Pre-replication complex

- proteins assemble at regions of low nucleosome occupancy within the Chinese hamster dihydrofolate reductase initiation zone. *Nucleic Acids Research*, (December 9), [Epub ahead of print], ISSN 0305-1048.
- MacAlpine, D.M. & Bell, S. P. (2005). A genomic view of eukaryotic DNA replication. *Chromosome Research*, Vol.13, No.3, pp. 309-326. ISSN 0967-3849.
- Mackiewicz, P.; Zakrzewska-Czerwinska, J.; Zawilak, A.; Dudek, M. R. & Cebrat, S. (2004). Where does bacterial replication start? Rules for predicting the oriC region. *Nucleic Acids Research*, Vol.32, No.13, (July 2004), pp. 3781-3791, ISSN 0305-1048.
- Marilley, M. & Pasero, P. (1996). Common DNA structural features exhibited by eukaryotic ribosomal gene promoters. *Nucleic Acids Research*, Vol.24, No.12, (June 1996), pp. 2204-2011. ISSN 0305-1048.
- Marilley, M.; Milani, P.; Thimonier, J.; Rocca-Serra, J. & Baldacci, G. (2007). Atomic force microscopy of DNA in solution and DNA modelling show that structural properties specify the eukaryotic replication initiation site. *Nucleic Acids Research*, Vol.35, No.20, (October 2007), pp. 6832-6845. ISSN 0305-1048.
- Marko, J. F. & Cocco, S. (2003). The Micromechanics of DNA. *Physics World*, Vol.16, No.3, (March 2003), pp. 37. ISSN: 0953-8585.
- Meisch, F. & Prioleau, M. N. (2011). Genomic approaches to the initiation of DNA replication and chromatin structure reveal a complex relationship. *Briefings in Functional Genomics*, Vol.10, No.1, (January 2011), pp. 30-36. ISSN 2041-2649.
- Milot, E.; Belmaaza, A.; Wallenburg, J. C.; Gusew, N.; Bradley, W. E. & Chartrand, P. (1992). Chromosomal illegitimate recombination in mammalian cells is associated with intrinsically bent DNA elements. *EMBO Journal*, Vol.11, No.13, (December 1992), pp. 5063-5070. ISSN 0261-4189.
- Mott, M. L. & Berger, J. M. (2007). DNA replication initiation: mechanisms and regulation in bacteria. *Nature Reviews Microbiology*, Vol.5, No.5, (May 2007), pp. 343-354, ISSN 1740-1526.
- Nair, T. M. (1998). Evidence for intrinsic DNA bends within the human cdc2 promoter. *FEBS Letters*, Vol.422, No.1, (January 1998), pp. 94-98. ISSN 0014-5793.
- Nair, T. M.; Madhusudan, K.; Nagaraja, V.; Kulkarni, B. D.; Majumdar, H. K. & Singh, R. (1994). On the mobility behavior of a curved DNA fragment located in circular permutation. *FEBS Letters*, Vol.351, No.3, (September 1994), pp. 321-324. ISSN 0014-5793.
- Newport, J. & Yan, H. (1996). Organization of DNA into foci during replication. *Current Opinion in Cell Biology*, Vol.8, No.3, (June 1996), pp. 365-368. ISSN 0955-0674.
- Ohyama, T. (2005). *DNA Conformation and Transcription* (Ed), Springer Science+Business Media, Inc., ISBN 0-387-25579-6, New York, U.S.A.
- Pasero, P.; Sjakste, N.; iBlettry, C.; Got, C. & Marilley, M. (1993). Long-range organization and sequence-directed curvature of *Xenopus laevis* satellite 1 DNA. *Nucleic Acids Research*, Vol.21, No.20, (October 1993), pp. 4703-4710. ISSN 0305-1048.
- Pitulescu, M.; Kessel, M. & Luo, L. (2005). The regulation of embryonic patterning and DNA replication by geminin. *Cellular and Molecular Life Sciences*, Vol.62, No.13, (July 2005), pp. 1425-1433. ISSN 1420-682X.
- Potaman, V. N. & Sinden, R. R. (2005). DNA: Alternative Conformations and Biology, In: *DNA Conformation and Transcription*, T. Ohyama, (Ed), 3-13, Springer Science+Business Media, Inc. ISBN 0-387-25579-6, New York, U.S.A.

- Remus, D.; Beall, E. & Botchan, M. R. (2004). DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC-DNA binding. *The EMBO Journal*, Vol.23, No.4, (February 2004), pp. 897-907. ISSN 0261-4189.
- Richmond, T. J. & Davey, C. A. (2003). The structure of DNA in the nucleosome core. *Nature*, Vol.423, No.6936, (May 2003), pp. 145-150. ISSN 0028-0836.
- Robinson, N. P. & Bell, S. D. (2005). Origins of DNA replication in the three domains of life. *The FEBS Journal*, Vol.272, No.15, (August 2005), pp. 3757-366, ISSN 1742-464X.
- Rodríguez-Lecompte, J. C.; Nagaraja, M. N. & Kibenge, F. S. B. (2001). A simple method for screening bacterial colonies for mutagenized sites in plasmid DNA. *The Journal of Virological Methods*, Vol.99, No.1-2, (January 2002), pp. 93-98. ISSN 0166-0934.
- Schroth, G. P.; Siino, J. S.; Cooney, C. A.; Th'ng, J. P.; Ho, P. S. & Bradbury, E. M. (1992). Intrinsically bent DNA flanks both sides of an RNA polymerase I transcription start site. Both regions display novel electrophoretic mobility. *Journal of Cellular Biochemistry*, Vol.267, No.14, (May 1992), pp. 9958-9964.
- Sharma, K.; Weinberger, M. & Huberman, J. A. (2001). Roles for internal and flanking sequences in regulating the activity of mating-type-silencer-associated replication origins in *Saccharomyces cerevisiae*. *Genetics*, Vol.159, No.1, (September 2001), pp. 35-45, ISSN 0016-6731.
- Speck, C.; Chen, Z.; Li, H. & Stillman, B. (2005). ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nature Structural & Molecular Biology*. Vol.12, No.11, (November 2005) pp. 965 - 971. ISSN 1545-9993.
- Sperbeck, S. J. & Wistow, G. J. (1998). pBendBlue: modification of the pBend system for color selectability. *Biotechniques*, Vol.24, No.1, (January 1998), pp. 66-68. ISSN 0736-620.
- Stehle, I. M.; Scinteie, M. F.; Baiker, A.; Jenke, A. C. & Lipps, H. J. (2003). Exploiting a minimal system to study the epigenetic control of DNA replication: the interplay between transcription and replication. *Chromosome Research*, Vol.11, No.5, (July 2003), pp. 413-421, ISSN 0967-3849.
- Stillman, B. (2005). Origin recognition and the chromosome cycle. *FEBS Letters*, Vol.579, No.4, (February 2005) pp. 877-884, ISSN 0014-5793.
- Toledo, F.; Baron, B.; Fernandez, M. A.; Lachages, A. M.; Mayau, V.; Buttin, G. & Debatisse, M. (1998). *oriGNAI3*: a narrow zone of preferential replication initiation in mammalian cells identified by 2D gel and competitive PCR replicon mapping techniques. *Nucleic Acids Research*, Vol.26, No.10, (May 1998), pp. 2313-2321, ISBN 0305-1048.
- Tower, J. (2004). Developmental gene amplification and origin regulation. *Annual Review of Genetics*, Vol. 38, (June 2004), pp. 273-304, ISSN 0066-4197.
- Travers, A. A. (2004). *The structural basis of DNA flexibility*. Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, Vol.362, No.1820, (July 2004), pp. 1423-1438. ISSN 1364-503X.
- Trifonov, E. N. & Sussman, J. I. (1980). The pitch of chromatin DNA is reflected in its nucleotide sequence. *Proceedings of the National Academy of Sciences USA*, Vol.77, No.7, (July 1980), pp. 3816-3820. ISSN 0027-8424.
- Wang, J. & Sugden, B. (2005). Origins of bidirectional replication of Epstein-Barr virus: models for understanding mammalian origins of DNA synthesis. *Journal of Cellular Biochemistry*, Vol.94, No.2, (February 2005), pp. 247-256, ISSN 0733-1959.

- Vlahovicek K, Kaján L, Pongor S. (2003). DNA analysis servers: plot.it, bend.it, model.it and IS. *Nucleic Acids Research*, Vol.31, No.13, (July 2003), pp. 3686-3687. ISSN 0305-1048.
- Wu, H. M. & Crothers, D. M. (1984). The locus of sequence-directed and protein-induced DNA bending. *Nature*, Vol.308, No.5959, (April 1984), pp. 509-513. ISSN 0028-0836.
- Zhu, J.; Carlson, D. L.; Dubey, D. D.; Sharma, K. & Huberman, J. A. (1994). Comparison of the two major ARS elements of the *ura4* replication origin region with other ARS elements in the fission yeast, *Schizosaccharomyces pombe*. *Chromosoma*, Vol.103, No.6, (October 1994), pp. 414-422, ISSN 0009-5915.
- Zuccheri, G.; Scipioni, A.; Cavaliere, V.; Gargiulo, G.; De Santis, P. & Samori, B. (2001). Mapping the intrinsic curvature and flexibility along the DNA chain. *Proceedings of the National Academy of Sciences*, Vol.98, No.6, (March 2001), pp. 3074-3079. ISSN 0027-8424.
- Zwieb, C. & Adhya, S. (1994). Improved plasmid vectors for the analysis of protein-induced DNA bending. In: *Methods in Molecular Biology*, G.G. Kneale (Ed.), Vol.30. Humana Press, Totowa. ISSN 1064-3745.

# Initiation of DNA Replication from Closed Circular DNA

Daniel Simmons  
*University of Delaware*  
USA

## 1. Introduction

The replication of SV40 DNA and that of related viruses (polyoma, papilloma) have been studied for several decades and we know a great deal about the molecules involved in this process. However, there are two major gaps in our understanding of viral DNA replication. First, in spite of a significant amount of structural information on the viral initiator proteins, it's not known exactly how the double stranded DNA is unwound at the origin. Second, although we have known for many years which cellular proteins participate in this reaction and what they do biochemically, we know next to nothing about how they work together as a machine to initiate the replication of the DNA, elongate initiated chains and terminate the process by making two covalently closed daughter circles. The intent of this review is to consolidate the available information into working models of DNA unwinding and replication.

## 2. The viruses and virus DNAs

The viruses in the polyoma and papilloma groups are unique in their mode of DNA replication. However, these viruses serve as important models of how a circular DS DNA is unwound and copied. The usefulness of these models is brought out by the fact that only one or two viral proteins are involved in the DNA replication process; all other replication factors come from the cell. It is therefore recognized that whatever function these cell factors have in replicating viral DNA, it is likely that the same function is utilized by the host to replicate its own DNA. These cellular proteins associate with each other and with the viral initiator protein to make functional complexes, nanomachines in fact, that most likely also form during the replication of cellular DNA. Importantly, these virus replication systems have the advantages of being relatively simple, are easy to set up in vitro with purified components and can work from a well characterized and easily recognized origin of replication. These features are lacking in cellular replication systems. The viral assays have allowed us to generate complex models of the interactions of a number of proteins with one another and with the DNA to gain insights into mechanisms of replication. Since the structures of many of these cellular proteins have been solved by X-ray crystallography, the models are approaching an atomic understanding of events prior to and during initiation of replication. A historical review of the SV40 DNA replication field and the many advantages of this system are described by Fanning and Zhao (Fanning and Zhao 2009).

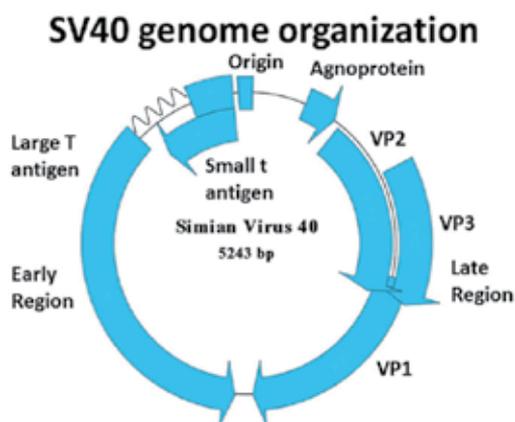


Fig. 1. Genome of SV40. The early genes are shown on the left and the late genes on the right. The origin of DNA replication is at the top.

The organization of genes in a typical polyomavirus (SV40) is shown in Fig. 1. The genome is divided into two parts: an early region and a late region (Simmons 2000). These regions are separated by the regulatory elements that include the origin of replication. Other sequences in the regulatory region are the early and late promoters, an enhancer, and a site (site I) used for repression of the early genes. The minimal origin is 64 bp long and is, by itself, sufficient for T antigen-driven DNA replication *in vitro*. The early genes are expressed shortly after infection. The early mRNAs are spliced differently to generate two main forms of T antigen, large and small, although numerous other forms can also be made. At about 8 hr PI, sufficient quantities of T antigen accumulate to allow DNA replication to begin. At about the same time, late transcription starts and gives rise to the agnoprotein, probably necessary for virus spread, and the three structural proteins VP1, VP2 and VP3. Early transcription continues late in infection but is down regulated when large T antigen binds site I near the origin.

### 3. The origin of replication

The origin of replication is tripartite. In the center, also called site II (Fig. 2), are four GAGGC pentanucleotides oriented as shown in Fig. 2 (Deb, DeLucia et al. 1986). Each pentanucleotide is a binding site for one T antigen monomer. Binding to one pentanucleotide is inefficient but the presence of multiple pentanucleotides permits cooperative interactions between T antigen monomers and together, site II constitutes a very sequence-specific strong binding site (Fanning 1992; Fanning and Knippers 1992; Bullock 1997; Simmons 2000; Fradet-Turcotte, Vincent et al. 2007). On the "early" side of site II is a region known as the early palindrome (EP) or inverted palindrome (IP). In fact, this region is not much of a palindrome, but it is the region where melting of the DNA takes place early on (Borowiec and Hurwitz 1988). On the "late" side of site II, there is a 17 bp AT rich stretch called the AT track. Structural changes take place there as well (Dean and Hurwitz 1991). When a monomer T antigen binds the origin, its origin binding domain (OBD) sits on a pentanucleotide and its helicase domain sits on the EP region or on the AT track depending on which pentanucleotide is bound (see Fig. 6) (Kim, Barbaro et al. 1999; Sreekumar, Prack et al. 2000; Reese, Sreekumar et al. 2004).

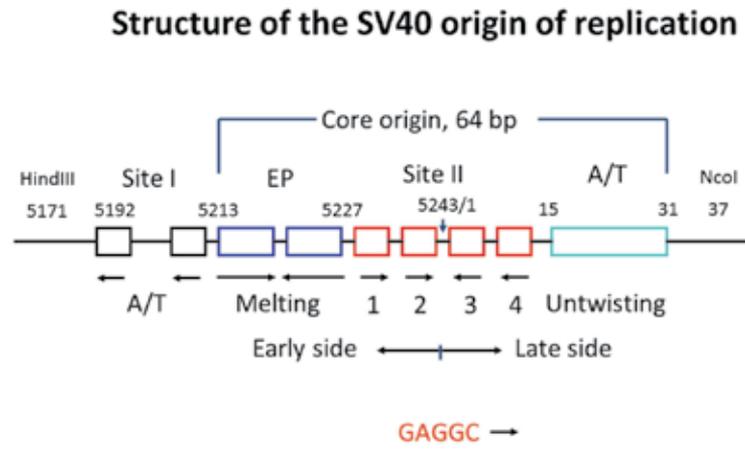


Fig. 2. The SV40 origin of replication. The minimal (core) origin is 64 bp long (Deb, DeLucia et al. 1986) and consists of three regions as shown. Auxiliary sequences that enhance replication, especially *in vivo*, are on the left and right of the core origin.

#### 4. Initiator proteins

T antigen encoded by the polyomaviruses or the E1/E2 proteins encoded by the papillomaviruses are the only proteins needed from the virus for DNA replication. These proteins are multifunctional and have activities dealing with virus replication and cell transformation/immortalization. In this review, I will concentrate only on the SV40 T antigen system. A domain map of this protein is shown in Fig. 3. Three regions of the protein are important for virus and DNA replication: the J domain and associated linker region (residues 1-131), the origin binding domain (residues 131-246) and the helicase domain (residues 247-627). The C-terminal region (residues 628-708) has host type specific functions but is not directly needed for DNA replication (Simmons 2000).

The structure of the J domain in association with a fragment of the retinoblastoma protein (Rb) has been determined (Kim, Ahn et al. 2001). This region of the protein associates with DNA K-like proteins such as HSP70 and this association is used to functionally alter proteins such as Rb (Srinivasan, McClellan et al. 1997; Stubdal, Zalvide et al. 1997; Zalvide, Stubdal et al. 1998; Sullivan and Pipas 2002). The function of this domain in virus DNA replication is not clearly understood. In SV40, the J domain appears to be needed to support DNA replication *in vivo* but not *in vitro* (Sawai, Rasmussen et al. 1994; Weisshart, Bradley et al. 1996; Campbell, Mullane et al. 1997). Fragments of T antigen (such as 110-708) missing the J domain support replication *in vitro* just fine (Weisshart, Bradley et al. 1996; Simmons, Roy et al. 1998). However, mutants with changes in the J domain are replication defective in monkey cells (Rutila, Imperiale et al. 1986). Binding of this region to cellular proteins, including RPA and topo I (see below), may be more important *in vivo*.

The origin binding domain of T antigen has a critical role in initiating the replication process. That domain interacts with the GAGGC pentanucleotides within site II in a sequence specific way. The structure of the domain has been determined by NMR and X-ray crystallography (Luo, Sanford et al. 1996; Bochkareva, Martynowski et al. 2006; Meinke, Bullock et al. 2006; Meinke, Phelan et al. 2007). It has a spheroidal structure with two loops

on one side (called A1 and B2) that make sequence specific contacts with at least 4 of the base pairs at each pentanucleotide site. This domain can bind single-stranded DNA in addition to double-stranded origin DNA (Reese, Meinke et al. 2006; Fradet-Turcotte, Vincent et al. 2007; Foster and Simmons 2010; Meinke, Phelan et al. 2010). It also has the ability to hexamerize into a lock-washer structure where there is a gap between the two terminal domains (Fig. 4A) (Meinke, Bullock et al. 2006). It is reasonable to propose that the protein engages the origin as a monomer or dimer attached to each of the four pentanucleotide sites since all four pentanucleotides are needed for DNA replication (Deb, Tsui et al. 1987). Subsequent events (see below) then occur to convert this domain into the lock washer hexamer where it functions during DNA unwinding.

## T antigen's structural and functional domains

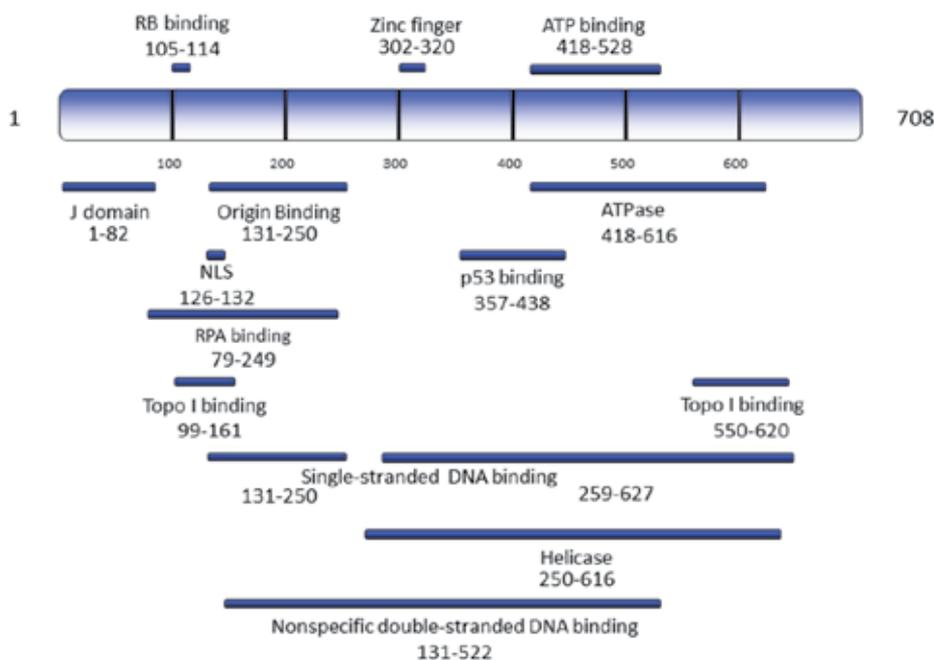


Fig. 3. Structural and functional domains of T antigen. The numbers correspond to the amino acid sequences responsible for each activity.

The helicase domain is made up of three subdomains called D1, D2 and D3 (Li, Zhao et al. 2003). D1:D1 and D2:D2 interactions are critical for the helicase to hexamerize into a functional helicase. When assembled as such, a center channel is created which is charged and accommodates DNA (Fig. 4B). The helicase domain also contains an ATPase domain. Various forms of the T antigen helicase domain have been crystallized including a nucleotide free form, one containing ADP and one with ATP (Gai, Zhao et al. 2004). These three forms are somewhat different from each other in overall dimensions, size of the inner channel and relationship between monomers. In the nucleotide free form, the channel is barely large enough to accommodate double stranded DNA but the other two forms are too small and only single stranded DNA can fit. The latter two are similar in this respect to a

structure of the BPV E1 helicase crystallized in association with single stranded DNA (Enemark and Joshua-Tor 2006). The E1 structure shows that the central channel interacts with single-stranded DNA through 6  $\beta$ -hairpin loops making contacts with 6 successive nucleotides in the DNA (Fig. 5). The proposed mechanism of helicase action from this structure is that the terminal  $\beta$ -hairpin releases its nucleotide and through an allosteric change associated with ATP hydrolysis, the hairpin extends and grabs a nucleotide 6 residues away. All the other  $\beta$ -hairpins move slightly but remain associated with their nucleotide. In this way, the DNA is pulled into the hexameric channel.

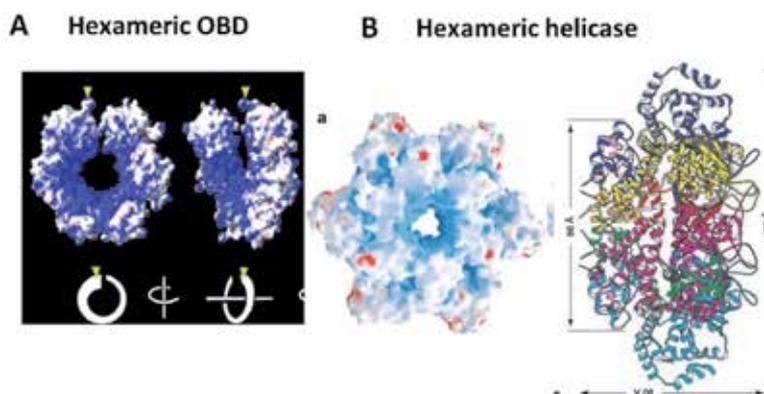


Fig. 4. The origin binding domain and helicase domain assemble into functional hexamers. (A) The crystal structures of hexameric SV40 OBD (Meinke, Bullock et al. 2006). (B) Front and side views of the hexameric helicase domain (Li, Zhao et al. 2003).

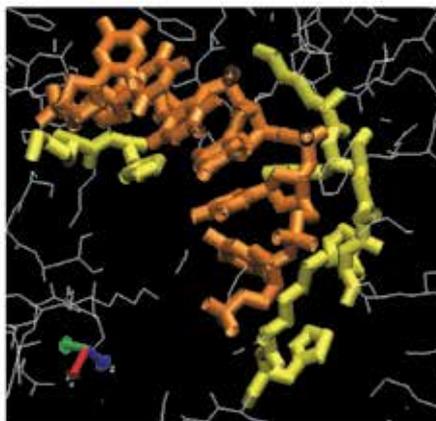


Fig. 5. Structure of the central channel of BPV E1 associated with single-stranded DNA. 12  $\beta$ -hairpin residues highlighted in yellow (K506 and H507 on each monomer) are attached to the backbone of 6 successive nucleotides of ssDNA (orange) lying within the E1 helicase central channel. [Complete structure is from Enemark and Joshua-Tor (Enemark and Joshua-Tor 2006)]. Most of the other residues have been cut away for clarity. Image generated by VMD program (Humphrey, Dalke et al. 1996).

The mechanism of helicase action is likely to be essentially the same for T antigen. In this case, interactions with the phosphates as well as hydrophobic packing interactions appear to be utilized to attach each nucleotide (Shen, Gai et al. 2005).

The structure of the complete T antigen is not known, but models can be generated from the crystal structures of the three solved domains. These models should be consistent with the imaging and biochemical data. In particular, it's known that in the presence of a nucleotide, T antigen assembles into a double hexamer that covers about 65 to 70 bp of DNA (Mastrangelo, Hough et al. 1989; Gomez-Lorenzo, Valle et al. 2003; Reese, Sreekumar et al. 2004; Cuesta, Nunez-Ramirez et al. 2010). DNase protection data show that the core origin and flanking sequences are protected by the protein (Tjian 1978; Han and Hurley 1996). EM imaging shows a similar relationship (Gomez-Lorenzo, Valle et al. 2003; Valle, Chen et al. 2006; Cuesta, Nunez-Ramirez et al. 2010). The other consideration is that the helicase domain should be able to access the EP and AT regions, and at least initially, the OBDs should be in contact with the recognition GAGGC sequences within site II (Kim, Barbaro et al. 1999; Sreekumar, Prack et al. 2000; Reese, Sreekumar et al. 2004). Nevertheless, several models can be drawn up and remain in agreement with most of the evidence. The positioning of the J domain has been determined by cryo EM by comparing images with or without this region (Cuesta, Nunez-Ramirez et al. 2010). It seems to lie on the surface of the OBDs where it might be able to interact with cellular proteins. In fact, both RPA and topo I make contacts there or in the neighboring linker region (see below).

## 5. Basic replication mechanism

Replication is divided into three phases: initiation, elongation and termination. Each one of those steps depends on a particular set of cellular proteins in addition to T antigen. Initiation requires replication protein A (RPA), DNA polymerase alpha/primase (pol/prim) and topoisomerase I (topo I). Elongation requires, in addition, replication factor C (RFC), PCNA and a processive polymerase (most likely pol delta). Termination needs topoisomerase II. Other factors (RNase H and ligase) are needed to remove the RNA primers and seal the resulting nicks.

## 6. Recognition and unwinding of the origin

In order to replicate the DNA, it must first be unwound. The resulting single strands then act as templates for DNA synthesis. One of the major obstacles in the field is to understand how sequence specific unwinding takes place from circular double stranded DNA. To appreciate the nature of the problem, we first need to consider how T antigen engages the origin. Since there are four pentanucleotide sequences in site II, it makes sense to postulate that all four sites become occupied. Extensive mutagenesis of this region has shown that all four pentanucleotides are not necessary for the formation of a double hexamer over the origin. Two of them (pentas 1 and 3 or 2 and 4) will suffice (Joo, Kim et al. 1998; Sreekumar, Prack et al. 2000), although pentas 1 and 3, which constitute one assembly site, are more effective than pentas 2 and 4 in permitting double hexamers to form. However, unwinding of the DNA (Dean, Borowiec et al. 1987) and its subsequent replication (Deb, Tsui et al. 1987) are both completely dependent on all four pentanucleotides. Therefore, all four pentanucleotides must be engaged with T antigen in order for unwinding and replication to occur. It is unclear if only two pentanucleotides are used initially to allow the double

hexamer to form and the other two (2 and 4) get involved at a later stage in preparation for unwinding. It might be difficult for two T antigen monomers to make sequence specific contacts with the DNA when they are present within a double hexamer since the origin binding domain is known to change orientation when it hexamerizes (Meinke, Bullock et al. 2006). It is somewhat more appealing to assume that all four sites are occupied at the outset (Fig. 6). These four monomers then quickly and effectively recruit additional monomers which become rearranged to form two hexamers. There is some evidence that the hexamer over the EP region (early side of origin) forms first (Parsons, Stenger et al. 1991), followed by the second hexamer on the late side. However, kinetic studies (Junfang Jiao, unpublished results) have demonstrated that the double hexamer forms very quickly (within one minute).

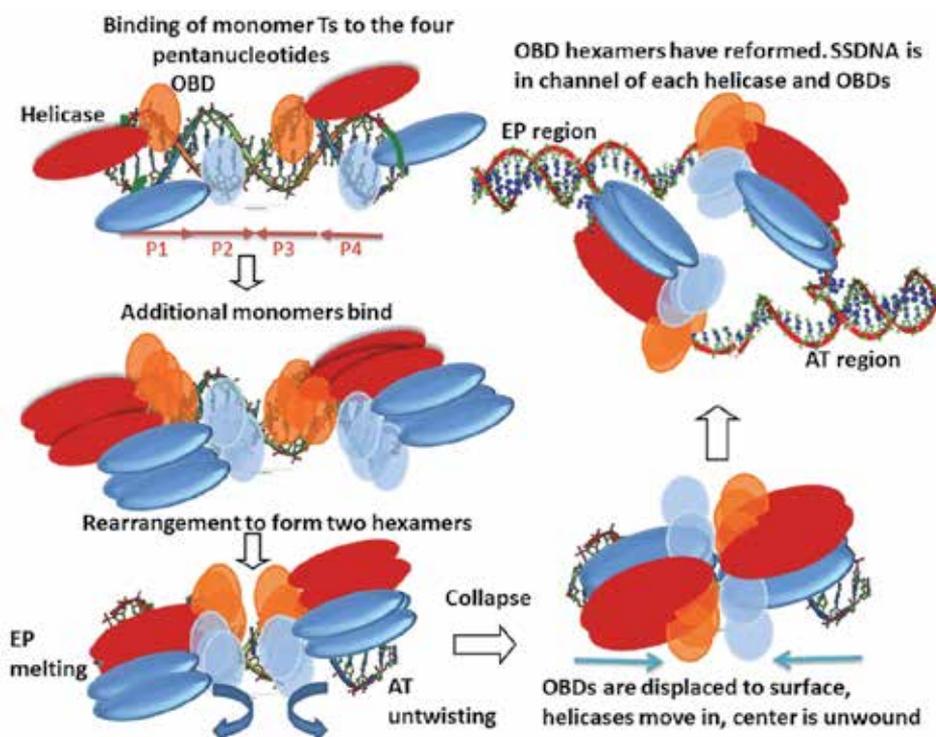


Fig. 6. Proposed model of DNA unwinding from the origin. The model is based on multiple data. Monomers attach first to the pentanucleotides at the origin, additional monomers are recruited to make an intermediate and then two hexamers form. The OBDs in the center undergo displacement to the surface of the helicase domains permitting the helicases to move in closer to the center (collapse reaction). After the center of the origin is unwound, the OBD hexamers form again and the DNA becomes threaded as shown in the top right. The hexamers then become attached to one another as shown in Fig. 8.

What is the actual structure of the double hexamers? In the presence of ATP, there is no question that single stranded DNA occupies the entire molecule (Borowiec, Dean et al. 1990; Valle, Gruss et al. 2000; VanLoock, Alexandrov et al. 2002; Gomez-Lorenzo, Valle et al. 2003; Valle, Chen et al. 2006; Cuesta, Nunez-Ramirez et al. 2010). However, in the presence of

ADP or a non-hydrolyzable ATP analogue, double hexamers still form (Borowiec and Hurwitz 1988) and these structures have the ability to melt about 8 bp of DNA at the “early” side of the EP region (Borowiec and Hurwitz 1988). The AT track is either unchanged or perhaps becomes partially untwisted (Borowiec and Hurwitz 1988; Borowiec, Dean et al. 1991) and straightened (Han and Hurley 1996). However, complete untwisting of the AT track depends on ATP (Borowiec and Hurwitz 1988). Therefore, in the absence of a usable energy source, most of the origin DNA remains double stranded within the double hexamer. It is therefore most likely that the double hexamer can accommodate double stranded DNA through most if not all of its length (Cuesta, Nunez-Ramirez et al. 2010). If this is the case, then there are at least two ways in which one strand could be moved to the surface of the helicase. The first possibility is that there is a major reorganization of the hexamer structure such that a DNA strand slips through between adjacent monomers in each hexamer and positions itself on the outside. The other mechanism involves breakage of a DNA strand, rethreading to the outside and religation of the strand. There is no experimental evidence whatsoever for the second case but there is some evidence for the first. Within the large tier of the helicase (Li, Zhao et al. 2003), there are six “hydrophilic channels” that separate individual monomers from one another and the displaced strand would have to go through one of them. Mutagenesis of each hydrophilic residue in this channel, not including the ones involved in monomer to monomer contacts, shows that nearly all of them are needed for origin unwinding (Wang, Manna et al. 2007). Importantly, a number of these mutants are not compromised in helicase activity indicating that these channels are needed specifically for unwinding of the origin.

Mutagenesis of certain residues of the OBD (Foster and Simmons 2010) shows that four amino acids placed close to the gap on one monomer form a trough that can allow single-stranded DNA to go through. One interpretation of the data is that the four residues in question participate in the threading of the single stranded DNA to the outside of the complex. Because single stranded DNA binding proteins relieve the mutants’ defect, it’s appealing to think that these residues are normally used to prevent back sliding of the single strand during unwinding. This model is shown in Fig. 7. The active form of the helicase is the double hexamer (Alexandrov, Botchan et al. 2002). Threading of single-stranded DNA through the double hexamer complex is of special interest because it allows us to consider how the strand on the surface of the hexamer becomes positioned so that it can be captured by the DNA polymerase for copying. Although we are still some distance from an actual structure, we can make several educated guesses based on the available evidence. First, the two strands of DNA may be separated just before one strand is pulled into the hexamer by the actions of the helicase. The other strand then surrounds the helicase and modeling shows that the OBD hexamer can be placed fairly closely to the small tier of the helicase so that it seems unlikely that the strand reenters the double hexamer between the helicase and OBD. On the other hand, both modeling and EM images of double hexamers of T antigen associated with origin DNA (Gomez-Lorenzo, Valle et al. 2003; Valle, Chen et al. 2006; Cuesta, Nunez-Ramirez et al. 2010) demonstrate that there are two significant holes between the two hexamers. These holes are made when two lock washer OBD hexamers are attached head to head. The placement of monomer OBDs in each hexamer prevents the two hexamers from coming too close to one another and two holes are generated. These holes are certainly large enough to accommodate single-stranded DNA and it makes sense to propose that the DNA reenters the structure there. One of several possible threading models, consistent with most of the data, is shown in Fig. 8.

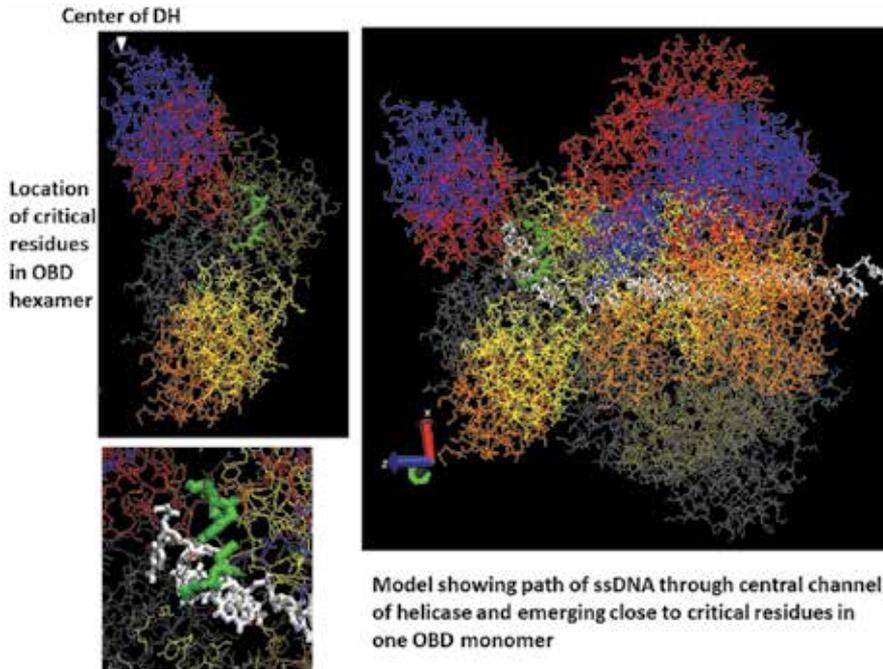


Fig. 7. Model of DNA threading through a hexamer. One strand of DNA (white) is shown to pass within the central channel of the helicase domain and partially through the channel of the OBD lock washer. There, it bends, passes very close to four residues (green) in one OBD monomer and exits the hexamer.

## Representation of DH with ssDNA

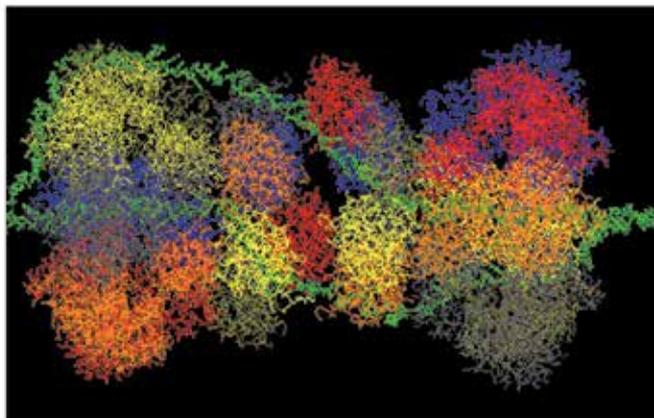


Fig. 8. 2 dimensional projection of an actual 3D model of the threading of single-stranded DNA through a double hexamer of T antigen. The DNA strands are shown in green. Each strand goes through the channel of one hexamer, emerges at the junction between hexamers and surrounds the other hexamer.

A slightly different model of DNA threading was originally proposed by Li et al (Li, Zhao et al. 2003) and Shen et al (Shen, Gai et al. 2005) and more recently by Cuesta (Cuesta, Nunez-Ramirez et al. 2010). They suggest that the  $\beta$ -hairpin structure is the point at which the two strands separate. At that point, one strand would continue to go through the central channel of the hexamer, the other would bend and become threaded through a side channel that exists between the two subdomains of the helicase. This is shown in Fig. 9. There is no direct evidence for either possibility, however. This latter mode of DNA threading appears to be at odds with footprinting data that shows that the DNA in the flanking regions (outside of site II) is accessible to the SSDNA reacting probe phenanthroline copper (Joo, Kim et al. 1998). In either case, a major reorganization of the hexamer structure would be required to transition one strand from the center of the channel to the side channel or to the outside of the helicase. This might be quite tricky for the model shown in Fig. 9. In the absence of strand breaks, one strand would have to move out through a space opened up by two neighboring helicases and then slide into the side channel, which would involve some impressive molecular gymnastics. In the model shown in Fig. 8, each strand would move towards the surface of the helicase altogether. This is the model that is used in the rest of this discussion.

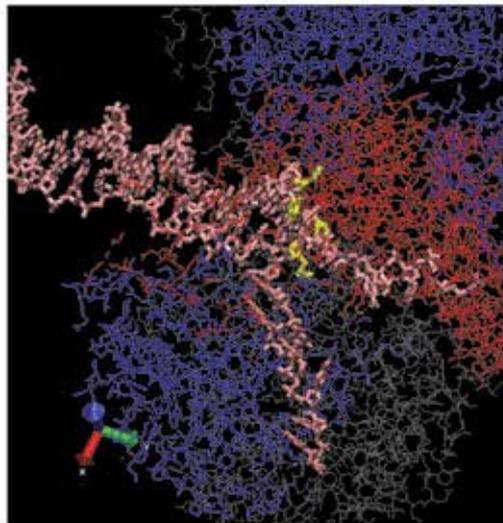


Fig. 9. Model of double stranded DNA (left) entering back end of helicase hexamer (ADP bound form) and separating at  $\beta$ -hairpins with one strand continuing into central channel (right) and the other threading through a side channel (bottom). The DNA is in pink, and four of the 6 His 513 residues in  $\beta$ -hairpin are highlighted in yellow. Some of the surface residues have been cut away for clarity.

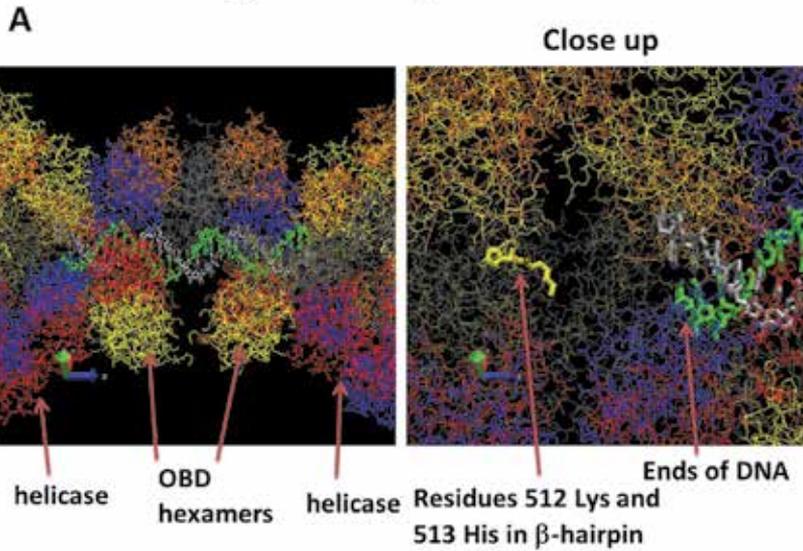
Although there is still a large gap in our understanding of how T antigen becomes a double hexamer with single-stranded DNA going through it, the general features of this structure are fairly well accepted in the field. As shown in Fig. 8, all DNA is single-stranded. Therefore, if this were origin DNA, the center of the origin would also be completely denatured. The difficulty is that there is no obvious way for T antigen to unwind the center of the origin. When the double hexamer first forms, the helicase domains are most likely positioned over the EP and AT regions (Kim, Barbaro et al. 1999; Reese, Sreekumar et al.

2004) where they induce structural distortion. Since T antigen is a 3' to 5' helicase, the 3' end of the strand that lies within the channel faces the center of the origin. Helicase activity would have to take place from the EP and AT regions OUTWARDS. The helicase has no apparent way to access the center of the origin in order to unwind it.

Evidence that the helicase does in fact structurally distort and unwind the center of the origin (mostly site II) exists. While WT T antigen can specifically unwind short substrate DNAs containing primarily site II, mutants with single substitutions in the helicase domain cannot (Wang and Simmons 2009). These mutants are structurally normal, have normal helicase activity, form double hexamers with origin DNA and bind to a variety of DNA substrates normally. Their defect appears to be primarily an inability to unwind origin DNA and, as a consequence, they are totally defective at DNA replication. It makes complete sense to presume, therefore, that helicase domain actively unwinds these short origin DNAs. However, modeling shows that in the context of a double hexamer, the ends of the DNA are far removed from the  $\beta$ -hairpins of the helicase domains. As shown in Fig. 10A, there is no way for the helicase domain to gain access to the ends of the DNA. Other data support the idea that a major structural change takes place in T antigen to unwind the center of the origin. First, Borowiec (Borowiec and Hurwitz 1988) showed that in the presence of ATP, the center of the origin becomes structurally distorted. Second, analysis of the unwinding of various mutant DNA substrates (Wang and Simmons 2009) demonstrated that pentanucleotides 1 and 4 are the most important for the unwinding of small origin DNAs. One way to solve this problem is to assume that the helicase domains are permitted to move into the center of the origin. The only way in which this could be done is by displacing the OBDs and relocating them to the surface of the helicase. The link between the OBD and helicase domain is a long flexible region that could easily be used to move the OBD some distance. Modeling shows that once the OBDs are displaced, the helicase domains could move close enough together so that the ends of these small DNAs can overlap with the  $\beta$ -hairpins (Fig. 10B). This proposed reaction is called "collapse". During collapse, the helicase would have to move 5' to 3' relative to the single stranded DNA within the channel. Although T antigen is primarily a 3' to 5' helicase, it does have a small amount of 5' to 3' activity (Goetz, Dean et al. 1988; Wiekowski, Schwarz et al. 1988).

The idea is also supported by what happens during unwinding of papilloma virus DNA. There, two initiator proteins are used: E1 the helicase and E2 which serves as the origin recognition protein. It has been pretty well documented that E2 binds the origin first, recruits E1 and when E1 assembles into a double hexamer, E2 is displaced and is no longer used (Sedman and Stenlund 1995; Sanders and Stenlund 2000). The situation is analogous to the collapse reaction for T antigen, but because the OBD and helicase are part of the same protein, the OBD can't leave altogether. Our model (Fig. 6) is that the collapse reaction occurs after the EP and AT regions are structurally distorted and this permits the helicase domains to move into the center of the origin where they melt all but about 15 bp. After collapse, the DNA would start moving the other (normal) direction relative to the helicases and the OBDs might very well reform their hexameric structures. The very center of the origin could become melted as the OBDs rehexamerize allowing only one strand to become threaded through the center of the hexamer (last image in Fig. 6). Simultaneously, the two hexamers would orient towards each other head to head, perhaps as illustrated in Fig. 8.

## Modeling of 29 bp DNA with DH



## Model of 29 bp DNA with rearranged DH

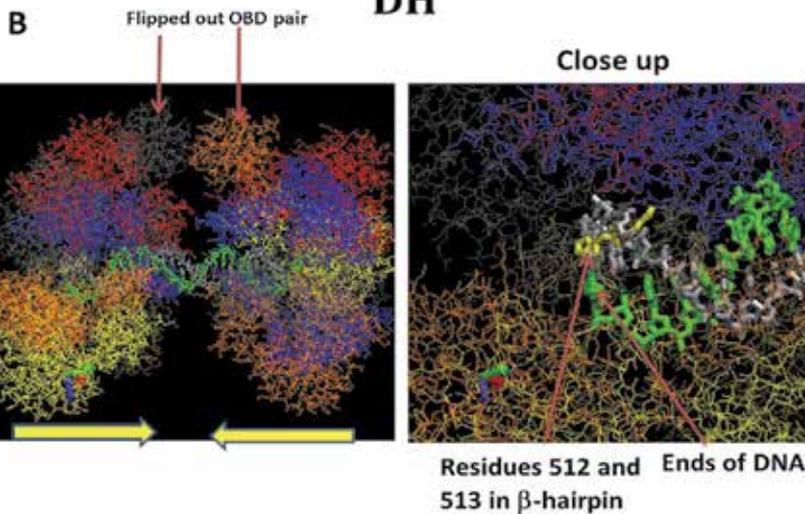


Fig. 10. (A) Modeling of a 29 bp length of DNA in the center of a double hexamer with the OBD lock washers facing each other head to head. In this model, the ends of the DNA are some distance away from the  $\beta$ -hairpins of the helicase domains (close up on right). (B) The OBD lock washers have been removed and the helicase domains have been relocated closer to one another. A pair of OBDs (Bochkareva, Martynowski et al. 2006; Meinke, Phelan et al. 2007) is shown on the surface of the helicases. In this structure, the  $\beta$ -hairpins are close to the ends of the DNA (right).

## 7. Initiation of DNA synthesis

Three cellular proteins are needed for efficient initiation of new chains. Although RPA and pol/prim are absolutely essential for DNA synthesis (Simmons 2000), some DNA synthesis can take place *in vitro* in the absence of topo I (Simmons, Trowbridge et al. 1998). However, topo I stimulates replication by up to 10-folds (Trowbridge, Roy et al. 1999). This observation holds for assays performed with crude cell extracts depleted of topo I as well as for *in vitro* assays with purified proteins. Topo I is needed for the efficient formation of completed form I DNA molecules as well as for the formation of replicative intermediates (Trowbridge, Roy et al. 1999). The cellular enzyme is needed from the beginning of DNA synthesis and is most likely a component of the replication initiation machine (Halmer, Vestner et al. 1998; Trowbridge, Roy et al. 1999). This was demonstrated in a number of ways, but the most convincing was that reactions started with catalytically inactive mutants of topo I could not be rescued by the later addition of WT topo I suggesting that once topo I becomes incorporated into the initiation complex, it cannot be substituted by soluble enzyme. The enzyme is therefore not merely one that is needed for the relaxation of torsionally strained DNA, but is an integral component of the replication complex, and this applies to the physiologically relevant chromatin state as well (Halmer, Vestner et al. 1998).

The order in which the three cellular proteins are recruited to the initiation complex has been studied (Simmons, Gai et al. 2004). In the presence of pol/prim, topo I binds first followed by RPA. Pol/prim was not detected in the complex probably because it has a fleeting association. Of the three cellular proteins, we know the most about how topo I is recruited to the replication machine.

## 8. Association of topo I with the initiation complex

There are two independent binding sites on T antigen for topo I (Simmons, Melendy et al. 1996; Roy, Trowbridge et al. 2003), one located near the C-terminal end between residues 550 and 627 and another near the N-terminal end between residues 99 to 161. The way in which topo I is recruited to T antigen at the C-terminal binding site is reasonably well understood. The binding site consists of a patch of 6 residues on the “back” side of the helicase (Khopde and Simmons 2008) (Fig. 11). Only double hexamers associated with DNA can recruit topo I (Gai, Roy et al. 2000). Single hexamers and intermediate structures are inactive. The stoichiometry of topo I to T antigen in the double hexamer is very close to 1:6 (Gai, Roy et al. 2000; Simmons, Gai et al. 2004), that is, one molecule of topo I per hexamer. It has been demonstrated that DNA flanking the core origin on both sides is necessary for efficient recruitment of topo I (Gai, Roy et al. 2000). The most obvious mechanism of topo I recruitment is therefore that one molecule first binds to the DNA on each side of the origin. Each enzyme then slides over and when it encounters a binding site on one of the helicase subunits, it “snaps” into place. This model (Fig. 12) explains all of the binding data as well as the known stoichiometry. Interestingly, given the predicted structure of the helicase-topo I complex (Khopde and Simmons 2008), DNA going through the center of topo I would have to bend around by about 165°. The idea then is that double-stranded DNA would thread through the topo I molecule, severely bend, and then separate just before entering the “back” of the helicase with one strand funneled through the central channel and the other delivered to the outside of the helicase (Fig. 12).

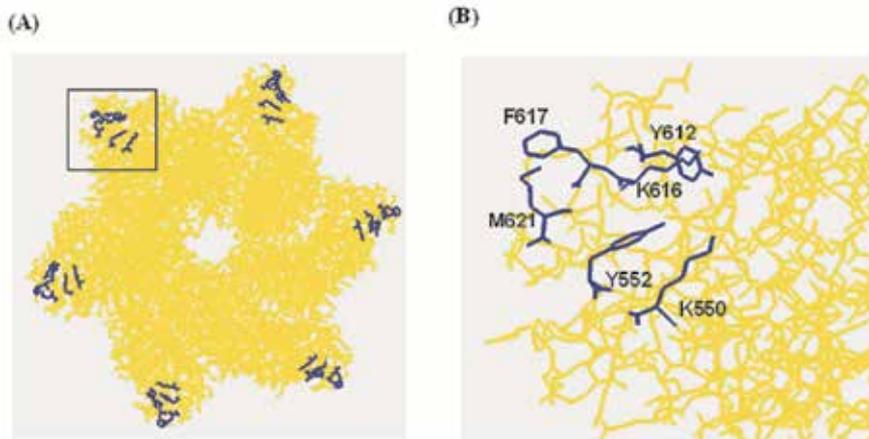


Fig. 11. (A) Structure of the “back” end of the hexameric helicase (Khopde and Simmons 2008) showing a patch of 6 residues that most likely contact topo I. (B) A close up of one of the monomers.

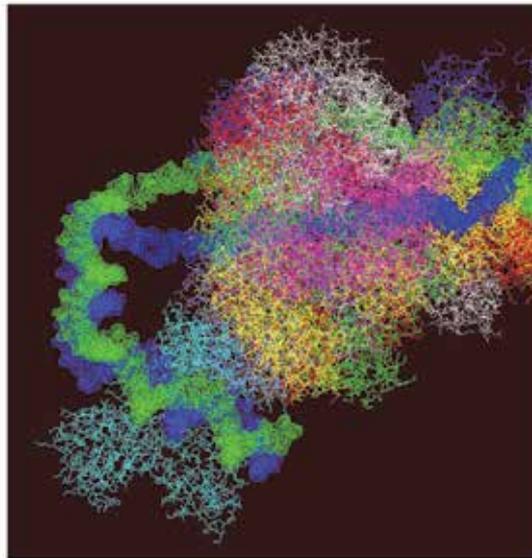


Fig. 12. Model of a T antigen hexamer with DNA threaded as shown in Fig. 8. Topo I (cyan) is shown associated with one of the T antigen helicase monomers through the six residues shown in Fig. 11. The predicted structure shows that double stranded DNA (in green and blue) going through topo I would have to bend before entering the back of the helicase with one strand going through the center channel and one strand going around the helicase domains.

Associated in this way, topo I would be in a perfect position to relax the DNA just before the DNA strands are separated by the helicase. Surprisingly, this is not its function. If this were the case, topo I binding defective mutants would be expected to be compromised in their

ability to unwind a circular DNA molecule in the presence of topo I. Khopde et al (Khopde, Roy et al. 2008) showed that all of these mutants are perfectly normal in DNA unwinding. What could the function of the bound topo I be? Since topo I stimulates DNA replication and is needed from the point of initiation, it makes sense to hypothesize that it is needed for activating or recruiting the polymerase. This was confirmed when it was shown that topo I stimulates pol/prim to synthesize about 6 to 7 times as much RNA-DNA primer (Khopde, Roy et al. 2008). These primers are made by the primase and polymerase activities of pol/prim and are about 36 nucleotides long (Bullock, Seo et al. 1991; Bullock, Tevosian et al. 1994). The most obvious way topo I could stimulate pol/prim is by directly communicating with the polymerase during initiation. At this time, however, there is no evidence that these two proteins can interact with one another.

The function of the topo I binding site near the N-terminal end of T antigen is not well understood. Mutagenesis of this region (unpublished results) has shown that it shares many of the same features as the C-terminal binding site. Mutants that are topo I binding defective cannot support DNA replication normally, although the effects are not quite as dramatic as they are for mutants with substitutions at the C-terminal region. Nevertheless, the data are convincing that both sites are needed for efficient DNA replication. If true, the stoichiometry of topo I to T antigen in the initiation complex should be 2 per hexamer, which is double the observed ratio (Gai, Roy et al. 2000; Simmons, Gai et al. 2004). There are several possible solutions to this dilemma. One is that the same topo I flips back and forth between its two binding sites (the same molecule cannot bind to both sites at once). In principle, this may be possible but it seems unlikely. Two, it's possible that binding to one site or the other is only transiently needed and no stable complex forms. For example, it's possible that the binding to the N-terminal site is only needed in the "collapsed" state described above. Since this structure is most likely very fleeting, a stable association of topo I at this site in the context of a preinitiation complex would not occur. Three, it's possible that the stoichiometry measured previously reflects that of a preinitiation complex, that is one that hasn't actually fired the origin and that initiation depends on another molecule of topo I binding each hexamer. At this time, it's not possible to clearly differentiate between these latter two choices.

## 9. Association of RPA with the initiation complex

RPA is a single-stranded DNA binding protein that is essential for SV40 DNA replication (Simmons 2000). The binding site for RPA on T antigen has been mapped to part of the origin binding domain, including residues 170-190, (Weisshart, Taneja et al. 1998; Arunkumar, Klimovich et al. 2005) and to the J domain (residues 79 to 97) (unpublished results). An NMR structure of the RPA 32 subunit bound to the OBD has been obtained and the contact sites on T antigen for this RPA subunit are near the C-terminal end of the domain (residues 240 to 260) (Arunkumar, Klimovich et al. 2005). The structure of part of RPA70 associated with single-stranded DNA has been determined (Bochkarev, Pfuetzner et al. 1997) but it is not known where exactly it binds to T antigen. There are probably a number of sites of interactions between the two proteins. RPA stimulates the DNA unwinding activity of T antigen (Iftode and Borowiec 1997) and an interaction with T antigen appears to be necessary for DNA replication (Weisshart, Taneja et al. 1998); (Melendy and Stillman 1993). The binding of RPA70 to the T antigen DNA binding domain stimulates the association of RPA with the emerging single stranded DNA during DNA

replication (Jiang, Klimovich et al. 2006). Given the known sites of interaction between T antigen and RPA, it is possible to construct several models that fit the data. Fig. 13 shows one such model. Only RPA70 associated with single-stranded DNA is included in the model and this protein is shown to interact with the OBD as well as with the J domain. How RPA32 and RPA 13 would fit in this structure is nearly impossible to predict based on available structures.

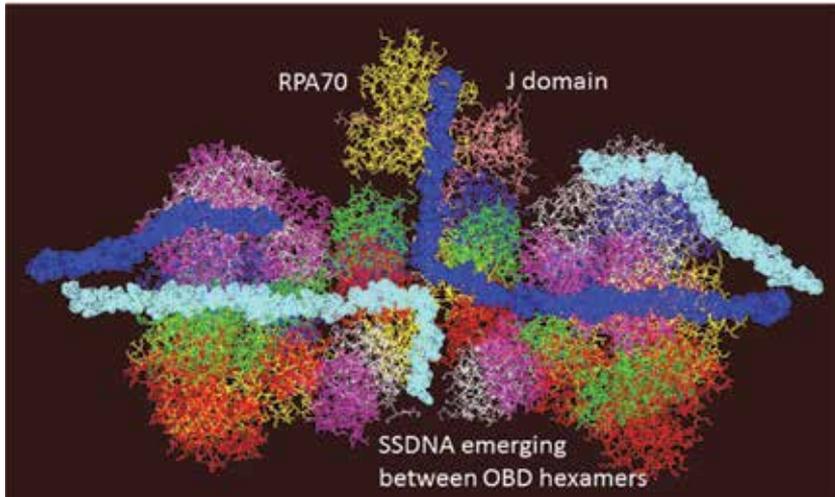


Fig. 13. One possible way in which RPA could associate with a double hexamer. RPA is shown to make contacts with the origin binding domain and with the J domain (mauve). Importantly, the single-strand shown to be emerging from the right hexamer becomes threaded through RPA (yellow). Another RPA molecule (not shown) would be associated with the light blue DNA strand at the bottom.

RPA is known to interact with pol/prim as well (Dornreiter, Erdile et al. 1992; Nasheuer, von Winkler et al. 1992; Braun, Lao et al. 1997). Its interaction with the polymerase complex most likely explains its ability to activate the polymerase during DNA synthesis (Wold 1997). The site on RPA 70K responsible for activating polymerase activity corresponds to the region involved in binding (Braun, Lao et al. 1997; Jacobs, Lipton et al. 1999). Given the critical nature of RPA in DNA synthesis, it makes perfect sense to assume that, in the initiation complex, the polymerase associates with the bound RPA and with T antigen at the same time (see below). The polymerase complex is a huge enzyme, more than two thirds the size of a T antigen hexamer. Its structure is not known, although the structure of the heterodimeric yeast pol alpha has been determined (Klinge, Nunez-Ramirez et al. 2009). The human enzyme has been shown to bind T antigen (Dornreiter, Hoss et al. 1990; Collins and Kelly 1991; Collins, Russo et al. 1993; Dornreiter, Copeland et al. 1993) and this interaction is needed for DNA replication (Taneja, Boche et al. 2007); (Huang, Zhao et al. 2010). Given its size, it might very well cover a substantial part of a hexamer. Nevertheless, these results, in summation, indicate that the "business" end of the T antigen hexamer is near the N-terminal end. This idea actually fits in quite nicely with where the single-stranded DNA is likely to emerge from the double hexamer (Fig.7). We could postulate that this same strand might be captured by an RPA molecule [stoichiometry in double hexamer preinitiation complex is the

same as topo I (Simmons, Gai et al. 2004)] and therefore thread into RPA as it emerges from the holes in the middle of the double hexamer. The polymerase might sit “behind” RPA and the DNA could thread into it as well. The threaded single stranded DNA would then be copied to make the RNA-DNA primer. This primer consists of about 7 nucleotides of RNA and 25-30 nucleotides of DNA (Bullock, Seo et al. 1991; Bullock, Tevosian et al. 1994).

An association between RPA and topo I has also been detected (Khopde, Roy et al. 2008). Since RPA binds to the N-terminal end of T antigen and topo I does as well, it is possible that all three proteins interact with one another there. A possible model incorporating these data is shown in Fig. 14.

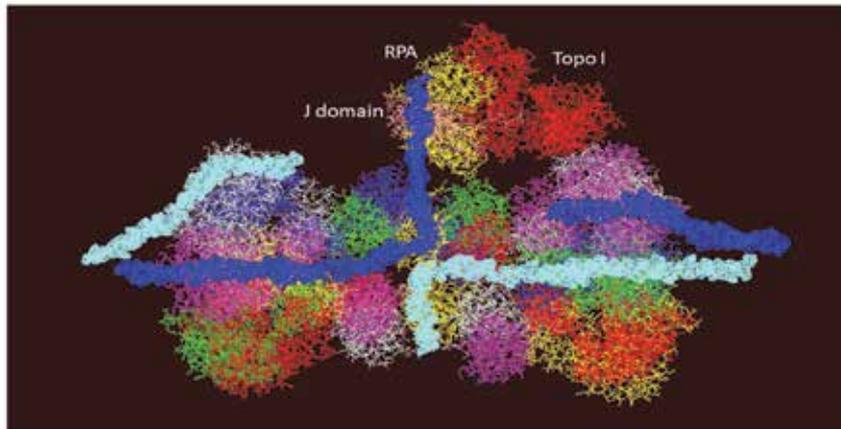


Fig. 14. Model of double hexamer with bound RPA and topo I associated with the N-terminal binding site on T antigen. Topo I binds primarily to the hinge region between the J domain and origin binding domain. In this orientation, topo I is likely to make contacts with RPA as well as with T antigen.

## 10. Overall mechanism of initiation

Taking all the data into accounting, we have come up with a model (Fig. 15) for how the replication machine forms and works to replicate the DNA. First, a double hexamer of T antigen forms over the origin as described above and shown in Fig. 6. The DNA is initially embedded in the channels of the OBD and helicase hexamers, mostly in double stranded form (structure 1). The only exception is probably at the EP region where about 8 nucleotides are single-stranded. How the two DNA strands of that region are threaded through the helicase is of immense interest because this information should tell us the detailed mechanism of unwinding from the origin. One possibility is that one strand is associated with the 6  $\beta$ -hairpins in the center of the hexameric helicase and the other strand is threaded in a different place of the helicase. This place could be one of the side channels, but modeling suggests instead that the second strand is in one of the hydrophilic channels opened up after two monomers become separated from each other. The AT track is probably initially double stranded but untwisted (Borowiec, Dean et al. 1991; Dean and Hurwitz 1991), so all of that region may fit into the channel of the other hexamer. After collapse and strand separation, one strand gets displaced to the surface of the helicase (structure 2). Topo I then binds, first to the DNA on both sides of the origin then snaps into

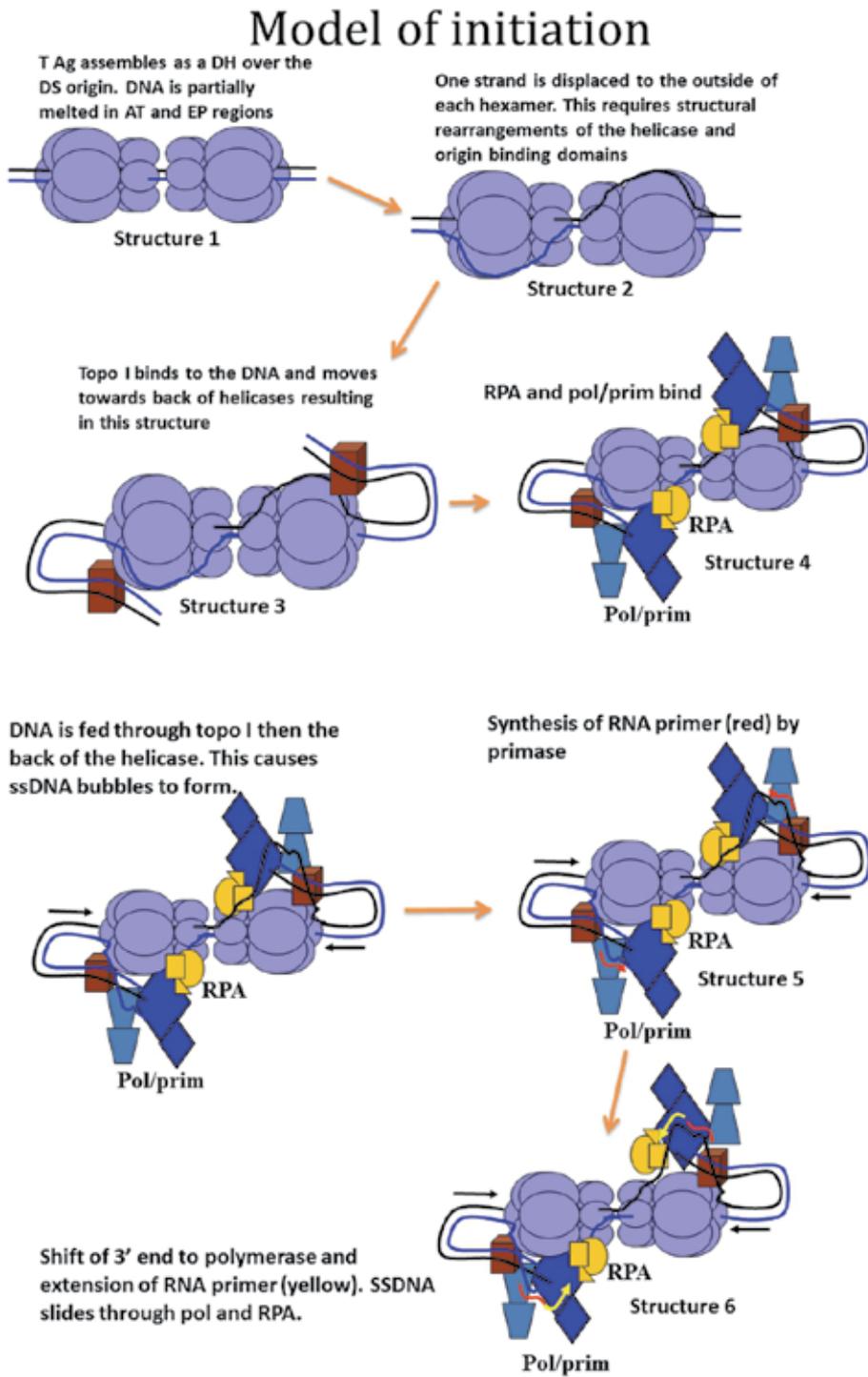


Fig. 15. Model of initiation of DNA synthesis from SV40 DNA. (see text)

place when it meets one of the six binding sites on the “back” end of each hexameric helicase (structure 3). The DNA bends severely before it enters the helicase. RPA is recruited next although it’s not known at what point pol/prim becomes associated with the complex. In this model, we show RPA and pol/prim binding together (structure 4). The RPA molecule is shown to bind to its site near the N-terminal end of T antigen. In order for topo I to stimulate the polymerase to synthesize the RNA/DNA primer as described above, we can assume that pol/prim will attach to the complex via interactions with topo I as well as with T antigen and with RPA. All of these interactions appear to be essential for efficient synthesis of DNA. In this way, pol/prim is now activated and can begin the synthesis of the RNA primer (red, structure 5) followed by DNA (yellow, structure 6). During this whole time, from the time structure 2 is formed, DNA is being pulled into the complex from both sides, and consequently, the single-stranded DNA loops get larger and larger. However, the polymerase probably quickly copies each strand at a particular site leaving single-stranded DNA only on the “back” side of the reaction (Fig. 16, structure 7).

## 11. Elongation reaction and synthesis of Okazaki fragments

Pol/prim is replaced with the processive polymerase after it has finished making the initial primer. This elongation reaction has been well characterized (Simmons 2000). Presumably, pol/prim is displaced when RFC, the clamp loader, or an RFC-PCNA (clamp) complex attaches to the DNA at the primer/template boundary (Ellison and Stillman 1998). Pol delta (Melendy and Stillman 1991; Tsurimoto and Stillman 1991) binds and processively copies the template to make the leading strands on both sides (structures 7-9) (Fig. 16). Behind the first primers, single-stranded DNA accumulates as the helicase separates the parental strands. This DNA is copied discontinuously into 200 nucleotide Okazaki sized segments. Each one is made when another complex of topo I, RPA and pol/prim associate with the template (structure 9). Pol/prim synthesizes only a short RNA/DNA primer as it did at first, and the elongation machinery takes over to make the remainder of the Okazaki fragment. In this way, about half of each strand is copied continuously from the point of initiation and the other half of each strand is copied discontinuously behind the point of initiation as described above (Fig. 17). Termination of DNA synthesis and separation of the two daughter strands require the action of topoisomerase II. Completion of each strand requires, in addition, repair enzymes such as RNase H and DNA ligase.

## 12. Remaining questions

The model described above does not consider the participation of topo I at the N-terminal side of T antigen. This is still a remaining question. The various possible explanations for this were described above. It’s not understood how topo I stimulates pol/prim to make an RNA/DNA primer. Does it just recruit the enzyme? The other major question is the exact way in which the double-stranded DNA becomes denatured and threaded through the double hexamer complex. Insights into this process probably will require the determination of additional T antigen-DNA structures. Another major gap in our knowledge is the interaction between the initiation and elongation machinery. Does the elongation machinery attach to T antigen? If so, where? It is thought that RPA is required for elongation as well as initiation. What is the connection between the RPA molecule that is used in initiation, those involved in coating the single-stranded DNA molecules and the molecules needed for elongation? Finally, how do these protein complexes dissociate from the DNA at

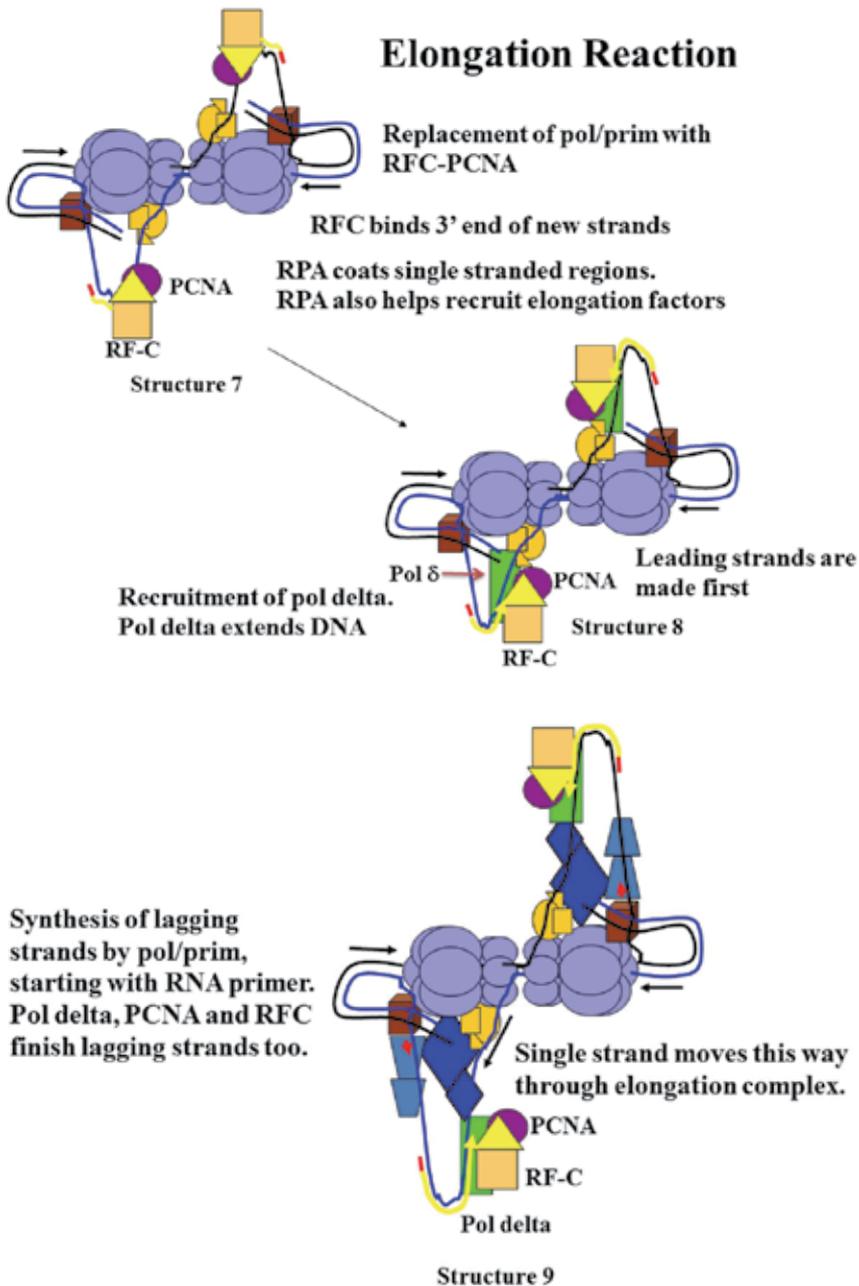


Fig. 16. Model of elongation reaction. After pol/prim has finished making an RNA-DNA primer from each strand, the enzyme is displaced and replaced by a complex of RFC-PCNA-pol  $\delta$ . This complex extends the primer (DNA is shown in yellow). As single-stranded DNA accumulates "behind" the first primer, it becomes associated with a new initiation complex which synthesizes another primer. This machinery is again replaced by an elongation complex to complete the synthesis of each Okazaki fragment.

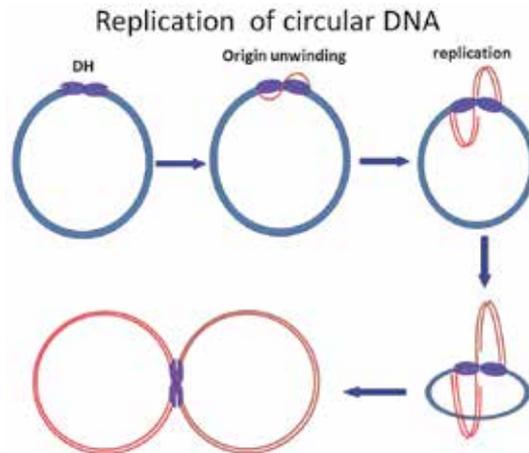


Fig. 17. Duplication of circular SV40 DNA. A double hexamer complex assembles over the origin, unwinds it and cellular proteins copy each strand as described in the text. Each strand is copied continuously and discontinuously.

termination? If either of the threading models in Fig. 9 and 10 is correct, the hexamers would have to be disrupted in order for the DNA to be unraveled. A related question is how the last bit of DNA gets replicated. Is it through a repair process?

Although we have made significant progress in our understanding of the mechanism of virus DNA replication, we are still a long way from understanding these steps at the atomic level. This must be our eventual goal. Additional structural and biochemical information will have to be used to complete this difficult task.

### 13. Acknowledgements

The work in the author's laboratory was supported by a grant from the National Cancer Institute.

### 14. References

- Alexandrov, A. I., M. R. Botchan, et al. (2002). "Characterization of simian virus 40 T-antigen double hexamers bound to a replication fork. The active form of the helicase." *J Biol Chem* 277(47): 44886-97.
- Arunkumar, A. I., V. Klimovich, et al. (2005). "Insights into hRPA32 C-terminal domain--mediated assembly of the simian virus 40 replisome." *Nat Struct Mol Biol* 12(4): 332-9.
- Bochkarev, A., R. A. Pfuetzner, et al. (1997). "Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA." *Nature* 385(6612): 176-81.
- Bochkareva, E., D. Martynowski, et al. (2006). "Structure of the origin-binding domain of simian virus 40 large T antigen bound to DNA." *Embo J* 25(24): 5961-9.
- Borowiec, J. A., F. B. Dean, et al. (1990). "Binding and unwinding--how T antigen engages the SV40 origin of DNA replication." *Cell* 60(2): 181-4.
- Borowiec, J. A., F. B. Dean, et al. (1991). "Differential induction of structural changes in the simian virus 40 origin of replication by T antigen." *J Virol* 65(3): 1228-35.

- Borowiec, J. A. and J. Hurwitz (1988). "Localized melting and structural changes in the SV40 origin of replication induced by T-antigen." *Embo J* 7(10): 3149-58.
- Braun, K. A., Y. Lao, et al. (1997). "Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms." *Biochemistry* 36(28): 8443-54 issn: 0006-2960.
- Bullock, P. A. (1997). "The initiation of simian virus 40 DNA replication in vitro." *Crit Rev Biochem Mol Biol* 32(6): 503-68.
- Bullock, P. A., Y. S. Seo, et al. (1991). "Initiation of simian virus 40 DNA synthesis in vitro." *Mol Cell Biol* 11(5): 2350-61.
- Bullock, P. A., S. Tevosian, et al. (1994). "Mapping initiation sites for simian virus 40 DNA synthesis events in vitro." *Mol Cell Biol* 14(8): 5043-55.
- Campbell, K. S., K. P. Mullane, et al. (1997). "DnaJ/hsp40 chaperone domain of SV40 large T antigen promotes efficient viral DNA replication." *Genes-Dev* 11(9): 1098-110 issn: 0890-9369.
- Collins, K. L. and T. J. Kelly (1991). "Effects of T antigen and replication protein A on the initiation of DNA synthesis by DNA polymerase alpha-primase." *Mol Cell Biol* 11(4): 2108-15.
- Collins, K. L., A. A. Russo, et al. (1993). "The role of the 70 kDa subunit of human DNA polymerase alpha in DNA replication." *Embo J* 12(12): 4555-66.
- Cuesta, I., R. Nunez-Ramirez, et al. (2010). "Conformational rearrangements of SV40 large T antigen during early replication events." *J Mol Biol* 397(5): 1276-86.
- Dean, F. B., J. A. Borowiec, et al. (1987). "Simian virus 40 large tumor antigen requires three core replication origin domains for DNA unwinding and replication in vitro." *Proc Natl Acad Sci U S A* 84(23): 8267-71.
- Dean, F. B. and J. Hurwitz (1991). "Simian virus 40 large T antigen untwists DNA at the origin of DNA replication." *J Biol Chem* 266(8): 5062-71.
- Deb, S., A. L. DeLucia, et al. (1986). "Domain structure of the simian virus 40 core origin of replication." *Mol Cell Biol* 6(5): 1663-70.
- Deb, S., S. Tsui, et al. (1987). "The T-antigen-binding domain of the simian virus 40 core origin of replication." *J Virol* 61(7): 2143-9.
- Dornreiter, I., W. C. Copeland, et al. (1993). "Initiation of simian virus 40 DNA replication requires the interaction of a specific domain of human DNA polymerase alpha with large T antigen." *Mol Cell Biol* 13(2): 809-20.
- Dornreiter, I., L. F. Erdile, et al. (1992). "Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen." *EMBO-J* 11(2): 769-76 issn: 0261-4189.
- Dornreiter, I., A. Hoss, et al. (1990). "SV40 T antigen binds directly to the large subunit of purified DNA polymerase alpha." *EMBO-J* 9(10): 3329-36 issn: 0261-4189.
- Ellison, V. and B. Stillman (1998). "Reconstitution of recombinant human replication factor C (RFC) and identification of an RFC subcomplex possessing DNA-dependent ATPase activity." *J Biol Chem* 273(10): 5979-87.
- Enemark, E. J. and L. Joshua-Tor (2006). "Mechanism of DNA translocation in a replicative hexameric helicase." *Nature* 442(7100): 270-5.
- Fanning, E. (1992). "Simian virus 40 large T antigen: the puzzle, the pieces, and the emerging picture." *J. Virol.* 66: 1289-1293.

- Fanning, E. and R. Knippers (1992). "Structure and function of simian virus 40 large tumor antigen." *Annu. Rev. Biochem.* 61: 55-85.
- Fanning, E. and K. Zhao (2009). "SV40 DNA replication: from the A gene to a nanomachine." *Virology* 384(2): 352-9.
- Foster, E. C. and D. T. Simmons (2010). "The SV40 large T-antigen origin binding domain directly participates in DNA unwinding." *Biochemistry* 49(10): 2087-96.
- Fradet-Turcotte, A., C. Vincent, et al. (2007). "Quantitative analysis of the binding of simian virus 40 large T antigen to DNA." *J Virol* 81(17): 9162-74.
- Gai, D., R. Roy, et al. (2000). "Topoisomerase I associates specifically with simian virus 40 large-T-antigen double hexamer-origin complexes." *J Virol* 74(11): 5224-32.
- Gai, D., R. Zhao, et al. (2004). "Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen." *Cell* 119(1): 47-60.
- Goetz, G. S., F. B. Dean, et al. (1988). "The unwinding of duplex regions in DNA by the simian virus 40 large tumor antigen-associated DNA helicase activity." *J Biol Chem* 263(1): 383-92.
- Gomez-Lorenzo, M. G., M. Valle, et al. (2003). "Large T antigen on the simian virus 40 origin of replication: a 3D snapshot prior to DNA replication." *Embo J* 22(23): 6205-13.
- Halmer, L., B. Vestner, et al. (1998). "Involvement of Topoisomerases in the Initiation of Simian Virus 40 Minichromosome Replication." *J Biol Chem* 273(52): 34792-34798.
- Han, F. X. and L. H. Hurley (1996). "A model for the T-antigen-induced structural alteration of the SV40 replication origin based upon experiments with specific probes for bent, straight, and unwound DNA." *Biochemistry* 35(24): 7993-8001.
- Huang, H., K. Zhao, et al. (2010). "A specific docking site for DNA polymerase {alpha}-primase on the SV40 helicase is required for viral primosome activity, but helicase activity is dispensable." *J Biol Chem* 285(43): 33475-84.
- Humphrey, W., A. Dalke, et al. (1996). "VMD: visual molecular dynamics." *J Mol Graph* 14(1): 33-8, 27-8.
- Iftode, C. and J. A. Borowiec (1997). "Denaturation of the simian virus 40 origin of replication mediated by human replication protein A." *Mol-Cell-Biol* 17(7): 3876-83 issn: 0270-7306.
- Jacobs, D. M., A. S. Lipton, et al. (1999). "Human replication protein A: global fold of the N-terminal RPA-70 domain reveals a basic cleft and flexible C-terminal linker." *J Biomol NMR* 14(4): 321-31.
- Jiang, X., V. Klimovich, et al. (2006). "Structural mechanism of RPA loading on DNA during activation of a simple pre-replication complex." *Embo J* 25(23): 5516-26.
- Joo, W. S., H. Y. Kim, et al. (1998). "Assembly of T-antigen double hexamers on the simian virus 40 core origin requires only a subset of the available binding sites." *Mol Cell Biol* 18(5): 2677-87.
- Khopde, S., R. Roy, et al. (2008). "The Binding of Topoisomerase I to T Antigen Enhances the Synthesis of RNA-DNA Primers during Simian Virus 40 DNA Replication." *Biochemistry* 47(36): 9653-60.
- Khopde, S. and D. T. Simmons (2008). "Simian virus 40 DNA replication is dependent on an interaction between topoisomerase I and the C-terminal end of T antigen." *J Virol* 82(3): 1136-45.
- Kim, H. Y., B. Y. Ahn, et al. (2001). "Structural basis for the inactivation of retinoblastoma tumor suppressor by SV40 large T antigen." *Embo J* 20(1-2): 295-304.

- Kim, H. Y., B. A. Barbaro, et al. (1999). "Sequence requirements for the assembly of simian virus 40 T antigen and the T-antigen origin binding domain on the viral core origin of replication." *J Virol* 73(9): 7543-55.
- Klinge, S., R. Nunez-Ramirez, et al. (2009). "3D architecture of DNA Pol alpha reveals the functional core of multi-subunit replicative polymerases." *Embo J* 28(13): 1978-87.
- Li, D., R. Zhao, et al. (2003). "Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen." *Nature* 423(6939): 512-8.
- Luo, X., D. G. Sanford, et al. (1996). "Solution structure of the origin DNA-binding domain of SV40 T-antigen." *Nat-Struct-Biol* 3(12): 1034-9 issn: 1072-8368.
- Mastrangelo, I. A., P. V. C. Hough, et al. (1989). "ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication." *Nature* 338: 658-662.
- Meinke, G., P. A. Bullock, et al. (2006). "Crystal structure of the simian virus 40 large T-antigen origin-binding domain." *J Virol* 80(9): 4304-12.
- Meinke, G., P. Phelan, et al. (2007). "The crystal structure of the SV40 T-antigen origin binding domain in complex with DNA." *PLoS Biol* 5(2): e23.
- Meinke, G., P. J. Phelan, et al. (2010). "Structure-based analysis of the interaction between the simian virus 40 T-antigen origin binding domain and single-stranded DNA." *J Virol* 85(2): 818-27.
- Melendy, T. and B. Stillman (1991). "Purification of DNA polymerase delta as an essential simian virus 40 DNA replication factor." *J Biol Chem* 266(3): 1942-9.
- Melendy, T. and B. Stillman (1993). "An interaction between replication protein A and SV40 T antigen appears essential for primosome assembly during SV40 DNA replication." *J Biol Chem* 268(5): 3389-95.
- Nasheuer, H. P., D. von Winkler, et al. (1992). "Purification and functional characterization of bovine RP-A in an in vitro SV40 DNA replication system." *Chromosoma* 102(1): S52-9.
- Parsons, R. E., J. E. Stenger, et al. (1991). "Cooperative assembly of simian virus 40 T-antigen hexamers on functional halves of the replication origin." *J Virol* 65(6): 2798-806.
- Reese, D. K., G. Meinke, et al. (2006). "Analyses of the interaction between the origin binding domain from simian virus 40 T antigen and single-stranded DNA provide insights into DNA unwinding and initiation of DNA replication." *J Virol* 80(24): 12248-59.
- Reese, D. K., K. R. Sreekumar, et al. (2004). "Interactions required for binding of simian virus 40 T antigen to the viral origin and molecular modeling of initial assembly events." *J Virol* 78(6): 2921-34.
- Roy, R., P. Trowbridge, et al. (2003). "The Cap Region of Topoisomerase I Binds to Sites near Both Ends of Simian Virus 40 T Antigen." *J Virol* 77(18): 9809-16.
- Rutila, J. E., M. J. Imperiale, et al. (1986). "Replication and transformation functions of in vitro-generated simian virus 40 large T antigen mutants." *J Virol* 58(2): 526-35.
- Sanders, C. M. and A. Stenlund (2000). "Transcription factor-dependent loading of the E1 initiator reveals modular assembly of the papillomavirus origin melting complex." *J Biol Chem* 275(5): 3522-34.
- Sawai, E. T., G. Rasmussen, et al. (1994). "Construction of SV40 deletion mutants and delimitation of the binding domain for heat shock protein to the amino terminus of large T-antigen." *Virus Res* 31(3): 367-378.

- Sedman, J. and A. Stenlund (1995). "Co-operative interaction between the initiator E1 and the transcriptional activator E2 is required for replicator specific DNA replication of bovine papillomavirus in vivo and in vitro." *Embo J* 14(24): 6218-28.
- Shen, J., D. Gai, et al. (2005). "The roles of the residues on the channel beta-hairpin and loop structures of simian virus 40 hexameric helicase." *Proc Natl Acad Sci U S A* 102(32): 11248-53.
- Simmons, D. T. (2000). "SV40 large T antigen: Functions in DNA replication and transformation [In Process Citation]." *Adv Virus Res* 55: 75-134.
- Simmons, D. T., D. Gai, et al. (2004). "Assembly of the replication initiation complex on SV40 origin DNA." *Nucleic Acids Res* 32(3): 1103-12.
- Simmons, D. T., T. Melendy, et al. (1996). "Simian virus 40 large T antigen binds to topoisomerase I." *Virology* 222(2): 365-74.
- Simmons, D. T., R. Roy, et al. (1998). "The Activity of Topoisomerase I Is Modulated by Large T Antigen during Unwinding of the SV40 Origin." *J Biol Chem* 273(32): 20390-6.
- Simmons, D. T., P. W. Trowbridge, et al. (1998). "Topoisomerase I stimulates SV40 T antigen-mediated DNA replication and inhibits T antigen's ability to unwind DNA at nonorigin sites." *Virology* 242(2): 435-43.
- Sreekumar, K. R., A. E. Prack, et al. (2000). "The simian virus 40 core origin contains two separate sequence modules that support T-antigen double-hexamer assembly." *J Virol* 74(18): 8589-600.
- Srinivasan, A., A. J. McClellan, et al. (1997). "The amino-terminal transforming region of simian virus 40 large T and small t antigens functions as a J domain." *Mol-Cell-Biol* 17(8): 4761-73 issn: 0270-7306.
- Stubdal, H., J. Zalvide, et al. (1997). "Inactivation of pRB-related proteins p130 and p107 mediated by the J domain of simian virus 40 large T antigen." *Mol Cell Biol* 17(9): 4979-90.
- Sullivan, C. S. and J. M. Pipas (2002). "T antigens of simian virus 40: molecular chaperones for viral replication and tumorigenesis." *Microbiol Mol Biol Rev* 66(2): 179-202.
- Taneja, P., I. Boche, et al. (2007). "Different activities of the largest subunit of replication protein A cooperate during SV40 DNA replication." *FEBS Lett* 581(21): 3973-8.
- Tjian, R. (1978). "Protein-DNA interactions at the origin of simian virus 40 DNA replication." *Cold Spring Harbor Symp. Quant. Biol.* 43: 655-662.
- Trowbridge, P. W., R. Roy, et al. (1999). "Human Topoisomerase I Promotes Initiation of Simian Virus 40 DNA Replication In Vitro." *Mol Cell Biol* 19(3): 1686-1694.
- Tsurimoto, T. and B. Stillman (1991). "Replication factors required for SV40 DNA replication in vitro. II. Switching of DNA polymerase alpha and delta during initiation of leading and lagging strand synthesis." *J Biol Chem* 266(3): 1961-8.
- Valle, M., X. S. Chen, et al. (2006). "Structural basis for the cooperative assembly of large T antigen on the origin of replication." *J Mol Biol* 357(4): 1295-305.
- Valle, M., C. Gruss, et al. (2000). "Large T-antigen double hexamers imaged at the simian virus 40 origin of replication." *Mol Cell Biol* 20(1): 34-41.
- VanLoock, M. S., A. Alexandrov, et al. (2002). "SV40 large T antigen hexamer structure: domain organization and DNA-induced conformational changes." *Curr Biol* 12(6): 472-6.

- Wang, W., D. Manna, et al. (2007). "Role of the hydrophilic channels of simian virus 40 T-antigen helicase in DNA replication." *J Virol* 81(9): 4510-9.
- Wang, W. and D. T. Simmons (2009). "Simian virus 40 large T antigen can specifically unwind the central palindrome at the origin of DNA replication." *J Virol* 83(7): 3312-22.
- Weisshart, K., M. K. Bradley, et al. (1996). "An N-terminal deletion mutant of simian virus 40 (SV40) large T antigen oligomerizes incorrectly on SV40 DNA but retains the ability to bind to DNA polymerase alpha and replicate SV40 DNA in vitro." *J-Virol* 70(6): 3509-16 issn: 0022-538x.
- Weisshart, K., P. Taneja, et al. (1998). "The replication protein A binding site in simian virus 40 (SV40) T antigen and its role in the initial steps of SV40 DNA replication [In Process Citation]." *J Virol* 72(12): 9771-81.
- Wiekowski, M., M. W. Schwarz, et al. (1988). "Simian virus 40 large T antigen DNA helicase. Characterization of the ATPase-dependent DNA unwinding activity and its substrate requirements." *J Biol Chem* 263(1): 436-42.
- Wold, M. S. (1997). "Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism." *Annu Rev Biochem* 66: 61-92.
- Zalvide, J., H. Stubdal, et al. (1998). "The J domain of simian virus 40 large T antigen is required to functionally inactivate RB family proteins." *Mol Cell Biol* 18(3): 1408-15.

# Function of DNA Polymerase $\alpha$ in a Replication Fork and its Putative Roles in Genomic Stability and Eukaryotic Evolution

Masaharu Takemura  
*Laboratory of Biology Education*  
*Graduate School of Mathematics and Science Education*  
*Tokyo University of Science*  
*Japan*

## 1. Introduction

DNA replication takes place during S phase (Kornberg & Baker, 1992). In eukaryotic cells, DNA replication is performed through the concerned action of plural DNA polymerases with their accessory proteins (Wang, 1991; Hübscher et al., 2002; Burgers, 2009). Eukaryotic cells contain three replicative DNA polymerases, named DNA polymerases  $\alpha$ ,  $\delta$ , and  $\epsilon$ , mitochondrial DNA polymerase  $\gamma$ , and at least 12 non-replicative DNA polymerases, named DNA polymerases  $\beta$ ,  $\zeta$ ,  $\eta$ ,  $\theta$ ,  $\iota$ ,  $\kappa$ ,  $\lambda$ ,  $\mu$ , and  $\nu$ , terminal deoxynucleotidyl transferase (TdT), and Rev1 (Hübscher et al., 2002; Livneh et al., 2010). DNA replication progresses according to the catalytic activity of three replicative DNA polymerases (pol-)  $\alpha$ ,  $\delta$ , and  $\epsilon$ , cooperatively (Burgers, 2009). Pol- $\alpha$  is necessary for the onset of DNA synthesis, which synthesizes RNA-DNA primer in both leading and lagging strands (Wang, 1991; Hübscher et al., 2002; Burgers, 2009). Pol- $\delta$  is probably involved in the synthesis of Okazaki fragment in the lagging strand, and pol- $\epsilon$  is thought to synthesize long DNA strand in the leading strand (Pavlov et al., 2006a; Pavlov et al., 2006b; Pursell et al., 2007) (Fig. 1).

Pol- $\alpha$  is the first eukaryotic DNA polymerase ever discovered (Bollum, 1960). Its complex type with DNA primase (pol- $\alpha$ -primase complex) is the only DNA polymerase that can initiate DNA synthesis *de novo*. Therefore, pol- $\alpha$  is an important enzyme for cell survival (Murakami et al., 1985; Budd & Campbell, 1987; Takada-Takayama et al., 1991). Pol- $\alpha$ -primase complex comprises four subunits, each of which is highly conserved in eukaryotes from yeast to human (Wang, 1991). The function of this 'classical' DNA polymerase at DNA replication reaction has been well established. Primase is an RNA polymerase. It is able to synthesize RNA primer, following DNA synthesis by the catalytic subunit of pol- $\alpha$ . Almost all DNA replication researchers consider its role in DNA replication reaction as a mere 'initiator' of DNA chain elongation. Nevertheless, several recent examinations of this 'classical' DNA polymerase have revealed its heretofore unknown but important roles in DNA replication. They have revealed a link between DNA replication and other cellular dynamism, in addition to advancing the evolution of eukarya. According to this recent progress, I describe several topics related to eukaryotic pol- $\alpha$  in this chapter, not only as an 'initiator', but also as a 'key regulator'.

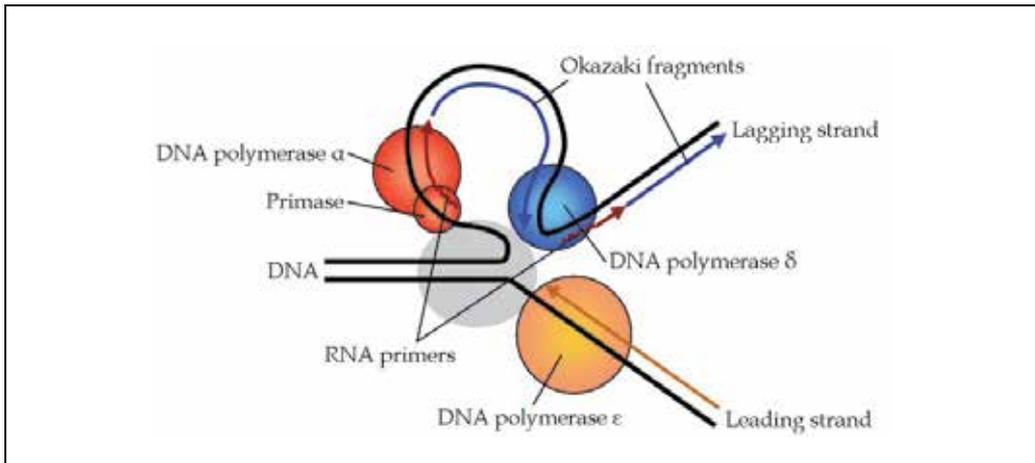


Fig. 1. Typical model of eukaryotic DNA replication fork.

## 2. Structure of pol- $\alpha$ -primase complex

In 1960, pol- $\alpha$  became the first discovered eukaryotic DNA polymerase from calf thymus (Bollum, 1960). Eukaryotic pol- $\alpha$  was found to be tightly associated with DNA primase (Yagura et al., 1982; Yagura et al., 1983; Frick & Richardson, 2001). Now pol- $\alpha$  has been purified as a multisubunit complex that is tightly associated with DNA primase activity from various eukaryotic cells and tissues, the so-called pol- $\alpha$ -primase complex (Wang, 1991; Frick & Richardson, 2001; Hübscher et al., 2002).

Pol- $\alpha$ -primase complex comprises four subunits: p180 (in mouse and human), the largest subunit associated with catalytic polymerase activity; p68 (in mouse and human) and p54 (in mouse, p58 in human), the second-largest and third-largest subunits whose function has been investigated to date; and p46 (in mouse, p48 in human), a primase subunit (Miyazawa et al., 1993). Subunit compositions of pol- $\alpha$ -primase complex in each typical organism are presented in Table 1.

In calf, pol- $\alpha$ -primase complex can be immunopurified from calf thymus using anti-pol- $\alpha$  monoclonal antibody SJK287-38 or MT-17 (Nasheuer & Grosse, 1987; Tamai et al., 1988; Takemura et al., 1997). Purified calf pol- $\alpha$  comprises at least 3 or 4 subunits (Nasheuer & Grosse, 1987; Tamai et al., 1988; Takemura et al., 1997). Reportedly, the immunopurified calf thymus pol- $\alpha$ -primase complex using monoclonal antibody SJK287-38 has 148-180, 73, 59, and 48 kDa subunits (Nasheuer & Grosse, 1987). However, Tamai et al. reported that the immunopurified calf pol- $\alpha$ -primase complex using MT-17 consists of at least three subunits, which have 140, 50, and 47 kDa, respectively (Tamai et al., 1988). It was also reported that the immunopurified pol- $\alpha$ -primase complex using MT-17 from calf thymus had 4 subunits with 140, 73, 50, and 47 kDa (Takemura et al., 1997), corresponding with the subunit composition in human and mouse (Mizuno et al., 1999; see also Mizuno's chapter).

The subunits of pol- $\alpha$ -primase complex in the yeast *Saccharomyces cerevisiae* are 167, 79, 62, and 48 kDa polypeptides (Hübscher et al., 2002; Biswas et al., 2003). The pol- $\alpha$ -primase complex from *Drosophila melanogaster* consists of 182, 73, 60, and 50 kDa subunits (Kaguni et al., 1983; Cotterill et al., 1987). The estimated roles of each subunit of pol- $\alpha$  is summarized in Fig. 2.

	Human	Mouse	Calf	Fly	Budding yeast	Fission yeast
Catalytic subunit	180	180	140	182	167	170
2nd-largest subunit	68	68	73	73	79	66
3rd-largest subunit	58	54	50	60	62	55
Primase subunit	48	46	47	50	48	45

Table 1. Subunit compositions (kDa) of pol- $\alpha$ -primase complex in six organisms.

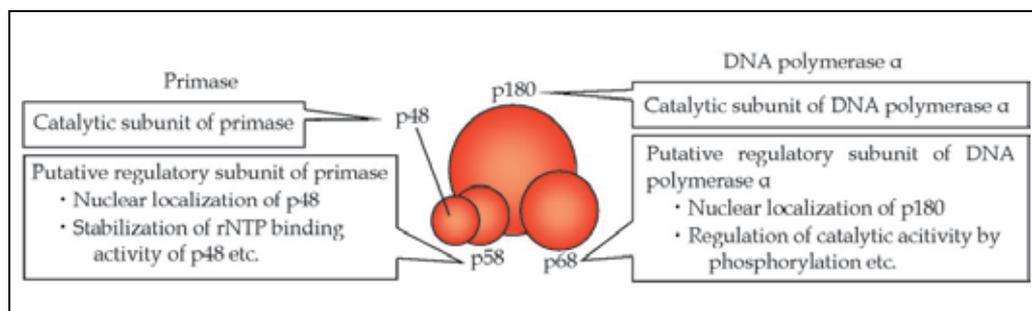


Fig. 2. The estimated roles of each subunit.

Catalytic subunit of pol- $\alpha$  is thought to bind several carbohydrate chains. Reportedly, several lectins such as Concanavalin A bound human catalytic subunit of pol- $\alpha$  (p180), and that ricin from castor beans bound and inhibited pol- $\alpha$ , suggesting that the p180 subunit has carbohydrate chains recognized by these lectins (Bhattacharya et al., 1979; Hsi et al., 1990). Because tunicamycin, which inhibits the glycosylation of protein, has been reported to decrease the activity of pol- $\alpha$  (Bhattacharya & Basu, 1982), it is suggested that the carbohydrate chain on pol- $\alpha$  is important for its activity. However, the detailed roles of carbohydrate chains on p180 subunit have not been elucidated to date.

In the p180 catalytic subunit, three separate domains were identified: an N-terminal domain (amino acids 1-329), a central domain (amino acids 330-1279), and a C-terminal domain (amino acids 1235-1465) (Hübscher et al., 2002). Immunohistochemical analysis using several mutant subunits whose abilities of interaction with other subunits were absent has revealed the mechanism of the subunit assembly, domain structure, and stepwise formation of pol- $\alpha$ -primase complex at the cellular level (Mizuno et al. 1999). The N-terminal domain is thought to be involved in the catalytic activity and the assembly of tetrameric complex, but the detailed roles have not been elucidated. The central domain contains conserved regions among B-family DNA polymerases, which are responsible for dNTP binding, template DNA binding, and the catalytic activity forming a phosphodiester bond (Hübscher et al., 2002). The C-terminal domain was thought to be necessary for its interaction with both p68 subunit and p54/p46 primase subunits (Mizuno et al., 1998; Mizuno et al., 1999) (see also Mizuno's chapter). Recently, the 3D architecture of *Saccharomyces cerevisiae* pol- $\alpha$  was constructed (Klinge et al., 2009). According to this architecture, the combination between X-ray crystallography and electron microscopy revealed the structural properties of the

subunit interaction, especially between p180 and p68 subunit. It was revealed that the C-terminal domain of the catalytic subunit of yeast pol- $\alpha$  contains two C4 zinc-binding motifs and interacts with the second-largest subunit through the  $\alpha 2$  helix (Klinge et al., 2009; Johansson & MacNeil, 2010).

Although the tetrameric complex of pol- $\alpha$ -primase is apparently tightly formed, it has some flexibility. As described above, calf thymus pol- $\alpha$ -primase complex is sometimes purified in the trimeric complex, which lacks the second-largest subunit (Tamai et al., 1988). In the case of the tetrameric complex purified, the amounts of the second-largest subunit are sometimes lower than those of other subunits (Nasheuer & Grosse, 1987; Takemura et al., 1997). It is particularly interesting that Mizuno et al. found that there are many 'free' second-largest (p68) subunit in COS-1 cells, in addition to that in tetrameric pol- $\alpha$ -primase complex (Mizuno et al., 1999). In *Coprinus cinereus*, one basidiomycete, pol- $\alpha$ -primase complex has been reported to dissociate, and the catalytic subunit (p135 in *Coprinus cinereus*) has been reported to be purified detached from other subunits during meiosis (Sawado & Sakaguchi, 1997; Namekawa et al., 2003). These findings suggest that pol- $\alpha$ -primase complex tightly forms in the DNA replication initiation step at the replication fork, and dissociates in other phases or function, before or after DNA replication reaction.

### 3. Function of pol- $\alpha$ in DNA replication initiation

According to the current models, the pol- $\alpha$ -primase complex initiates RNA/DNA primer synthesis on both leading and lagging strands (Waga & Stillman, 1998; Burgers 2009). The primase subunit of the complex synthesizes RNA primer with subsequent short DNA primer synthesis by the catalytic subunit of the complex. In one putative model, pol- $\alpha$  is switched to another processive pol- $\epsilon$  on the leading strand, and to pol- $\delta$  for complete Okazaki fragment synthesis on the lagging strand.

Pol- $\delta$  and pol- $\epsilon$  catalytic subunits are known to both contain a 3'-5' proofreading exonuclease domain. Therefore, they have 3'-5' exonuclease activity, otherwise pol- $\alpha$  catalytic subunit has no intrinsic exonuclease activity (Shevelev & Hübscher, 2002), except for the study of *Drosophila melanogaster* and recombinant *Saccharomyces cerevisiae* pol- $\alpha$  (Cotterill et al., 1987; Brooke et al., 1991). Because of the lack of proofreading activity, the initial deoxyribonucleotides in each Okazaki fragment and the leading strand are expected to contain misincorporated bases, which are able to cause mutation. Instead of the lack of 3'-5' exonuclease activity, previous reports describe nuclear factors that are thought to be responsible for proofreading, thereby covering a deficit of intrinsic exonuclease of pol- $\alpha$ , or which have been co-purified with pol- $\alpha$ -primase complex (Bialek & Grosse, 1993; Brown et al., 2002). It was shown that the p53, a tumour suppressor protein, contains an intrinsic 3'-5' exonuclease activity, and physically binds to pol- $\alpha$  (Mummenbrauer et al., 1996; Huang, 1998; Melle & Nasheuer, 2002). It remains unclear whether these nuclear factors work as 3'-5' exonuclease substituted to the lost exonuclease activity of pol- $\alpha$  or not. A recent model suggests that the errors that occurred on short DNA primer by pol- $\alpha$  are removed by other proofreading (or editing) mechanisms. A possible model is that the errors made by pol- $\alpha$  are edited out by FEN1, flap endonuclease 1, during a process that removes RNA primer (Bae et al., 2001). This model is supported by a finding that the yeast cells carrying FEN1 endonuclease-defective mutation had a 25-fold increase in rates of base substitution, and mice carrying the FEN1 mutation caused in cancer (Zheng et al., 2007; Zheng & Shen, 2011).

Another model is that the intrinsic 3'-5' exonuclease activity of pol- $\delta$  proofreads errors made by pol- $\alpha$  (Pavlov et al., 2006b).

To date, various studies have examined the link mechanism between the cell cycle and the DNA replication progression. Phosphorylation of proteins by cdk-cyclin families as cell cycle regulators is the most important mechanism to regulate the cell cycle. The activities of cdk2/cyclin E and cdk2/cyclin A complexes respectively peak in G<sub>1</sub> and S/G<sub>2</sub> phases of the cell cycle (Dulić et al., 1992; Koff et al., 1992; Rosenblatt et al., 1992). The cell cycle-dependent phosphorylation on pol- $\alpha$  is also found by several research groups. Catalytic p180 subunit of pol- $\alpha$  is reported to be phosphorylated throughout the cell cycle and is hyper-phosphorylated in the G<sub>2</sub>/M phase (Nasheuer et al., 1991). The hyper-phosphorylated type of pol- $\alpha$  has lower affinity for single-stranded DNA (Nasheuer et al., 1991), suggesting that the phosphorylation of pol- $\alpha$  suppresses the activity of pol- $\alpha$  in its DNA synthesis via change in its conformation from active to inactive. The second-largest subunit is also shown to be phosphorylated in a cell cycle-dependent manner (Nasheuer et al., 1991). These subunits are phosphorylated by both cdk2/cyclin E and cdk2/cyclin A (Voitenleitner et al., 1999). Reportedly, the pol- $\alpha$  activity was decreased as a result of the phosphorylation of the second-largest subunit (Voitenleitner et al., 1999). It is particularly interesting that cdk2/cyclin A executes both stimulatory and inhibitory effects on pol- $\alpha$  activity in initiating *in vitro* DNA replication, which is putatively caused by distinct phosphorylation events on pol- $\alpha$  by cdk2/cyclin A (Schub et al., 2001). In yeast, dephosphorylation of the second-largest subunit occurs while cells exit mitosis (Foiani et al., 1995). The second-largest subunit is thought to have a regulatory role in an early stage of S phase. It is reportedly associated with the origin recognition complex (ORC) for initiation of DNA replication (Uchiyama & Wang). Huang et al. also reported that the interaction of the second-largest subunit with both catalytic p180 subunit and helicase is necessary for the SV40 primosome activity (Huang et al., 2010). The author and colleagues reported that the interaction of the catalytic p180 subunit of pol- $\alpha$  with hyper-phosphorylated retinoblastoma protein is regulated by the phosphorylation of the second-largest subunit (Takemura et al., 2006). These results cumulatively suggest that the second-largest subunit of pol- $\alpha$  regulates various stages of the activity of pol- $\alpha$ , depending upon its phosphorylation status (Fig. 2).

#### 4. Various proteins bound to pol- $\alpha$

Although the mechanisms that regulate DNA polymerases and DNA replication are not yet fully understood, *in vitro* DNA replication systems using tumour viruses such as simian virus 40 (SV40) have provided a useful tool for study in this area (Kelly, 1988; Hurwitz et al., 1990). SV40 DNA replication is accomplished by DNA replication machinery of the host cultured cells with the assistance of a single viral protein, large T antigen, which unwinds the double DNA strand at replication origin, thereby forming a preinitiation complex (Dean et al., 1987; Dodson et al., 1987; Stillman, 1989; Tsurimoto et al., 1990; Collins & Kelly, 1991; Murakami et al., 1992). With these *in vitro* DNA replication studies as a start, many studies of accessory factors associated with pol- $\alpha$  have been conducted to date (Fig. 3).

SV40 large T antigen can form a complex with pol- $\alpha$ , and stimulate its activity (Smale & Tjian, 1986; Dornreiter et al., 1990; Collins & Kelly, 1991; Dornreiter et al., 1992). Although DNA polymerase  $\alpha$  has been shown to exist in the cell nucleus throughout the cell cycle (Nakamura et al., 1984), it might act only in the S phase. Therefore, pol- $\alpha$  might be regulated post-translationally by some nuclear factors. A number of stimulatory / accessory proteins

have been reported in addition to SV40 large T antigen. C-factor ( $C_1C_2$  factor), microtubule-associated protein (MAP)-2, and factor T were, respectively, found to interact with and activate pol- $\alpha$  (Lamothe et al., 1981; Pritchard & DePamphilis, 1983; Pritchard et al., 1983; Yoshida et al., 1989; Shioda et al., 1991). CDC45 were bound to pol- $\alpha$ , by which DNA polymerase  $\alpha$  is loaded onto the chromatin (Mimura & Takisawa, 1998). Chk1, an intra-S and DNA damage checkpoint protein, is also reported to bind physically with pol- $\alpha$ , suggesting that pol- $\alpha$  is also an important component of the signal transduction cascade activating the checkpoint (Toricani et al., 2009).

Simbulan et al. have reported that the poly (ADP-ribose) polymerase (PARP) is able to stimulate the activity of pol- $\alpha$ , and that auto-poly(ADP-ribosylation) of PARP itself decreases its stimulatory activity (Simbulan et al., 1993). Actually, PARP is known to post-translationally modulate target proteins by transfer and polymerization of ADP-ribose, using NAD as substrates (Miwa & Masutani, 2007). PARP (especially PARP-1) was found in the nuclei and was shown to be activated by DNA strand breaks (Miwa & Masutani, 2007). Although the role of PARP in DNA replication progression remains unclear, it was suggested that PARP constitutes a signal that switches off DNA synthesis temporarily via the interaction to pol- $\alpha$ , to ensure that lesions are not replicated before DNA repair (Fig. 3).

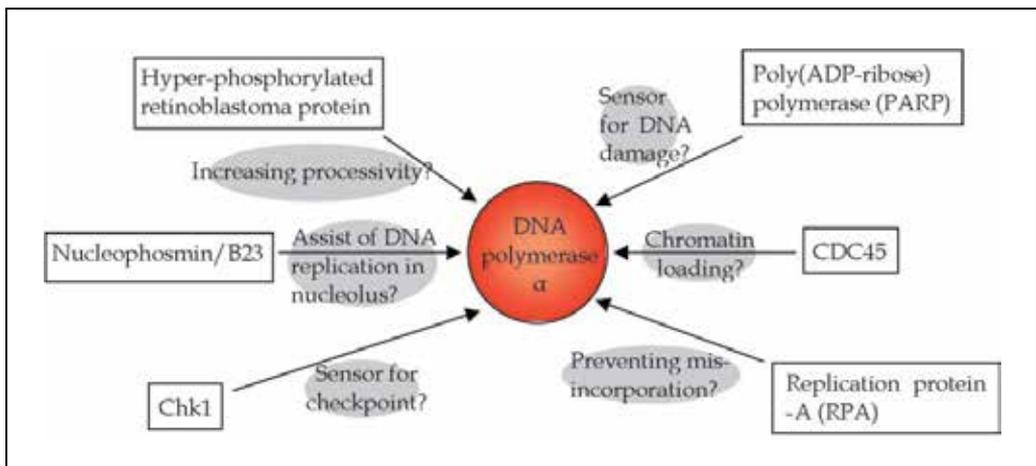


Fig. 3. Pol- $\alpha$ -binding proteins reported.

The author and colleagues found that the nucleophosmin/B23, a major nucleolar protein, is capable of binding to pol- $\alpha$  directly and stimulate its activity (Takemura et al., 1994; Takemura et al., 1999). Nucleophosmin/B23 exists in two isoforms, which are produced through alternative splicing, designed as B23.1 and B23.2. Among them, results demonstrated that purified B23.1 stimulates pol- $\alpha$  activity, and B23.2 does so more weakly than B23.1 (Takemura et al., 1994; Umekawa et al., 2001). Feuerstein et al. have reported that nucleophosmin/B23 is copurified with pol- $\alpha$  on conventional purification (Feuerstein et al., 1990). These results suggest that B23.1 might be involved in DNA replication, aside from the ribosome assembly process in the nucleolus (Orrick et al., 1973; Olson et al., 1974; Schmidt-Zachmann et al., 1987; Chan et al., 1989). Hyper-phosphorylated retinoblastoma protein was reported to bind with pol- $\alpha$  and to raise its processivity (Takemura et al., 1997; Takemura, 2002).

Replication protein-A (RPA) is a single-strand DNA binding protein that is necessary for SV40 DNA replication *in vitro*, stabilizing single-strand DNA after unwinding. RPA comprises three subunits. It reportedly interacts with pol- $\alpha$ . It is known that pol- $\alpha$  pauses at some sites on the natural DNA template. These pause sites are known to be hot spots for nucleotide misincorporation (Fry & Loeb, 1992). RPA reportedly leads pol- $\alpha$  to overcome these error-prone pause sites, suggesting that RPA might reduce misincorporation at some pause sites by pol- $\alpha$  (Suzuki et al., 1994). Maga et al. found that RPA, as an auxiliary factor for pol- $\alpha$ , stabilizes pol- $\alpha$ -primase complex and reduces the misincorporation efficiency of pol- $\alpha$  (Maga et al., 2001).

## 5. How does pol- $\alpha$ contribute to genomic stability?

Recent studies have revealed various functions of pol- $\alpha$ , not only in RNA/DNA synthesis, but also in the mechanism linking DNA replication to the other cellular regulatory mechanisms. One is the relation between pol- $\alpha$  and mutation, which sometimes causes cancer. Another is the role of pol- $\alpha$  on genomic stability including telomere maintenance.

As described above, pol- $\alpha$  lacks proofreading activity, an intrinsic 3'-5' exonuclease activity, so it performs DNA synthesis with a fidelity that is lower than the other DNA polymerases pol- $\delta$  and pol- $\epsilon$  by ten-fold or more (Pavlov et al., 2006b). Limsirichaikul et al. found that pol- $\alpha$  would leave more mismatched primers without additional extension without proofreading (Limsirichaikul et al., 2003). In one estimation, pol- $\alpha$  is thought to generate 30,000 mutations in one mammalian cell division (Albertson & Preston, 2006). In *Saccharomyces cerevisiae* pol- $\alpha$ , Gly952 residue which is conserved amino acid residue, is necessary for catalytic activity of pol- $\alpha$ , and is therefore important for correct deoxyribonucleotide incorporation (Ogawa et al., 2003; Limsirichaikul et al., 2003). Amino acid substitution (Gly to Ala) of this Gly952 caused the decrease of replication fidelity (Limsirichaikul et al., 2003). On the other hand, amino acid substitution of another amino acid residue, Leu868, of the catalytic subunit of pol- $\alpha$  decreased the replication fidelity of pol- $\alpha$  and increased mutation rates in *Saccharomyces cerevisiae* (Niimi et al., 2004). This mutant, L868F, had a spontaneous error frequency of 3 in 100 nucleotides, and performed 570-fold lower replication fidelity than wild type. Another research group reported that the several mutants including temperature sensitive (ts) mutants of pol- $\alpha$  increases mutation rates (Liu et al., 1999). Recently, Tanaka et al. reported that the base selection step functions to prevent replication errors also in human pol- $\alpha$  and to maintain genomic stability (Tanaka et al., 2010). These results suggest that the replication fidelity of pol- $\alpha$  is important for avoiding mutagenesis and for maintaining genomic integrity.

Various replication stresses engender genomic instabilities. Various mutational analyses of pol- $\alpha$  have shown that the functional aberration affects transcriptional gene silencing, homologous recombination, and chromosomal instability. In *Arabidopsis*, mutation of the catalytic subunit of pol- $\alpha$  reduced histone demethylation in some promoter domains and delayed G<sub>2</sub>/M phase with high expression of a G<sub>2</sub>/M marker gene, suggesting that pol- $\alpha$  plays a role in some epigenetic states (Liu et al., 2010). Liu et al. also reported that the mutation of pol- $\alpha$  increased the frequency of homologous recombination (Liu et al., 2010). The author and colleagues reported that the inhibition of DNA replication by selective inhibitor of pol- $\alpha$  caused chromosomal instability of cultured opossum (*Didelphis marsupialis virginiana*) cells (Takemura et al., 2011). In *Saccharomyces cerevisiae*, low levels of pol- $\alpha$  gene expression in cells reportedly induced chromosomal translocation (Lemoine et al., 2005) and mitotic and meiotic instability in rRNA gene cluster (Casper et al., 2008), suggesting that pol- $\alpha$  contributes to chromosomal stability throughout the cell cycle directly in some mechanism.

Telomere maintenance is an important task that must be done to ensure chromosome stability, especially in germ line and cancer cells. Using pol- $\alpha$ -ts mutant, tsFT20 (Murakami et al., 1985), results showed that the telomere G-tail was induced to increase markedly in the overall telomere length when pol- $\alpha$  was inactivated (Nakamura et al., 2005). It is particularly interesting that the inhibition of pol- $\alpha$  also caused a considerable increase of Robertsonian chromosome fusions (Nakamura et al., 2005). These observations correspond to our *in vitro* data indicating that pol- $\alpha$  itself performs replication slippage on a telomeric repeat (Nozawa et al., 2000), and that hyper-phosphorylated retinoblastoma protein suppresses replication slippage by pol- $\alpha$  in normal cells (Takemura et al., 2008). These results suggest that the steady state expression of pol- $\alpha$  and its normal regulation by other accessory factors suppresses chromosome abnormality, including the expansion of telomeric repeat.

In telomeric DNA replication, Cdc13 protein is known to be a key factor for replication and telomere capping in yeast (Evans & Lundblad, 2000). Cdc13 is also known to interact with the catalytic subunit of pol- $\alpha$  (Qi & Zakian, 2000), suggesting that pol- $\alpha$  has an important role in telomere DNA replication. Grossi et al. described that the second-largest subunit of pol- $\alpha$  physically interacts with Stn1 protein, which is known to be the regulator of telomere end (Grossi et al., 2004). In pol- $\alpha$  (*cdc17/pol1*), a mutant of yeast revealed defects in telomeric lagging strand synthesis (Martin et al., 2000). These results suggest that pol- $\alpha$  acts as a regulator of telomeric lagging strand synthesis completion linking to the action of telomerase, via the interaction with Cdc13 and Stn1.

## 6. Remaining mysteries of pol- $\alpha$

To date, much research about one eukaryotic key DNA polymerase, pol- $\alpha$ , has been performed by numerous research groups. The initiation mechanism of pol- $\alpha$  has been well studied, and the interaction of pol- $\alpha$  with other DNA polymerases and other replication factors has been clarified. In addition to these previous efforts by previous researchers, several studies have revealed unknown profiles of pol- $\alpha$ .

### 6.1 Detailed structure, attached molecules, and intracellular amounts of pol- $\alpha$

It was reported that purified pol- $\alpha$  from calf thymus showed microheterogeneity in its catalytic subunit, especially in their molecular weights: 150 kDa, 145 kDa, and 140 kDa (Masaki et al., 1982). Pol- $\alpha$ , pol- $\delta$ , and pol- $\epsilon$  of rat ascites hepatoma cells, Novikoff hepatoma cells, were shown to differ from those of normal rat liver cells in their  $K_m$  values of DNA-binding (Popanda et al., 1995). It has not been clarified how these fluctuations about molecular 'species' of pol- $\alpha$  have been revealed, in spite of the suggestion that the 'programmed proteolysis' of the catalytic subunit of pol- $\alpha$  might be performed (Masaki et al., 1982).

Simbulan et al. reported that sialic acid-containing glycolipids inhibit the activity of pol- $\alpha$  (Simbulan et al., 1994). Although sialic acid is an extremely important carbohydrate attached to the carbohydrate chain, such as that on cell membrane, it is not known whether it functions in the cell nucleus. Reportedly, pol- $\alpha$  itself is one such glycoprotein (Bhattacharya et al., 1979; Hsi et al., 1990). Detailed roles of the carbohydrate chain on pol- $\alpha$  have not yet been clarified.

One more mystery of pol- $\alpha$  is why it is abundantly present in mammalian cells. Pol- $\alpha$  has been known to have large quantities in the cell nucleus compared with other pol- $\delta$  and pol- $\epsilon$ . Among them, only pol- $\alpha$  molecules required for DNA replication initiation are thought to

participate in DNA replication machinery. The reasons for the large number of pol- $\alpha$  molecules apparently standing by in the cell nucleus remain unknown.

### 6.2 How has pol- $\alpha$ evolved?

It has been proposed that three DNA polymerases-- pol- $\alpha$ , pol- $\delta$ , and pol- $\epsilon$ --have evolved from ancestral  $\epsilon$ -like DNA polymerase (Edgell et al., 1998). Molecular evolutionary studies suggest that pol- $\epsilon$  is an ancestral type of these replicative DNA polymerases in eukaryotes (Edgell et al., 1998). Actually, pol- $\alpha$  and pol- $\delta$  were thought to have arisen from two gene duplications, emerging from ancestral pol- $\epsilon$  (Edgell et al. 1998). Filée et al. suggested that the evolutionary history of DNA replication proteins such as DNA polymerases involve significant exchanges among viruses, plasmids, and their host cells (Filée et al., 2002). According to the molecular evolutionary analysis, the author advanced the hypothesis that ancestors of pol- $\alpha$  gene were derived from a DNA virus-like organism, which had some eukaryote-like characteristics, such as poxviruses (Takemura, 2001). Villarreal & DeFilippis proposed that the origin of the eukaryotic pol- $\delta$  should be some algal viral DNA polymerases (Villarreal & DeFilippis, 2000). In spite of these attractive studies, the mystery of the evolution of the eukaryotic DNA polymerases remains to be elucidated.

### 6.3 How has pol- $\alpha$ contributed to eukaryotic evolution?

In mammals, a gene of catalytic subunit of pol- $\alpha$  is established in the X-added region of an X chromosome in eutherians, although it is on the autosome in metatherians (marsupials). This locational property is thought to have arisen from chromosomal translocation that occurred in the eutherian lineage 105 million years ago (Fig. 4). One large segment of one autosome was thought to translocate to X chromosomes: the so-called 'X-added' region (Graves, 1995). Although the effects of this large-scale translocation on the evolution of eutheria are entirely unknown, according to the simulation of spontaneous mutation rates in these two mammalian groups, the author has proposed that the X-linked pol- $\alpha$  gene contributes to greater diversity of eutherian mammals (Takemura, 2008). The solution of this problem, however, will require much more time and study.

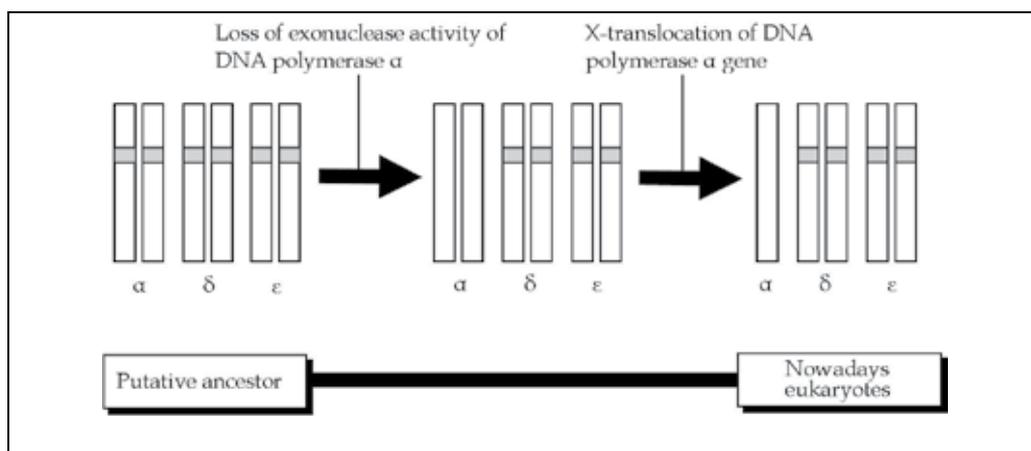


Fig. 4. Remaining mystery of pol- $\alpha$ .

## 7. Acknowledgements

I thank Drs. Shonen Yoshida (a Professor Emeritus of Nagoya University, Japan) and Hayato Umekawa (Mie University, Japan), supervisor of laboratories in Nagoya University or in Mie University to which I had been assigned previously. I also thank S. Y., H. U., and Dr. Motoshi Suzuki (Nagoya University) for their reviews of manuscript. My works described in this chapter were supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

## 8. References

- Albertson, T. M. & Preston, B. D. (2006). DNA replication fidelity: proofreading in trans. *Current Biology*, Vol.16, No.6, (March 2006), pp. R209-R211
- Bae, S. H.; Bae, K. H.; Kim, J. A. & Seo, Y. S. (2001). RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes. *Nature*, Vol.412, No.6845, (July 2001), pp. 456-461
- Bhattacharya, P.; Simet, I. & Basu, S. (1979). Inhibition of human neuroblastoma DNA polymerase activities by plant lectins and toxins. *Proceedings of the National Academy of Science, USA*, Vol.76, No.5, (May 1979), pp. 2218-2221
- Bhattacharya, P. & Basu, S. (1982). Probable involvement of a glycoconjugate in IMR-32 DNA synthesis: decrease of DNA polymerase  $\alpha_2$  activity after tunicamycin treatment. *Proceedings of the National Academy of Science, USA*, Vol.79, No.5, (March 1982), pp. 1488-1491
- Bialek, G. & Groose, F. (1993). An error-correcting proofreading exonuclease-polymerase that copurifies with DNA-polymerase- $\alpha$ -primase. *The Journal of Biological Chemistry*, Vol.268, No.8, (March 1993), pp.6024-6033
- Biswas, S. B.; Khopde, S. M.; Zhu, F. & Biswas, E. E. (2003). Subunit interactions in the assembly of *Saccharomyces cerevisiae* DNA polymerase  $\alpha$ . *Nucleic Acids Research*, Vol.31, No.8, (April 2003), pp. 2056-2065
- Bollum, F. J. (1960). Calf thymus polymerase. *The Journal of Biological Chemistry*, Vol.235, No.8, (August 1960), pp. 2399-2403
- Brooke, R. G.; Singhal, R.; Hinkle, D. C. & Dumas, L. B. (1991). Purification and characterization of the 180- and 86-kilodalton subunits of the *Saccharomyces cerevisiae* DNA primase-DNA polymerase protein complex. The 180-kilodalton subunit has both DNA polymerase and 3'→5'-exonuclease activities. *The Journal of Biological Chemistry*, Vol.266, No.5, (February 1991), pp. 3005-3015
- Brown, K. R.; Weatherdon, K. L.; Galligan, C. L. & Skalski, V. (2002). A nuclear 3'-5' exonuclease proofreads for the exonuclease-deficient DNA polymerase  $\alpha$ . *DNA Repair*, Vol.1, No.10, (October 2002), pp. 795-810
- Budd, M. & Campbell, J. L. (1987). Temperature-sensitive mutations in the yeast DNA polymerase I gene. *Proceedings of the National Academy of Science, USA*, Vol.84, No.9, (May 1987), pp. 2838-2842
- Burgers, P. M. J. (2009). Polymerase dynamics at the eukaryotic DNA replication fork. *The Journal of Biological Chemistry*, Vol.284, No.7, (February 2009), pp. 4041-4045
- Casper, A. M.; Mieczkowski, P. A.; Gawel, M. & Petes, P. D. (2008). Low levels of DNA polymerase alpha induce mitotic and meiotic instability in the ribosomal DNA

- gene cluster of *Saccharomyces cerevisiae*. *PLoS Genetics*, Vol.4, No.6, (June 2008), e1000105
- Chan, W. Y.; Liu, Q. R.; Borjigin, J.; Busch, H.; Rennert, O. M., Tease, L. A. & Chan, P. K. (1989). Characterization of the cDNA encoding human nucleophosmin and studies of its role in normal and abnormal growth. *Biochemistry*, Vol.28, No.3, (February 1989), pp. 1033-1039
- Collins, K. L. & Kelly, T. J. (1991). Effects of T antigen and replication protein A on the initiation of DNA synthesis by DNA polymerase  $\alpha$ -primase. *Molecular and Cellular Biology*, Vol.11, No.4, (April 1991), pp. 2108-2115
- Cotterill, S. M.; Reyland, M. E.; Loeb, L. A. & Lehman, I. R. (1987). A cryptic proofreading 3'→5' exonuclease associated with the polymerase subunit of the DNA polymerase-primase from *Drosophila melanogaster*. *Proceedings of the National Academy of Science, USA*, Vol.84, No.16, (August 1987), pp. 5635-5639
- Dean, F. B.; Bullock, P.; Murakami, Y.; Wobbe, C. R.; Weissbach, L. & Hurwitz, J. (1987). Simian virus 40 (SV40) DNA replication: SV40 large T antigen unwinds DNA containing the SV40 origin of replication. *Proceedings of the National Academy of Science, USA*, Vol.84, No.1, (January 1987), pp. 16-20
- Dodson, M.; Dean, F. B.; Bullock, P.; Echols, H. & Hurwitz, J. (1987). Unwinding of duplex DNA from the SV40 origin of replication by T antigen. *Science*, Vol.238, No.4829, (November 1987), pp. 964-967
- Dornreiter, I.; Höss, A.; Arthur, A. K. & Fanning, E. (1990). SV40 T antigen binds directly to the large subunit of purified DNA polymerase  $\alpha$ . *The EMBO Journal*, Vol.9, No.10, (October 1990), pp. 3329-3336
- Dornreiter, I.; Erdile, L. F.; Gilbert, I. U.; von Winkler, D.; Kelly, T. J. & Fanning, E. (1992). Interaction of DNA polymerase  $\alpha$ -primase with cellular replication protein A and SV40 T antigen. *The EMBO Journal*, Vol.11, No.2, (February 1992), pp. 769-776
- Dulić V.; Lees, E. & Reed, S. I. (1992). Association of human cyclin E with a periodic G1-S phase protein kinase. *Science*, Vol.257, No.5078, (September 1992), pp. 1958-1961
- Edgell, D. R.; Malik, S. B. & Doolittle, W. F. (1998). Evidence of independent gene duplications during the evolution of archaeal and eukaryotic family B DNA polymerases. *Molecular Biology and Evolution*, Vol.15, No.9, (September 1998), pp. 1207-1217
- Evans, S. K. & Lundblad, V. (1999). Est1 and Cdc13 as comediators of telomerase access. *Science*, Vol.286, No.5437, (October 1999), pp. 117-120
- Feuerstein, N.; Mond, J. J.; Kinchington, P. R.; Hickey, R.; Karjalainen-Lindsberg, M. L.; Hay, I. & Ruyechan, W. T. (1990). Evidence for DNA binding activity of numatrin (B23), a cell cycle-regulated nuclear matrix protein. *Biochimica et Biophysica Acta*, Vol.1087, No.2, (October 1990), pp. 127-136
- Filée, J.; Forterre, P.; Sen-Lin, T. & Laurent, J. (2002). Evolution of DNA polymerase families: evidences for multiple gene exchange between cellular and viral proteins. *Journal of Molecular Evolution*, Vol.54, No.6, (June 2002), pp. 763-773
- Foiani, M.; Liberi, G.; Lucchini, G. & Plevani, P. (1995). Cell cycle-dependent phosphorylation and dephosphorylation of the yeast DNA polymerase  $\alpha$ -primase B subunit. *Molecular and Cellular Biology*, Vol.15, No.2, (February 1995), pp. 883-891

- Frick, D. N. & Richardson, C. C. (2001). DNA primases. *Annual Reviews of Biochemistry*, Vol.70, pp. 39-80
- Fry, M. & Loeb, L. A. (1992). A DNA polymerase  $\alpha$  pause site is a hot spot for nucleotide misinsertion. *Proceedings of the National Academy of Science, USA*, Vol.89, No.2, (January 1992), pp. 763-767
- Graves, J. A. (1995). The origin and function of the mammalian Y chromosome and Y-borne genes – an evolving understanding. *BioEssays*, Vol.17, No.4, (April 1995), pp. 311-320
- Grossi, S.; Puglisi, A.; Dmitriev, P. V.; Lopes, M. & Shore, D. (2004). Pol12, the B subunit of DNA polymerase  $\alpha$ , functions in both telomere capping and length regulation. *Genes & Development*, Vol.18, No.9, (May 2004), pp. 992-1006
- Hsi, K.-L.; Copeland, W. C. & Wang, T. S.-F. (1990). Human DNA polymerase  $\alpha$  catalytic polypeptide binds ConA and RCA and contains a specific labile site in the N-terminus. *Nucleic Acids Research*, Vol.18, No.21, (November 1990), pp. 6231-6237
- Huang, P. (1998). Excision of mismatched nucleotides from DNA: a potential mechanism for enhancing DNA replication fidelity by wild-type p53 protein. *Oncogene*, Vol.17, No.3, (July 1998), pp. 261-270
- Huang, H.; Weiner, B. E.; Zhang, H.; Fuller, B. E.; Gao, Y.; Wile, B. M.; Zhao, K.; Arnett, D. R.; Chazin, W. J. & Fanning, E. (2010). Structure of a DNA polymerase  $\alpha$ -primase domain that docks on the SV40 helicase and activates the viral primosome. *The Journal of Biological Chemistry*, Vol.285, No.22, (May 2010), pp. 17112-17122
- Hübscher, U.; Maga, G. & Spadari, S. (2002). Eukaryotic DNA polymerases. *Annual Reviews of Biochemistry*, Vol.71, pp. 133-163
- Hurwitz, J.; Dean, F. B.; Kwong, A. D. & Lee, S.-H. (1990). The in vitro replication of DNA containing the SV40 origin. *The Journal of Biological Chemistry*, Vol.265, No.30, (October 1990), pp. 18043-18046
- Johansson, E. & MacNeill, S. A. (2010). The eukaryotic replicative DNA polymerases take shape. *Trends in Biochemical Sciences*, Vol.35, No.6, (June 2010), pp.339-347
- Kaguni, L. S.; Rossignol, J. M.; Conaway, R. C. & Lehman, I. R. (1983). Isolation of an intact DNA polymerase-primase from embryos of *Drosophila melanogaster*. *Proceedings of the National Academy of Science, USA*, Vol.80, No.8, (April 1983), pp. 2221-2225
- Kelly, T. J. (1988). SV40 DNA replication. *The Journal of Biological Chemistry*, Vol.,263, No.34, (December 1988), pp. 17889-17892
- Klinge, S.; Núñez-Ramírez, R.; Llorca, O. & Pellegrini, L. (2009). 3D architecture of DNA pol  $\alpha$  reveals the functional core of multi-subunit replicative polymerases. *The EMBO Journal*, Vol.28, No.13, (July 2009), pp. 1978-1987
- Koff, A.; Giordano, A.; Desai, D.; Yamashita, K.; Harper, J. W.; Elledge, S.; Nishimoto, T.; Morgan, D. O.; Franza, B. R. & Roberts, J. M. (1992). Formation and activation of a cyclin E-cdk2 complex during the G1 phase of the human cell cycle. *Science*, Vol.257, No.5077, (September 1992), pp. 1689-1694
- Kornberg, A. & Baker, T. A. (1992). *DNA Replication*, 2<sup>nd</sup> edition, Freeman WD and Co., ISBN 1-891389-44-0, New York, USA
- Lamothe, P.; Baril, B.; Chi, A.; Lee, L. & Baril, E. (1981). Accessory proteins for DNA polymerase  $\alpha$  activity with single-strand DNA templates. *Proceedings of the National Academy of Science, USA*, Vol.78, No.8, (August 1981), pp. 4723-4727

- Lemoine, F. J.; Degtyareva, N. P.; Lobachev, K. and Petes, T. D. (2005). Chromosomal translocations in yeast induced by low levels of DNA polymerase: a model for chromosome fragile sites. *Cell*, Vol.120, No.5, (March 2005), pp. 587-598
- Limsirichaikul, S.; Ogawa, M.; Niimi, A.; Iwai, S.; Murate, T.; Yoshida, S. & Suzuki, M. (2003). The Gly952 residue of *Saccharomyces cerevisiae* DNA polymerase  $\alpha$  is important in discriminating correct deoxyribonucleotides from incorrect ones. *The Journal of Biological Chemistry*, Vol.278, No.21, (May 2003), pp. 19079-19086
- Liu, J.; Ren, X.; Yin, H.; Wang, Y.; Xia, R.; Wang, Y. & Gong, Z. (2010). Mutation in the catalytic subunit of DNA polymerase  $\alpha$  influences transcriptional gene silencing and homologous recombination in *Arabidopsis*. *The Plant Journal*, Vol 61, No.1, (January 2010), pp. 36-45
- Liu, V. F.; Bhaumik, D. & Wang, T. S.-F. (1999). Mutator phenotype induced by aberrant replication. *Molecular and Cellular Biology*, Vol.19, No.2, (February 1999), pp. 1126-1135
- Livneh, Z.; Ziv, O. & Shachar, S. (2010). Multiple two-polymerase mechanisms in mammalian translesion DNA synthesis. *Cell Cycle*, Vol.9, No.4, (February 2010), pp. 729-735
- Maga, G.; Frouin, I.; Spadari, S. & Hübscher, U. (2001). Replication protein A as a „fidelity clamp“ for DNA polymerase  $\alpha$ . *The Journal of Biological Chemistry*, Vol.276, No.21, (May 2001), pp. 18235-18242
- Martin, A. A.; Dionne, I.; Wellinger, R. J. & Holm, C. (2000). The function of DNA polymerase  $\alpha$  at telomeric G tails is important for telomere homeostasis. *Molecular and Cellular Biology*, Vol.20, No.3, (February 2000), pp. 786-796
- Masaki, S.; Koiwai, O. & Yoshida, S. (1982). 10 S DNA polymerase  $\alpha$  of calf thymus shows a microheterogeneity in its large polypeptide component. *The Journal of Biological Chemistry*, Vol.257, No.12, (June 1982), pp. 7172-7177
- Melle, C. & Nasheuer, H. P. (2002). Physical and functional interactions of the tumor suppressor protein p53 and DNA polymerase  $\alpha$ -primase. *Nucleic Acids Research*, Vol.30, No.7, (April 2002), pp.1493-1499
- Mimura, S. & Takisawa, H. (1998). *Xenopus* cdc45-dependent loading of DNA polymerase  $\alpha$  onto chromatin under the control of S-phase cdk. *The EMBO Journal*, Vol.17, No.19, (October 1998), pp. 5699-5707
- Miwa, M. & Masutani, M. (2007). PolyADP-ribosylation and cancer. *Cancer Science*, Vol.98, No.10, (October 2007), pp. 1528-1535
- Miyazawa, H.; Izumi, M.; Tada, S.; Takada, R.; Masutani, M.; Ui, M. & Hanaoka, F. (1993). Molecular cloning of the cDNAs for the four subunits of mouse DNA polymerase  $\alpha$ -primase complex and their gene expression during cell proliferation and the cell cycle. *The Journal of Biological Chemistry*, Vol.268, No.11, (April 1993), pp. 8111-8122
- Mizuno, T.; Ito, N.; Yokoi, M.; Kobayashi, A.; Tamai, K.; Miyazawa, H. & Hanaoka, F. (1998). The second-largest subunit of mouse DNA polymerase  $\alpha$ -primase complex facilitates both production and nuclear translocation of the catalytic subunit of DNA polymerase  $\alpha$ . *Molecular and Cellular Biology*, Vol.18, No.6, (June 1998), pp. 3552-3562

- Mizuno, T.; Yamagishi, K.; Miyazawa, H. & Hanaoka, F. (1999). Molecular architecture of the mouse DNA polymerase  $\alpha$ -primase complex. *Molecular and Cellular Biology*, Vol.19, No.11, (November 1999), pp. 7886-7896
- Mummenbrauer, T.; Janus, F.; Müller, B.; Wiesmüller, L.; Deppert, W. & Grosse, F. (1996). p53 protein exhibits 3'-to-5' exonuclease activity. *Cell*, Vol.85, No.7, (June 1996), pp. 1089-1099
- Murakami, Y.; Yasuda, H.; Miyazawa, H.; Hanaoka, F. & Yamada, M. (1985). Characterization of a temperature-sensitive mutant of mouse FM3A cells defective in DNA replication. *Proceedings of the National Academy of Science, USA*, Vol.82. No.6, (March 1985), pp. 1761-1765
- Murakami, Y.; Eki, T. & Hurwitz, J. (1992). Studies on the initiation of simian virus 40 replication in vitro: RNA primer synthesis and its elongation. *Proceedings of the National Academy of Science, USA*, Vol.89, No.3, (February 1992), pp. 952-956
- Nakamura, H.; Morita, T.; Masaki, S.; & Yoshida, S. (1984). Intracellular localization and metabolism of DNA polymerase  $\alpha$  in human cells visualized with monoclonal antibody. *Experimental Cell Research*, Vol.151, No.1, (March 1984), pp. 123-133
- Nakamura, M.; Nabetani, A.; Mizuno, T.; Hanaoka, F. & Ishikawa, F. (2005). Alterations of DNA and chromatin structures at telomeres and genetic instability in mouse cells defective in DNA polymerase  $\alpha$ . *Molecular and Cellular Biology*, Vol.25, No.4, (December 2005), pp. 11073-11088
- Namekawa, S.; Hamada, F.; Sawado, T.; Ishii, S.; Nara, T.; Ishizaki, T.; Ohuchi, T.; Arai, T. & Sakaguchi, K. (2003). Dissociation of DNA polymerase  $\alpha$ -primase complex during meiosis in *Coprinus cinereus*. *European Journal of Biochemistry*, Vol.270, No.10, (May 2003), pp. 2137-2146
- Nasheuer, H.-P. & Grosse, F. (1987). Immunoaffinity-purified DNA polymerase  $\alpha$  displays novel properties. *Biochemistry*, Vol.26, No.25, (December 1987), pp. 8458-8466
- Nasheuer, H.-P.; Moore, A.; Wahl, A. F. & Wang, T. S.-F. (1991). Cell cycle-dependent phosphorylation of human DNA polymerase  $\alpha$ . *The Journal of Biological Chemistry*, Vol.266, No.12, (April 1991), pp. 7893-7903
- Niimi, A.; Limsirichaikul, S.; Yoshida, S.; Iwai, S.; Masutani, C.; Hanaoka, F.; Kool, E. T.; Nishiyama, Y. & Suzuki, M. (2004). Palm mutants in DNA polymerases  $\alpha$  and  $\eta$  alter DNA replication fidelity and translesion activity. *Molecular and Cellular Biology*, Vol.24, No.7, (April 2004), pp. 2734-2746
- Nozawa, K.; Suzuki, M.; Takemura, M. & Yoshida, S. (2000). In vitro expansion of mammalian telomere repeats by DNA polymerase  $\alpha$ -primase. *Nucleic Acids Research*, Vol.28, No.16, (August 2000), pp. 3117-3124
- Ogawa, M.; Limsirichaikul, S.; Niimi, A.; Iwai, S.; Yoshida, S. & Suzuki, M. (2003). Distinct function of conserved amino acids in the fingers of *Saccharomyces cerevisiae* DNA polymerase  $\alpha$ . *The Journal of Biological Chemistry*, Vol.278, No.21, (May 2003), pp. 19071-19078
- Olson, M. O. J.; Orrick, L. R.; Jones, C. & Busch, H. (1974). Phosphorylation of acid-soluble nucleolar proteins of Novikoff hepatoma ascites cells in vivo. *The Journal of Biological Chemistry*, Vol.249, No.9, (May 1974), pp. 2823-2827
- Orrick, L. R.; Olson, M. O. J. & Busch, H. (1973). Comparison of nucleolar proteins of normal rat liver and Novikoff hepatoma ascites cells by two-dimensional polyacrylamide

- gel electrophoresis. *Proceedings of the National Academy of Science, USA*, Vol.70, No.5, (May 1973), pp. 1316-1320
- Pavlov, Y. I.; Shcherbakova, P. V. & Rogozin, I. B. (2006a). Roles of DNA polymerases in replication, repair, and recombination in eukaryotes. *International Reviews of Cytology*, Vol.255, pp. 41-132
- Pavlov, Y. I.; Frahm, C.; Nick McElhinny, S. A.; Niimi, A.; Suzuki, M. & Kunkel, T. A. (2006b). Evidence that errors made by DNA polymerase  $\alpha$  are corrected by DNA polymerase  $\delta$ . *Current Biology*, Vol.16, No.2, (January 2006), pp. 202-207
- Popanda, O.; Fox, G. & Thielmann, H. W. (1995). DNA polymerases  $\alpha$ ,  $\delta$ , and  $\epsilon$  of Novikoff hepatoma cells differ from those of normal rat liver in physicochemical and catalytic properties. *Journal of Molecular Medicine*, Vol.73, No.5, (May 1995), pp. 259-268
- Pritchard, C. G. & DePamphilis, M. L. (1983) Preparation of DNA polymerase  $\alpha$ •C<sub>1</sub>C<sub>2</sub> by reconstituting DNA polymerase  $\alpha$  with its specific stimulatory cofactors, C<sub>1</sub>C<sub>2</sub>. *The Journal of Biological Chemistry*, Vol.258, No.16, (August 1983), pp. 9801-9809
- Pritchard, C. G.; Weaver, D. T.; Baril, E. F. & DePamphilis, M. L. (1983). DNA polymerase  $\alpha$  cofactors C<sub>1</sub>C<sub>2</sub> function as primer recognition proteins. *The Journal of Biological Chemistry*, Vol.258, No.16, (August 1983), pp. 9810-9819
- Pursell, Z. F.; Isoz, I.; Lundström, E.-B.; Johansson, E. & Kunkel, T. A. (2007). Yeast DNA polymerase  $\epsilon$  participates in leading-strand DNA replication. *Science*, Vol.317, No.5834, (July 2007), pp. 127-130
- Qi, H. & Zakian, V. A. (2000). The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase  $\alpha$  and the telomerase-associated est1 protein. *Genes & Development*, Vol.14, No.14, (July 2000), pp. 1777-1788
- Rosenblatt J.; Gu, Y. & Morgan, D. O. (1992). Human cyclin-dependent kinase 2 is activated during S and G2 phases of the cell cycle and associates with cyclin A. *Proceedings of the National Academy of Science, USA*, Vol.89, No.7, (April 1992), pp. 2824-2828
- Sawado, T. & Sakaguchi, K. (1997). A DNA polymerase  $\alpha$  catalytic subunit is purified independently from the tissue at meiotic prometaphase I of a basidiomycete, *Coprinus cinereus*. *Biochemical and Biophysical Research Communications*, Vol.232, No.2, (March 1997), pp. 454-460
- Schmidt-Zachmann, M. S.; Hügle-Dörr, B. & Franke, W. W. (1987). A constitutive nucleolar protein identified as a member of the nucleoplasmin family. *The EMBO Journal*, Vol.6, No.7, (July 1987), pp. 1881-1890
- Schub, O.; Rohaly, G.; Smith, R. W. P.; Schneider, A.; Dehde, S.; Dornreiter, I. & Nasheuer, H.-P. (2001). Multiple phosphorylation sites of DNA polymerase  $\alpha$ -primase cooperate to regulate the initiation of DNA replication *in vitro*. *The Journal of Biological Chemistry*, Vol.276, No.41, (October 2001), pp. 38076-38083
- Shevelev, I. V. & Hübscher, U. (2002). The 3'-5' exonucleases. *Nature Reviews Molecular Cell Biology*, Vol.3, No.5, (May 2002), pp. 364-375
- Shioda, M.; Okuhara, K.; Murofushi, H.; Mori, A.; Sakai, H.; Murakami-Murofushi, K.; Suzuki, M. & Yoshida, S. (1991). Stimulation of DNA polymerase  $\alpha$  activity by microtubule-associated proteins. *Biochemistry*, Vol.30, No.48, (December 1991), pp. 11403-11412.

- Simbulan, C. M. G.; Suzuki, M.; Izuta, S.; Sakurai, T.; Savoysky, E.; Kojima, K.; Miyahara, K.; Shizuta, Y. & Yoshida, S. (1993). Poly(ADP-ribose) polymerase stimulates DNA polymerase  $\alpha$  by physical association. *The Journal of Biological Chemistry*, Vol.268, No.1, (January 1993), pp.93-99
- Simbulan, C. M. G.; Taki, T.; Tamiya-Koizumi, K.; Suzuki, M.; Savoysky, E.; Shoji, M. & Yoshida, S. (1994). Sulfate- and sialic acid-containing glycolipids inhibit DNA polymerase  $\alpha$  activity. *Biochimica et Biophysica Acta*, Vol.1205, No.1, (March 1994), pp. 68-74
- Smale, S. T. & Tjian, R. (1986). T-antigen-DNA polymerase  $\alpha$  complex implicated in simian virus 40 DNA replication. *Molecular and Cellular Biology*, Vol.6, No.11, (November 1986), pp. 4077-4087
- Stillman, B. (1989). Initiation of eukaryotic DNA replication in vitro. *Annual Review of Cell Biology*, Vol.5, pp. 197-245
- Suzuki, M.; Izuta, S. & Yoshida, S. (1994). DNA polymerase  $\alpha$  overcomes an error-prone pause site in the presence of replication protein-A. *The Journal of Biological Chemistry*, Vol.269, No.14, (April 1994), pp. 10225-10228
- Takada-Takayama, R.; Hanaoka, F.; Yamada, M. & Ui, M. (1991). Immunoaffinity-purified DNA polymerase  $\alpha$  from a mouse temperature-sensitive mutant, tsFT20 strain, is heat-labile. *The Journal of Biological Chemistry*, Vol.266, No.24, (August 1991), pp. 15716-15718
- Takemura, M.; Ohta, N.; Furuichi, Y.; Takahashi, T.; Yoshida, S.; Olson, M. O. J. & Umekawa, H. (1994). Stimulation of calf thymus DNA polymerase  $\alpha$  activity by nucleolar protein B23. *Biochemical and Biophysical Research Communications*, Vol.199, No.1, (February 1994), pp. 46-51
- Takemura, M.; Kitagawa, T.; Izuta, S.; Wasa, J.; Takai, A.; Akiyama, T. & Yoshida, S. (1997). Phosphorylated retinoblastoma protein stimulates DNA polymerase  $\alpha$ . *Oncogene*, Vol.15, No.20, (November 1997), pp. 2483-2492
- Takemura, M.; Sato, K.; Nishio, M.; Akiyama, T.; Umekawa, H. & Yoshida, S. (1999). Nucleolar protein B23.1 binds to retinoblastoma protein and synergistically stimulates DNA polymerase  $\alpha$  activity. *Journal of Biochemistry*, Vol.125, No.5, (May 1999), pp. 904-909
- Takemura, M. (2001). Poxviruses and the origin of the eukaryotic nucleus. *Journal of Molecular Evolution*, Vol.52, No.5, (May 2001), pp. 419-425
- Takemura, M. (2002) Biochemical properties of the stimulatory activity of DNA polymerase  $\alpha$  by the hyper-phosphorylated retinoblastoma protein. *Biochimica et Biophysica Acta*, Vol.1571, No.2, (June 2002), pp. 151-156
- Takemura, M.; Yoshida, S.; Akiyama, T.; Kitagawa, M. & Yamada, Y. (2006). Role of the second-largest subunit of DNA polymerase  $\alpha$  in the interaction between the catalytic subunit and hyperphosphorylated retinoblastoma protein in late S phase. *Biochimica et Biophysica Acta*, Vol.1764, No.9, (September 2006), pp. 1447-1453
- Takemura, M.; Sugimura, K.; Okumura, K.; Limsirichaikul, S.; Suzuki, M.; Yamada, Y. & Yoshida, S. (2008). Hyper-phosphorylated retinoblastoma protein suppresses telomere elongation. *Bioscience, Biotechnology, and Biochemistry*, Vol.72, No.2, (February 2008), pp. 630-635

- Takemura, M. (2008). Eutherian intrinsically run a higher risk of replication deficiency. *BioSystems*, Vol.92, No.2, (May 2008), pp. 117-121
- Takemura, M.; Kazama, T.; Sakuma, K.; Mizushina, Y. & Oshima, T. (2011). Growth inhibition and chromosomal instability of cultured marsupial (opossum) cells after treatment with DNA polymerase  $\alpha$  inhibitor. *Bioscience, Biotechnology, and Biochemistry*, in press
- Tamai, K.; Kojima, K.; Hanaichi, T.; Masaki, S.; Suzuki, M.; Umekawa, H. & Yoshida, S. (1988). Structural study of immunoaffinity-purified DNA polymerase  $\alpha$ -DNA primase complex from calf thymus. *Biochimica et Biophysica Acta*, Vol.950, No.3, (September 1988), pp. 263-273
- Tanaka, S.; Cao, K.; Niimi, A.; Limsirichaikul, S.; Miao, H. Q.; Nakamura, N.; Murate, T.; Hasegawa, Y.; Takahashi, T. & Suzuki, M. (2010). Functions of base selection step in human DNA polymerase  $\alpha$ . *DNA Repair*, Vol.9, No.5, (May 2010), pp. 534-541
- Toricani, L.; Shanahan, S. & Parry, D. (2009). Replication stress activates DNA polymerase alpha-associated Chk1. *Cell Cycle*, Vol.8, No.3, (February 2009), pp. 482-489
- Tsurimoto, T.; Melendy, T. & Stillman, B. (1990). Sequential initiation of lagging and leading strand synthesis by two different polymerase complexes at the SV40 DNA replication origin. *Nature*, Vol.346, No.6284, (August 1990), pp. 534-539
- Uchiyama, M. & Wang, T. S.-F. (2004). The B-subunit of DNA polymerase  $\alpha$ -primase associates with the origin recognition complex for initiation of DNA replication. *Molecular and Cellular Biology*, Vol.24, No.17, (September 2004), pp. 7419-7434
- Umekawa, H.; Sato, K.; Takemura, M.; Watanabe, Y.; Usui, S.; Takahashi, T.; Yoshida, S.; Olson, M. O. J. & Furuichi, Y. (2001). The carboxyl terminal sequence of nucleolar protein B23.1 is important in its DNA polymerase  $\alpha$ -stimulatory activity. *Journal of Biochemistry*, Vol.130, No.2, (August 2001), pp. 199-205
- Villarreal, L. P. & DeFilippis, V. R. (2000). A hypothesis for DNA viruses as the origin of eukaryotic replication proteins. *Journal of Virology*, Vol.74, No.15, (August 2000), pp. 7079-7084
- Voitenleitner, C.; Rehfuess, C.; Hilmes, M.; O'rear, L.; Liao, P.-C.; Gage, D. A.; Ott, R.; Nasheuer, H.-P. & Fanning, E. (1999). Cell cycle-dependent regulation of human DNA polymerase  $\alpha$ -primase activity by phosphorylation. *Molecular and Cellular Biology*, Vol.19, No.1, (January 1999), pp. 646-656
- Waga, S. & Stillman, B. (1998). The DNA replication fork in eukaryotic cells. *Annual Reviews of Biochemistry*, Vol.67, pp. 721-751
- Wang, T. S.-F. (1991). Eukaryotic DNA polymerases. *Annual Reviews of Biochemistry*, Vol.60, pp. 513-552
- Yagura, T.; Kozu, T. & Seno, T. (1982). Mouse DNA polymerase accompanied by a novel RNA polymerase activity: purification and partial characterization. *Journal of Biochemistry (Tokyo)*, Vol.91, No.2, (February 1982), pp. 607-618
- Yagura, T.; Tanaka, S.; Kozu, T.; Seno, T. & Korn, D. (1983). Tight association of DNA primase with a subspecies of mouse DNA polymerase  $\alpha$ . *The Journal of Biological Chemistry*, Vol.258, No.11, (June 1983), pp. 6698-6700
- Yoshida, S.; Tamai, K.; Umekawa, H.; Suzuki, M. & Kojima, K. (1989). A novel stimulating protein of mammalian DNA polymerase  $\alpha$ . *Journal of Biochemistry*, Vol.106, No.3, (September 1989), pp. 389-395.

- Zheng, L.; Dai, H.; Zhou, M.; Li, M.; Singh, P.; Qui, J.; Tsark, W.; Huang, Q.; Kernstine, K.; Zhang, X.; Lin, D. & Shen, B. (2007). Fen1 mutations result in autoimmunity, chronic inflammation and cancers. *Nature Medicine*, Vol.13, No.7, (July 2007), pp. 812-819
- Zheng, L. & Shen, B. (2011). Okazaki fragment maturation : nucleases take centre stage. *Journal of Molecular Cell Biology*, Vol.3, No.1, (February 2011), pp. 23-30

# Quality Control of DNA Polymerase $\alpha$

Takeshi Mizuno<sup>1</sup>, Masako Izumi<sup>2</sup> and Christian S. Eichinger<sup>3</sup>

<sup>1</sup>*The Cellular Dynamics Laboratory, ASI, RIKEN*

<sup>2</sup>*Radiation Biology Team, RNC, RIKEN*

<sup>3</sup>*Department of Biochemistry, University of Oxford*

<sup>1,2</sup>*Japan*

<sup>3</sup>*United Kingdom*

## 1. Introduction

In the human genome, at least 14 DNA polymerases were identified. The highly conserved DNA polymerase  $\alpha$ -primase complex is the only eukaryotic polymerase that can initiate DNA synthesis *de novo*. Thus, its recruitment is a crucial step in the tightly regulated stepwise assembly of the replication machinery in eukaryotic cells. This complex is required for the synthesis of RNA primers, an essential prerequisite for the initiation of replication, and for the discontinuous synthesis of Okazaki fragments on the lagging strand (Bell and Dutta 2002, Hubscher et al. 2002, Waga and Stillman 1998, Wang 1991). Moreover, DNA polymerase  $\alpha$  plays a fundamental role in coordinating DNA replication, DNA repair and cell cycle progression (Hubscher et al. 2002), in telomere capping and length regulation (Adams Martin et al. 2000, Dahlen et al. 2003, Grossi et al. 2004, Qi and Zakian 2000), and in the epigenetic control of transcriptional silencing and nucleosome reorganization (Nakayama et al. 2001, Zhou and Wang 2004).

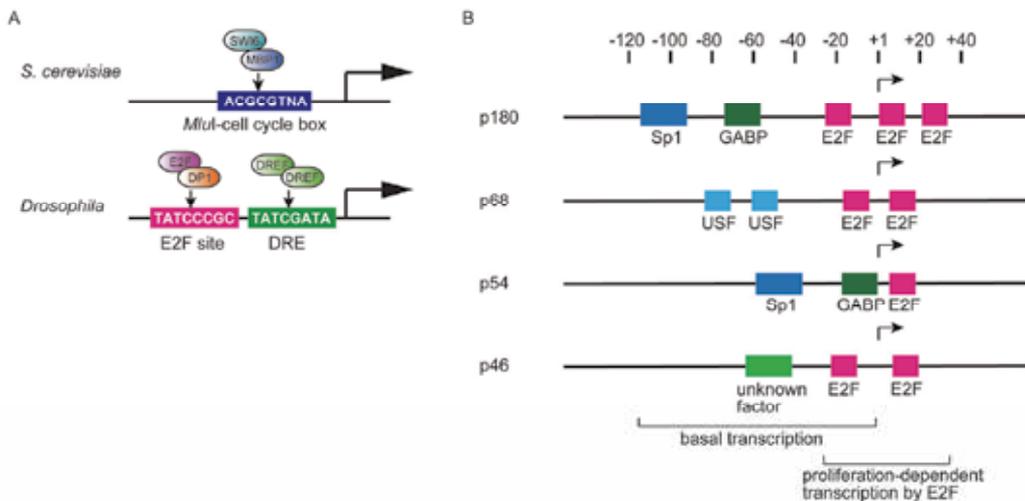
The DNA polymerase  $\alpha$ -primase complex consists of four subunits, each of which is conserved in eukaryotes; in yeast, all four subunits are essential for viability (Wang 1991). The largest subunit, 180 kDa (p180), harbors the catalytic polymerase activity. The two smallest subunits, 54 kDa (p54) and 46 kDa (p46), provide primase activity. The p46 protein, which is coupled to p180 by p54, synthesizes RNA primers and is involved in regulating their length; it also functions in cell cycle checkpoints (Arezi and Kuchta 2000). The 68 kDa subunit (p68) plays a crucial regulatory role in the early stage of chromosomal replication in yeast and has been shown to be essential for the nuclear import of p180 in mouse cells (Foiani et al. 1994, Mizuno et al. 1998).

In this review, we discuss several layers of control that exist to ensure proper functioning of DNA polymerase  $\alpha$  in the cell: regulation of transcription via *cis*- and *trans*-acting elements (part 2), ordered assembly of its four subunits into a functional complex (part 3.1), regulation of nuclear import (part 3.2), co-translational regulation (part 3.3), and quality control mechanisms to exclude aberrant proteins from acting in the nucleus (part 4).

## 2. Transcriptional regulation of DNA polymerase $\alpha$

In eukaryotes, the expression of the genes encoding DNA replication proteins is tightly coupled with cell cycle control. In budding yeast, more than 20 genes involved in DNA

replication and nucleotide metabolism (DNA polymerases, thymidylate synthase, ribonucleotide reductase, DNA topoisomerase II, etc.) show periodic fluctuations in mRNA levels during the cell cycle (McIntosh 1993). These genes have a common *cis*-acting element called the *Mlu*I-cell cycle box (MCB) in their 5' non-transcribed regions (Fig.1A). In late G1 shortly after the execution point, the transcription factor SWI6, which is activated by p34<sup>CDC28</sup>, binds to MCB with MBP1 (*Mlu*I-binding protein 1) to transactivate a set of genes. Also in higher eukaryotes, MCB-like elements control DNA replication-related genes. For instance in *Drosophila*, DNA replication-related elements (DREs) and E2F-binding sites are identified in the promoter regions of DNA replication factors including polymerase  $\alpha$ , PCNA, and Orc2 and involved in proliferation-dependent expression (Matsukage et al. 2008). Although the human homolog of DRE-binding factor (DREF) is reported as a putative positive transcriptional factor, it is unknown whether human DREF is involved in the transcriptional regulation of DNA replication factors. On the other hand, E2F-binding sites are generally found in the upstream regions of DNA replication-related genes as well as proliferation-related genes such as *c-myc*, *N-myc*, and *c-myb* protooncogenes in mammalian cells (Muller and Helin 2000). E2F transcription factors are activated by release from retinoblastoma protein (Rb), which is hyperphosphorylated by cyclin-dependent kinase (CDK) at the G1 to S transition.



A. Swi6 interacts with MBP1 at *Mlu*I-cell cycle box sequence and regulates proliferation-dependent transcription in budding yeast, while DREF homodimer as well as E2Fs are involved in *Drosophila*. B. The promoter regions of the four subunits of mouse DNA polymerase  $\alpha$  are presented. The E2F-binding sites are localized adjacent to the transcription initiation sites and involved in proliferation-dependent transcription, whereas constitutive transcription is controlled by different factors.

Fig. 1. Transcriptional regulation of DNA replication-related genes.

As for DNA polymerase  $\alpha$ , the mRNA and protein levels of all four subunits are coordinately regulated in mammalian cell (Miyazawa et al. 1993). The mRNA levels of quiescent cells (i.e. prior to DNA replication) become 10-fold upregulated when stimulated to proliferate and to enter S phase. We isolated the promoter region of the four subunits of mouse DNA polymerase  $\alpha$  and identified *cis*-acting elements and *trans*-activating factors

(Izumi et al. 2000, Nishikawa N. et al. 2001, Nishikawa N. S. et al. 2000). The mechanism of transcriptional regulation of the four subunits have common features as summarized in Fig. 1B. The upstream sequences of all four DNA polymerase  $\alpha$  subunits exhibit a high G/C content which is similar to many other housekeeping genes and lack TATA or CAAT boxes. In addition, basal transcription and the proliferation-dependent transcription can be distinguished as a common theme for the regulation of each subunit and are controlled by a different set of transcription factors and *cis*-regulatory elements. A unique factor that stimulates proliferation-dependent transcription of all four DNA polymerase  $\alpha$  subunits is E2F. E2F binding sites are located near the transcription initiation sites and 20-90 bp downstream of the *cis*-elements for basal transcription. Other than the common feature of proliferation-dependent transcription of each subunit by E2F, the regulation of basal transcription of each subunit seems to be subunit-specific. Namely, GA-binding protein (GABP), a member of Ets-binding transcription factors, and Sp1 are involved in the transcription of p180 and p54, whereas USF (upstream transcription factor) is involved in the transcription of p68. Interestingly, GABP is shown to be involved in the transcriptional control of Skp2, which is required for degradation of CDK inhibitors including p21 and p27 (Yang et al. 2007). Therefore, GABP may regulate the cell cycle-dependent expression of DNA polymerase  $\alpha$  by a pathway that is distinct from E2F.

### 3. Functional analysis of individual DNA polymerase $\alpha$ subunits

#### 3.1 Complex assembly and domain organization of DNA polymerase $\alpha$

In order to functionally analyse subunit assembly, cDNAs of each DNA polymerase  $\alpha$  subunit were subcloned into mammalian expression plasmids and expressed in cultured mammalian cells. To examine interactions among the four subunits, we transfected four subunits in various combinations, and tested interactions of co-expressed subunits by co-immunoprecipitation analysis (summarized as a cartoon in Fig. 2; Mizuno et al. 1998, 1999). A crucial tool to express different combinations of subunits from a single plasmid was the introduction of IRES (Internal Ribosomal Entry Site) elements, which allowed analysis of subunit interdependency concerning protein expression and subcellular localisation (Fig. 4). Moreover, a set of deletion constructs of p180 revealed that p180 can be divided into three domains (Fig. 2). 1) The amino-terminal (N-terminal) region is not only dispensable for interaction to other subunits but also nonessential for DNA polymerase activity. In addition, it has been reported that the N-terminal domain serves as a binding site for other replication factors (large T antigen, RPA, CDC13, Pol32, Mcm10). 2) The central core-domain is solely responsible for DNA polymerase activity, namely, single-stranded DNA binding activity, and deoxynucleotide triphosphate binding activity as well as phosphoryl-transfer catalytic reaction. 3) The carboxyl terminal (C-terminal) domain is a binding site for other DNA polymerase  $\alpha$  subunits. It contains a conserved cysteine rich region, which resembles most closely the iron-binding motif found in rubredoxin (Klinge et al. 2009). These cluster of cysteines are essential for binding to the second-largest subunit. Primase activity is located in the smallest subunit, p46. This subunit does not directly bind to p180 or p68, but is tethered to p180 via a fourth subunit, p54.

DNA polymerase  $\alpha$  belongs to the B-type ( $\alpha$ -type) DNA polymerase family, and we speculate, that domain arrangement of the largest, catalytical subunit as well as general complex architecture of DNA polymerase  $\alpha$  might well be conserved among members of this family like polymerases  $\alpha$ ,  $\delta$ ,  $\epsilon$  and  $\zeta$  (Fig. 3). All B-type polymerases have a cysteine rich



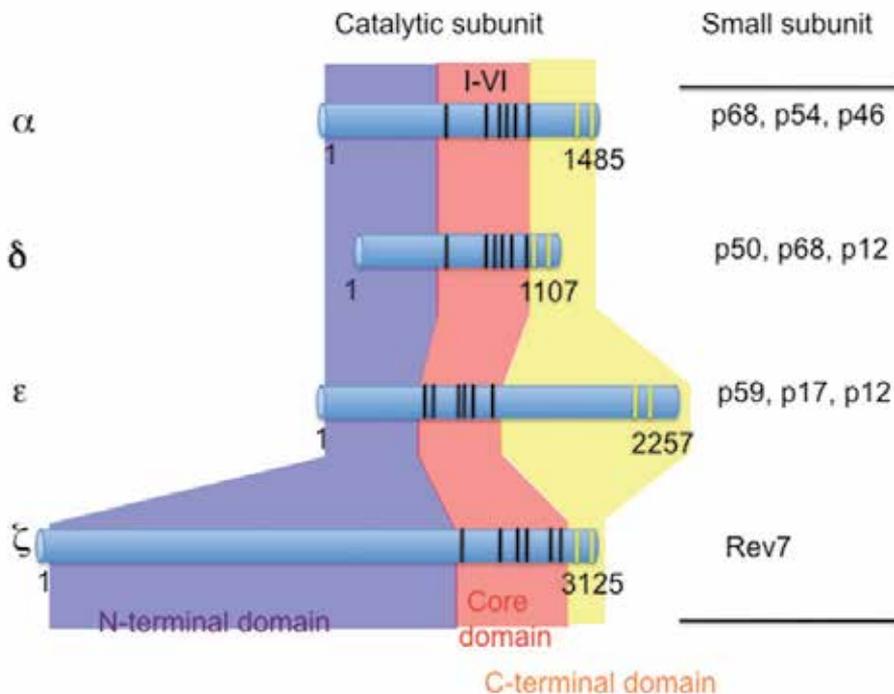


Fig. 3. B-type DNA polymerase family in mammalian cells. Proposed domain organization of catalytic subunits is depicted. Functional domains of catalytic subunit p180 are divided into three domains. Number of amino acids for catalytic subunit and highly conserved polymerase catalytic motifs I-VI (black bars) and cysteine rich motif (yellow bars) are shown.

(Bouvier D and Baldacci, 1995). Also recently, the crystal structure of the C-terminal part of p180 in *S. cerevisiae* suggested that at least in yeast, the C-terminus of yeast p180 is not responsible for nuclear entry (Klinge et al. 2008).

During the course of our studies, we also found a dependency p54 and p46 to enter into the nucleus (Fig. 4; Mizuno et al. 1997). The general idea emerging from these experiments is a necessity of regulated subunit pre-assembly to expose functional NLS and trigger subsequent nuclear translocation. This mechanism might potentially ensure the right protein stoichiometry of nuclear multi-subunit complexes such as DNA polymerase  $\alpha$ .

### 3.3 Co-translational interaction of p68 with p180 enhances p180 protein expression

Our extensive analysis of co-expressing several subunits of polymerase  $\alpha$  also revealed that p180 expression is strictly dependent on co-expressed p68. To clarify the effect of p68 on p180 protein expression, we analyzed transcriptional levels as well as the protein turn-over rate in the presence or absence of p68, and found that p68 affects protein expression rate or protein stability at a post-transcriptional level (as mRNA levels are not affected; Mizuno et al. 1998). This finding suggests an important role of p180-p68 interaction at a co-translational step to ensure full expression of each protein (Fig. 5).

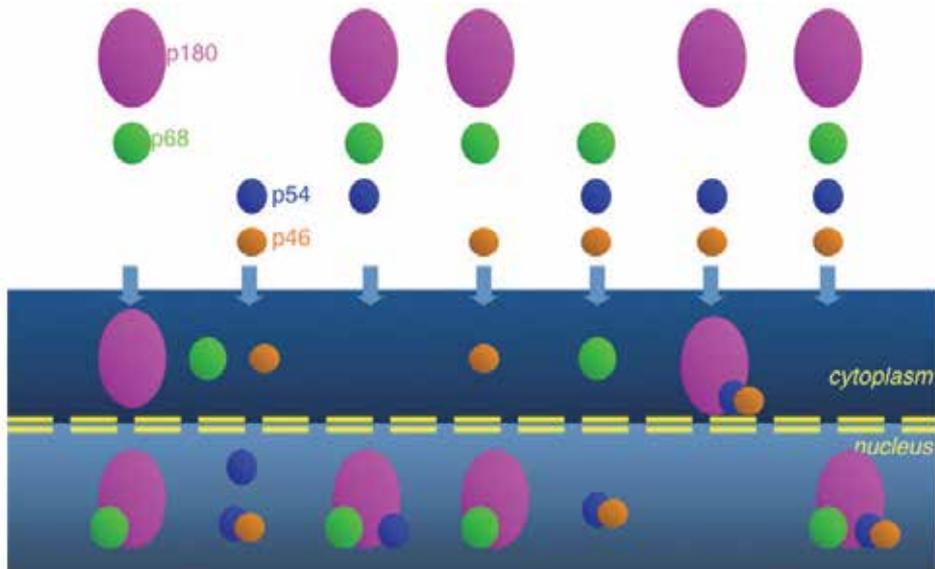


Fig. 4. Nuclear translocation of polymerase  $\alpha$  subunits. Various combination of co-expression of four subunits revealed, that nuclear import of the catalytic subunit of polymerase  $\alpha$  is dependent on accompanying subunits. In summary, p68 or p54 are essential for p180 or p46 nuclear translocation, respectively.

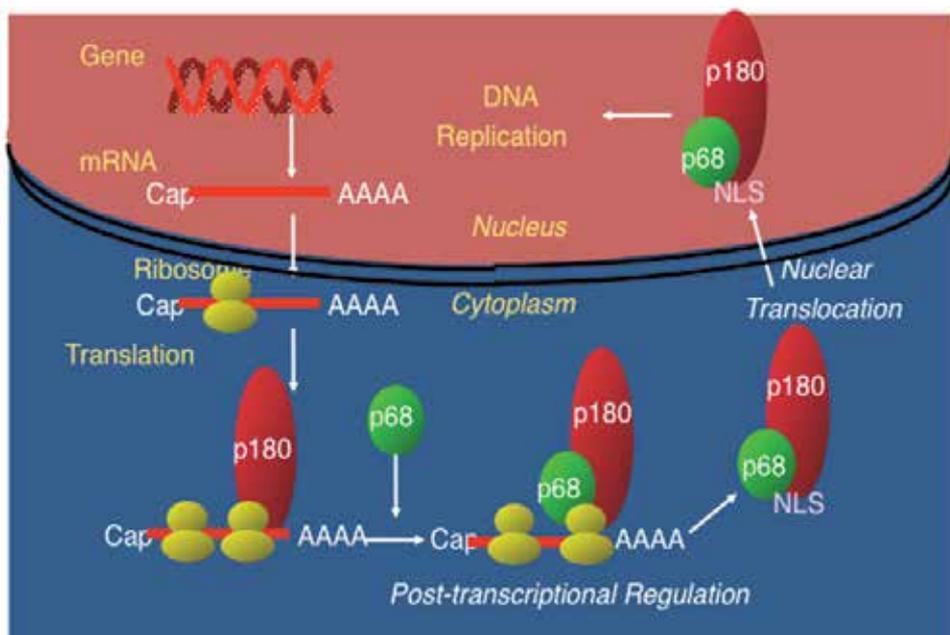


Fig. 5. The second-largest subunit p68 associated with p180 at a post-transcriptional step to enhance translation of p180 and stabilize p180 protein level. In the cytoplasm, free p68 immediately associates with p180 C-terminal region at a co-translational step and it seems association of p68 with p180 may directly stabilize and enhance translation of p180.

#### 4. Quality control of polymerase $\alpha$ in the nucleus

When we found that p68 is important for both p180 protein expression and nuclear translocation, we first could not understand the entire physiological meaning of this observation. Unexpectedly, our investigation using an aberrant mutant version of p180 (p180<sup>tsFT20</sup>) and live cell imaging techniques allowed us to recognize novel protein quality control mechanisms acting on DNA polymerase  $\alpha$  which ensure exclusion of aberrant forms of p180 from the nucleus (Eichinger et al. 2009). In general, misfolded and non-functional proteins must be degraded to assure the correct functioning of cellular processes. Several mechanisms that ensure the removal of defective proteins have been described, such as the unfolded protein response and ER-associated degradation (ERAD); the latter process governs the degradation of unfolded ER proteins in the cytoplasm (Richly et al. 2005, Ye et al. 2001). Degradation systems that function in quality control have been identified in the secretory pathway and in mitochondria (Goldberg 2003). These processes are carried out by proteasomes, organelles consisting of several proteases that are located in the cytoplasm as well as in the nucleus (Bennett et al. 2005, Hershko and Ciechanover 1998). Proteasomes function in regulatory processes, allowing cells to balance protein expression and to eliminate misfolded proteins that are useless or even harmful (Hershko and Ciechanover 1998). In contrast, the degradation-dependent protein quality control mechanisms that act in the nucleus are still poorly understood. Although several proteins such as the transcription factors p53 and Myc as well as the replication-associated factors Orc1 and Cdt1 are degraded in the nucleus by a proteasome-dependent pathway over the course of the cell cycle (Nishitani and Lygerous. 2002), the principles underlying the degradation of damaged nuclear proteins have not been uncovered to date. Recently, Gardner et al. reported that San1-mediated degradation acts as a protein quality control system in *Saccharomyces cerevisiae* nuclei (Gardner et al. 2005, Rosenbaum JC, et al. 2011). However, it is still unclear how aberrant nuclear proteins are recognized by E3 proteins and whether analogous systems exist in higher eukaryotes.

##### 4.1 tsFT20 is a model system of aberrant p180 in the cell

The tsFT20 cell line, a temperature-sensitive mutant clone identified in a screen of N-methyl-N'-nitro-N-nitrosoguanidine-treated mouse mammary carcinoma FM3A cells, has been the subject of many studies during the last 30 years (Fig.6; Eki et al. 1986, Eki et al. 1987a, 1988, Eki et al. 1987b, Eki et al. 1990, Ikehata 1994, Ikehata et al. 1997, Izumi et al. 1994, Miyazawa et al. 1986, Murakami et al. 1985, Murakami et al. 1986, Pendergrass et al. 1994, Takada-Takayama et al. 1991, Yamaguchi et al. 1995). Compared to parental FM3A cells, this cell line grows normally at 33°C. However at the restrictive temperature of 39.5°C, they exhibit a phenotype closely resembling that conferred by arrest at the G1/S boundary. At restrictive temperature, tsFT20 cells were further characterized as defective in DNA replication, with a highly decreased DNA synthesis rate and a reduced frequency of replicon initiation. Purified DNA polymerase  $\alpha$  from tsFT20 cells was found to be temperature-sensitive although composed of the same hetero-tetrameric complex as that of FM3A cells, and this defect was ascribed to a single point mutation that changes amino acid 1180 of the p180 subunit from serine to phenylalanine (Izumi et al. 1994). Compared to the wild-type form, the heat-labile DNA polymerase activity of the mutant protein exhibits altered properties with respect to optimal pH and KCl concentration, as well as changes in the dCTP-binding site (Izumi et al. 1994, Takada-Takayama et al. 1991). Furthermore, tsFT20 cells exhibit

extensive chromosomal aberrations after incubation from two hours to 4 hours at restrictive temperature, due to the point mutation. The abnormal cessation of DNA replication in tsFT20 cells at restrictive temperature results in cell death via the induction of DNA double-strand breaks, suggesting that aberrant DNA polymerase  $\alpha$  is cytotoxic and that a specific quality control mechanism for DNA polymerase  $\alpha$  may be crucial to ensure genomic stability and accurate DNA replication under normal as well as stress conditions in a cell.

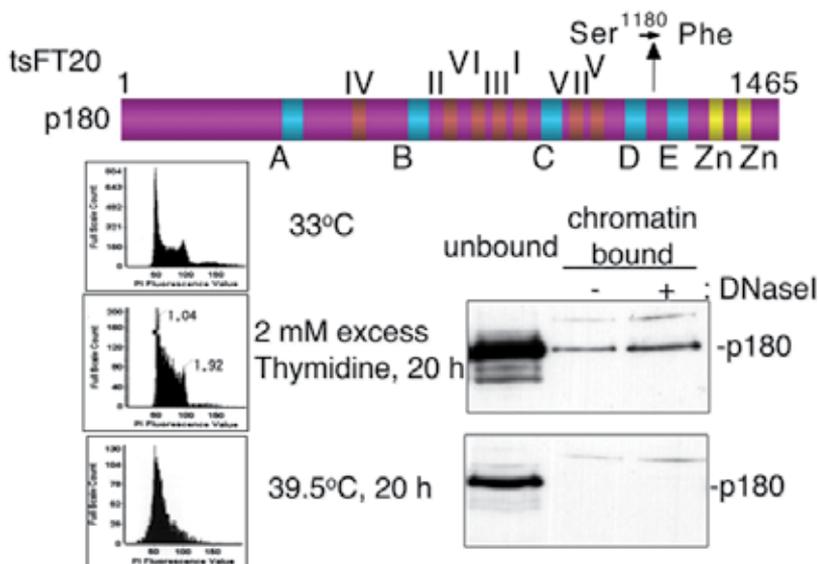


Fig. 6. tsFT20 is a temperature-sensitive mutant cell line derived from mouse FM3A cells. In tsFT20 cells, the DNA polymerase  $\alpha$  subunit p180 contains a point mutation altering serine at amino acid position 1180 to phenylalanine. Due to this mutation, tsFT20 cells arrest at the G1/S transition at the restricted temperature (39.5°C) as shown in FACS analysis with total protein levels and chromatin bound p180 gradually decreasing as shown by Western blot analysis.

#### 4.2 Exclusion of aberrant p180 from the nucleus due to defective complex assembly

To test localization of aberrant nuclear proteins and monitor their sub-cellular distribution directly, we constructed p180 or p180<sup>tsFT20</sup> with GFP. At permissive temperature (33°C), both wild-type p180GFP and mutant p180<sup>tsFT20</sup>GFP clearly localized in the nucleus of mouse NIH3T3 cells in the presence of ectopically over-expressed p68 protein (Fig. 7A). However, at restrictive temperature (39.5°C) in the presence of p68, the wild-type protein remained in the nucleus, whereas the mutant form was found exclusively in the cytoplasm (Fig. 7B). By using co-immunoprecipitation analysis, we found that at restrictive temperature, p180<sup>tsFT20</sup> cannot fully associate with p68, presumably due to a conformational change. Thus, we concluded that cytoplasmic expression of p180<sup>tsFT20</sup> at restrictive temperature is caused by its deficiency to bind p68. Moreover, knock down p68 by siRNA results in expression of endogenous p180 exclusively in the cytoplasm. Taken together, both unbalanced subunit composition and defective interaction of cognate subunits cause nuclear protein sequestration in the cytoplasm.

This finding provides a novel principle of a quality control mechanism to ensure exclusion of misfolded proteins from nucleus. Very recently, Boulon *et al.* reported that RNA polymerase II subunits Rpb1 and Rpb3 accumulate in the cytoplasm when assembly is prevented (Boulon S, et al. 2011). This illustrates a similar example of cytoplasmic sequestration of subunits (of a complex normally acting in the nucleus) that are not properly incorporated into a functional complex or sub-complex. We speculate that subunit-subunit interactions within a complex or pre-assembled sub-complex play a pivotal role to confirm their correct folding state, ability to interact and stoichiometry. Inability of subunits to form pre-assemblies in the right ratio, will mark them as damaged or misfolded proteins and potentially exclude them from the cellular environment (e.g. the nucleus) where they usually act.

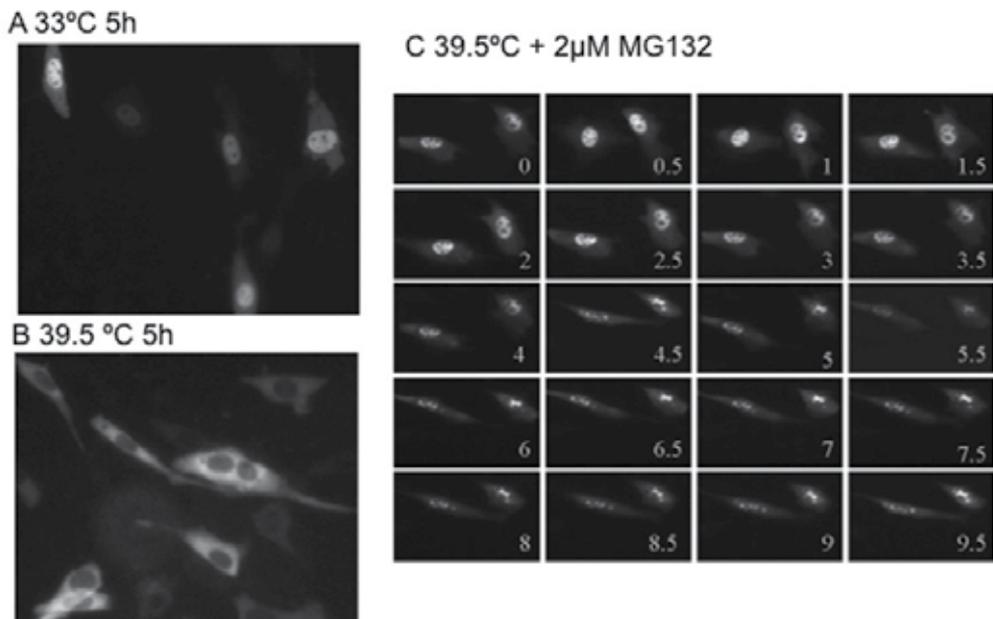


Fig. 7. Subcellular distribution of p180<sup>tsFT20</sup>GFP. p180<sup>tsFT20</sup>GFP is expressed in NIH3T3 cells at permissive temperature for 2 days. Afterwards, cells were incubated for further 5 hours either at permissive temperature (A) or at restrictive temperature (B). In the presence MG132, cells transfected with p180<sup>tsFT20</sup>GFP were incubated at restrictive temperature for the indicated hours in a live cell time-course (C).

#### 4.3 Degradation of aberrant proteins in the nucleus by ubiquitin-mediated proteasome system

In addition to cytoplasmic sequestration, we found that nuclear degradation of misfolded proteins is another way to exclude aberrant proteins from the nucleus. Upon temperature shift to restrictive temperature, nuclear localized p180<sup>tsFT20</sup>GFP is rapidly removed from the nucleus, a process that can be blocked by addition of MG132, a specific proteasome inhibitor (Fig. 7C; Eichinger, et al. 2009). Thus, we concluded that aberrant p180 is degraded in the nucleus in a proteasome-dependent manner. Taken together, we suggest that the turnover of p180<sup>tsFT20</sup>GFP after a temperature shift to 39.5°C may be separated into distinct processes:

first, at restrictive temperature, the mutant protein is rapidly synthesized *de novo* in the cytoplasm but is unable to enter the nucleus due to a mutation-specific, temperature-sensitive effect that affects binding to the second largest subunit p68. Second, nuclear-localized protein is degraded in a proteasome-dependent manner in the nucleus and to a minor extent in the cytoplasm after export from the nucleus (Fig. 8; Eichinger, et al. 2009). Very recently it has been reported that misfolded nuclear proteins are degraded by a unique E3 ubiquitin ligase in *S. cerevisiae*, San1 (Gardner et al. 2005; Rosenbaum JC, et al. 2011). However, the exact mechanism for such a pathway to control the quality of nuclear proteins is not known in yeast and conserved mechanisms are yet to be identified in higher eukaryotes. Given, the importance of ensuring that only properly folded and assembled multi-subunit complexes can enter and act in the nucleus; this field of biology provides many exciting open questions.

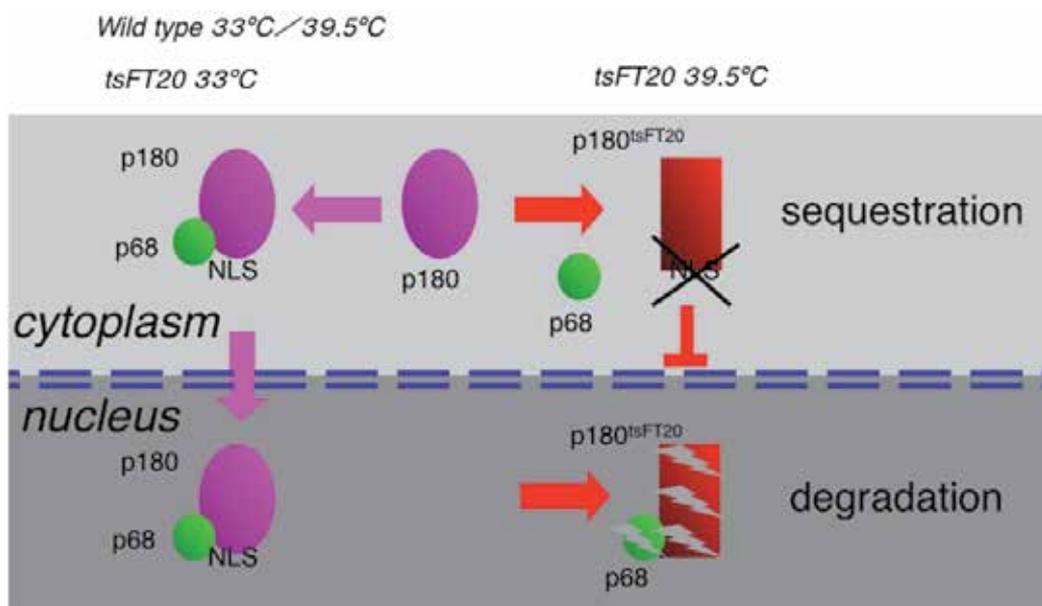


Fig. 8. Two independent quality control mechanisms to exclude aberrant proteins from the nucleus: First, sequestration of aberrant proteins in the cytoplasm, and second, degradation of aberrant proteins through ubiquitin-mediated degradation pathway in the nucleus.

## 5. Conclusion

DNA polymerase  $\alpha$ , a member of the very conserved B-type polymerase family, is a crucial factor to initiate DNA replication and therefore its activity has to be tightly regulated at several stages of its life cycle. Important steps of regulation include basal- and proliferation-dependent control of transcription by *cis*-acting elements and *trans*-activating factors, coordinated complex assembly, regulation of nuclear import, co-translational control of subunit expression as well as quality control mechanisms that ensure exclusion of aberrant proteins acting in the nucleus. These principles to keep a check on such central cellular regulators like DNA polymerase  $\alpha$ , might be a common feature shared with other B-type polymerases like DNA polymerases  $\delta$ ,  $\epsilon$  and  $\zeta$ .

## 6. Acknowledgment

This work was supported by grants from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and an incentive research of RIKEN. CSE was supported by a Marie Curie Intra-European Fellowship. The authors would appreciate to the valuable encouragement of Prof. F. Hanaoka and Prof. H. Miyazawa and Prof. T. Eki.

## 7. References

- Adams Martin A, Dionne I, Wellinger RJ, Holm C. 2000. The function of DNA polymerase  $\alpha$  at telomeric G tails is important for telomere homeostasis. *Mol Cell Biol* 20: 786-796.
- Arezi B, Kuchta RD. 2000. Eukaryotic DNA primase. *Trends Biochem Sci* 25: 572-576.
- Bell SP, Dutta A. 2002. DNA replication in eukaryotic cells. *Annu Rev Biochem* 71: 333-374.
- Bennett EJ, Bence NF, Jayakumar R, Kopito RR. 2005. Global impairment of the ubiquitin-proteasome system by nuclear or cytoplasmic protein aggregates precedes inclusion body formation. *Mol Cell* 17: 351-365.
- Boulon S, et al. 2011. HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol Cell* 39: 912-924.
- Bouvier D, Baldacci G. 1995. The N-terminus of fission yeast DNA polymerase alpha contains a basic pentapeptide that acts in vivo as a nuclear localization signal. *Mol Biol Cell*. 6:1697-1705.
- Dahlen M, Sunnerhagen P, Wang TS. 2003. Replication proteins influence the maintenance of telomere length and telomerase protein stability. *Mol Cell Biol* 23: 3031-3042.
- Eichinger CS, Mizuno T, Mizuno K, Miyake Y, Yanagi K, Imamoto N, Hanaoka F. 2009. Aberrant DNA polymerase alpha is excluded from the nucleus by defective import and degradation in the nucleus. *J Biol Chem* 284: 30604-30614.
- Eki T, Murakami Y, Enomoto T, Hanaoka F, Yamada M. 1986. Characterization of DNA replication at a restrictive temperature in a mouse DNA temperature-sensitive mutant, tsFT20 strain, containing heat-labile DNA polymerase  $\alpha$  activity. *J Biol Chem* 261: 8888-8893.
- Eki T, Enomoto T, Murakami Y, Hanaoka F, Yamada M. 1987a. Characterization of chromosome aberrations induced by incubation at a restrictive temperature in the mouse temperature-sensitive mutant tsFT20 strain containing heat-labile DNA polymerase alpha. *Cancer Res* 47: 5162-5170.
- Eki T, Enomoto T, Murakami Y, Hanaoka F, Yamada M. 1988. Characterization of DNA polymerase alpha activity from a mouse DNA temperature-sensitive mutant, strain tsFT20, which shows a defect in DNA polymerase alpha activity at restrictive temperatures. *Arch Biochem Biophys* 260: 552-560.
- Eki T, Enomoto T, Murakami Y, Miyazawa H, Hanaoka F, Yamada M. 1987b. Characterization of revertants derived from a mouse DNA temperature-sensitive mutant strain, tsFT20, which contains heat-labile DNA polymerase alpha activity. *Exp Cell Res* 171: 24-36.
- Eki T, Enomoto T, Miyajima A, Miyazawa H, Murakami Y, Hanaoka F, Yamada M, Ui M. 1990. Isolation of temperature-sensitive cell cycle mutants from mouse FM3A cells. Characterization of mutants with special reference to DNA replication. *J Biol Chem* 265: 26-33.

- Foiani M, Marini F, Gamba D, Lucchini G, Plevani P. 1994. The B subunit of the DNA polymerase  $\alpha$ -primase complex in *Saccharomyces cerevisiae* executes an essential function at the initial stage of DNA replication. *Mol Cell Biol* 14: 923-933.
- Gardner RG, Nelson ZW, Gottschling DE. 2005. Degradation-mediated protein quality control in the nucleus. *Cell* 120: 803-815.
- Goldberg AL. 2003. Protein degradation and protection against misfolded or damaged proteins. *Nature* 426: 895-859.
- Grossi S, Puglisi A, Dmitriev PV, Lopes M, Shore D. 2004. Pol12, the B subunit of DNA polymerase alpha, functions in both telomere capping and length regulation. *Genes Dev* 18: 992-1006.
- Hershko A, Ciechanover A. 1998. The ubiquitin system. *Annu Rev Biochem* 67: 425-479.
- Hubscher U, Maga G, Spadari S. 2002. Eukaryotic DNA polymerases. *Annu Rev Biochem* 71: 133-163.
- Ikehata H. 1994. Comparison of DNA polymerases alpha, delta, and epsilon of mouse cell line FM3A and its temperature-sensitive mutant tsFT20. *Tohoku J Exp Med* 172: 65-81.
- Ikehata H, Kaneda S, Yamao F, Seno T, Ono T, Hanaoka F. 1997. Incubation at the nonpermissive temperature induces deficiencies in UV resistance and mutagenesis in mouse mutant cells expressing a temperature-sensitive ubiquitin-activating enzyme (E1). *Mol Cell Biol* 17: 1484-1489.
- Izumi M, Miyazawa H, Harakawa S, Yatagai F, Hanaoka F. 1994. Identification of a point mutation in the cDNA of the catalytic subunit of DNA polymerase  $\alpha$  from a temperature-sensitive mouse FM3A cell line. *J Biol Chem* 269: 7639-7644.
- Izumi M, Yokoi M, Nishikawa NS, Miyazawa H, Sugino A, Yamagishi M, Yamaguchi M, Matsukage A, Yatagai F, Hanaoka F. 2000. Transcription of the catalytic 180-kDa subunit gene of mouse DNA polymerase  $\alpha$  is controlled by E2F, an Ets-related transcription factor, and Sp1. *Biochim Biophys Acta* 1492: 341-352.
- Klinge S, Nunez-Ramirez R, Llorca O, Pellegrini L. 2009. 3D architecture of DNA Pol alpha reveals the functional core of multi-subunit replicative polymerases. *Embo J* 28: 1978-1987.
- Mäkinieniemi M, Pospiech H, Kilpeläinen S, Jokela M, Vihinen M, Syväoja JE. 1999. A novel family of DNA-polymerase-associated B subunits. *Trends Biochem Sci*. 24:14-16.
- Matsukage A, Hirose F, Yoo MA, Yamaguchi M. 2008. The DRE/DREF transcriptional regulatory system: a master key for cell proliferation. *Biochim Biophys Acta* 1779: 81-89.
- McIntosh EM. 1993. MCB elements and the regulation of DNA replication genes in yeast. *Curr Genet* 24: 185-192.
- Miyazawa H, Tandai M, Hanaoka F, Yamada M, Hori T, Shimizu K, Sekiguchi M. 1986. Identification of a DNA segment containing the human DNA polymerase alpha gene. *Biochem Biophys Res Commun* 139: 637-643.
- Miyazawa H, Izumi M, Tada S, Takada R, Masutani M, Ui M, Hanaoka F. 1993. Molecular cloning of the cDNAs for the four subunits of mouse DNA polymerase alpha-primase complex and their gene expression during cell proliferation and the cell cycle. *J Biol Chem* 268: 8111-8122.
- Mizuno T, Yamagishi K, Miyazawa H, Hanaoka F. 1999. Molecular architecture of the mouse DNA polymerase  $\alpha$ -primase complex. *Mol Cell Biol* 19: 7886-7896.

- Mizuno T, Ito N, Yokoi M, Kobayashi A, Tamai K, Miyazawa H, Hanaoka F. 1998. The second-largest subunit of the mouse DNA polymerase  $\alpha$ -primase complex facilitates both production and nuclear translocation of the catalytic subunit of DNA polymerase  $\alpha$ . *Mol Cell Biol* 18: 3552-3562.
- Mizuno T, Okamoto T, Yokoi M, Izumi M, Kobayashi A, Hachiya T, Tamai K, Inoue T, Hanaoka F. 1996. Identification of the nuclear localization signal of mouse DNA primase: nuclear transport of p46 subunit is facilitated by interaction with p54 subunit. *J Cell Sci* 109 : 2627-2636.
- Muller H, Helin K. 2000. The E2F transcription factors: key regulators of cell proliferation. *Biochim Biophys Acta* 1470: M1-12.
- Murakami Y, Yasuda H, Miyazawa H, Hanaoka F, Yamada M. 1985. Characterization of a temperature-sensitive mutant of mouse FM3A cells defective in DNA replication. *Proc Natl Acad Sci U S A* 82: 1761-1765.
- Murakami Y, Eki T, Miyazawa H, Enomoto T, Hanaoka F, Yamada M. 1986. Further characterization of a murine temperature-sensitive mutant, tsFT20 strain, containing heat-labile DNA polymerase alpha-activity. *Exp Cell Res* 163: 135-142.
- Nakayama J, Allshire RC, Klar AJ, Grewal SI. 2001. A role for DNA polymerase  $\alpha$  in epigenetic control of transcriptional silencing in fission yeast. *Embo J* 20: 2857-2866.
- Nishikawa N, Izumi M, Yokoi M, Miyazawa H, Hanaoka F. 2001. E2F regulates growth-dependent transcription of genes encoding both catalytic and regulatory subunits of mouse primase. *Genes Cells* 6: 57-70.
- Nishikawa NS, Izumi M, Uchida H, Yokoi M, Miyazawa H, Hanaoka F. 2000. Cloning and characterization of the 5'-upstream sequence governing the cell cycle-dependent transcription of mouse DNA polymerase alpha 68 kDa subunit gene. *Nucleic Acids Res* 28: 1525-1534.
- Nishitani H, Lygerou Z. 2002. Control of DNA replication licensing in a cell cycle. *Genes Cells*. 7:523-534.
- Pendergrass WR, Saulewicz AC, Hanaoka F, Norwood TH. 1994. Murine temperature-sensitive DNA polymerase alpha mutant displays a diminished capacity to stimulate DNA synthesis in senescent human fibroblast nuclei in heterokaryons at the nonpermissive condition. *J Cell Physiol* 158: 270-276.
- Qi H, Zakian VA. 2000. The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase  $\alpha$  and the telomerase-associated est1 protein. *Genes Dev* 14: 1777-1788.
- Richly H, Rape M, Braun S, Rumpf S, Hoegel C, Jentsch S. 2005. A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting. *Cell* 120: 73-84.
- Rosenbaum JC, Fredrickson EK, Oeser ML, Garrett-Engele CM, Locke MN, Richardson LA, Nelson ZW, Hetrick ED, Milac TI, Gottschling DE, Gardner RG. 2011. Disorder targets misorder in nuclear quality control degradation: a disordered ubiquitin ligase directly recognizes its misfolded substrates. *Mol Cell* 41: 93-106.
- Takada-Takayama R, Hanaoka F, Yamada M, Ui M. 1991. Immunoaffinity-purified DNA polymerase  $\alpha$  from a mouse temperature-sensitive mutant, tsFT20 strain, is heat-labile. *J Biol Chem* 266: 15716-15718.
- Waga S, Stillman B. 1998. The DNA replication fork in eukaryotic cells. *Annu Rev Biochem* 67: 721-751.

- Wang TS. 1991. Eukaryotic DNA polymerases. *Annu Rev Biochem* 60: 513-552.
- Wong SW, Wahl AF, Yuan PM, Arai N, Pearson BE, Arai K, Korn D, Hunkapiller MW, Wang TS. 1988. Human DNA polymerase alpha gene expression is cell proliferation dependent and its primary structure is similar to both prokaryotic and eukaryotic replicative DNA polymerases. *Embo J* 7: 37-47.
- Yamaguchi K, Kobayashi M, Yamauchi Y, Tanaka A, Miyazawa H, Izumi M, Hanaoka F. 1995. Temperature-sensitive mutation of DNA polymerase alpha induces growth-suppressive phenotypes involving retinoblastoma protein and cyclin D1. *Cell Struct Funct* 20: 285-291.
- Yang ZF, Mott S, Rosmarin AG. 2007. The Ets transcription factor GABP is required for cell-cycle progression. *Nat Cell Biol* 9: 339-346.
- Ye Y, Meyer HH, Rapoport TA. 2001. The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol. *Nature* 414: 652-656.
- Zhou Y, Wang TS. 2004. A coordinated temporal interplay of nucleosome reorganization factor, sister chromatin cohesion factor, and DNA polymerase  $\alpha$  facilitates DNA replication. *Mol Cell Biol* 24: 9568-9579.

# Mechanisms and Controls of DNA Replication in Bacteria

César Quiñones-Valles, Laura Espíndola-Serna  
and Agustino Martínez-Antonio  
*Departamento de Ingeniería Genética, Cinvestav Irapuato  
México*

## 1. Introduction

DNA is the polymeric molecule that contains all the genetic information in a cell. This genetic information encodes the instructions to make a copy of itself, for the cellular structure, for the operative cellular machinery and also contains the regulatory signals, which determine when parts of this machinery should be *on* or *off*. The operative machinery in turn, is responsible for the cells functions either metabolically or in interactions with the environment. Part of this cellular machinery devoted to DNA metabolism is responsible for DNA replication, DNA-repair and for the regulation of gene expression. In this chapter we will focus our discussion on the mechanisms and controls that conduct DNA replication in bacteria, including the components, functions and regulation of replication machinery. Most of our discourse will consider this biological process in *Escherichia coli* but when possible we will compare it to other bacterial models, mainly *Bacillus subtilis* and *Caulobacter crescentus* as examples of organisms with asymmetrical cell division.

In order to maintain a bacterial population it is necessary that cells divide, but before the physical division of a daughter cell from its mother, it is necessary among other check points, that the DNA has been replicated accurately. This is done by the universal semi-conservative replication process of DNA-strands, which generates two identical strand copies from their parent templates. To better understand this process it has been divided into three phases: initiation, elongation and termination of DNA replication. In each of these steps, multiple stable and transient interactions are involved and we have summarized them below.

## 2. Components and mechanisms of the general process of DNA replication

Bacteria are subject to sudden changes in their surroundings, so they have adapted diverse strategies to allow them to persist through time. One of the adaptive changes consists in modifying growth rates, which is accompanied by adjusting mechanisms that control the timing of the cell-cycle. This adjustment ensures that the process of cell division is coordinated with the doubling of cell-mass and with the proper replication and segregation of the chromosome. The study of the cell-cycle in bacteria is usually divided into three stages: the period between cell-division (cell birth) and the initiation of chromosome replication, the period required to complete DNA replication (elongation of DNA) and, the

final phase, which goes from the end of DNA replication until the completion of cell-division (Wang & Levine, 2009).

Under the best growing-conditions, DNA replication starts immediately after cell division in most cells (Wang et al., 2005). Since replication of the chromosome takes more time than that the necessary for cell division under optimal culture conditions, such as *E. coli* growing in rich media, at 37°C with good aeration, it can happen that more than one event of DNA replication can occur per cell cycle (Zakrzewska-Czerwinska et al., 2007). For the purposes of this work we shall divide the DNA replication process in bacteria into three steps: initiation, elongation and termination as follows.

## 2.1 Initiation of DNA replication

In bacteria, the process of DNA replication initiates in a specific DNA region called “origin of replication” (*ori*) where multi-protein complexes are positioned and recruits additional initiator proteins to form the Pre-Replicative complex (pre-RC) whose main function is to facilitate the aperture of duplex DNA to permit the loading of the replicative DNA helicase. The activity of this DNA helicase assists the entrance and assembly of a large multi-subunit molecular machine, the replisome (Zakrzewska-Czerwinska et al., 2007; Ozaki & Katayama, 2009).

### 2.1.1 *oriC* and its *cis* regulatory regions

The origin of replication in *E. coli* (*oriC*) is a small DNA sequence of about 245 bp (Figure 1), which contains three AT-rich repeats named L, M, and R for left, middle and right positions respectively, each 13 bp long (Hwang & Kornberg, 1992). The *oriC* region also contains multiple boxes of 9 bp each where DnaA (replication initiation factor) proteins bind. These DnaA boxes recruit DnaA in two forms; DnaA-ATP and DnaA-ADP, although they show more affinity for the first form, which is the active replication initiation complex of DnaA. There are three DnaA-boxes of high affinity named R1, R2 and R4 and seven of low affinity (I1, I2, I3,  $\tau$ 1,  $\tau$ 2, R5M and R3), (Katayama et al., 2010; Ozaki & Katayama, 2009). The *oriC* region also contains GATC DNA motifs dispersed throughout, the GATC motif is recognized as a target for DNA-methylation by the Dam enzyme (DNA adenine methyltransferase). Finally, the *oriC* region also has DNA-binding sites for the union of several regulatory proteins such as Fis (Factor for inversion stimulation) and IHF (integration host factor), which assist in bending the DNA at this region (Leonard & Grimwade, 2009).

The comparison of the DNA sequence used as origin of replication in *E. coli* versus genomes of other sequenced bacteria indicates that the nucleotide composition and size of these regions is similar (Bramhill & Kornberg, 1988). A database of *ori* regions in bacterial genomes, the DoriC database, which contains a compilation of known and predicted DNA origins of replication in bacteria has been developed (Gao & Zhang, 2007).

### 2.1.2 DnaA is the key protein required to form the pre-RC

The critical step for the successful replication of DNA is the unfolding of the DNA strands at the *oriC* region, action that is assisted by the orisome (proteins-*oriC* complex) (Leonard & Grimwade, 2005). This complex mainly comprises of the activity of the initiator protein DnaA. This protein belongs to the ubiquitous AAA<sup>+</sup> superfamily of ATPases (ATPases

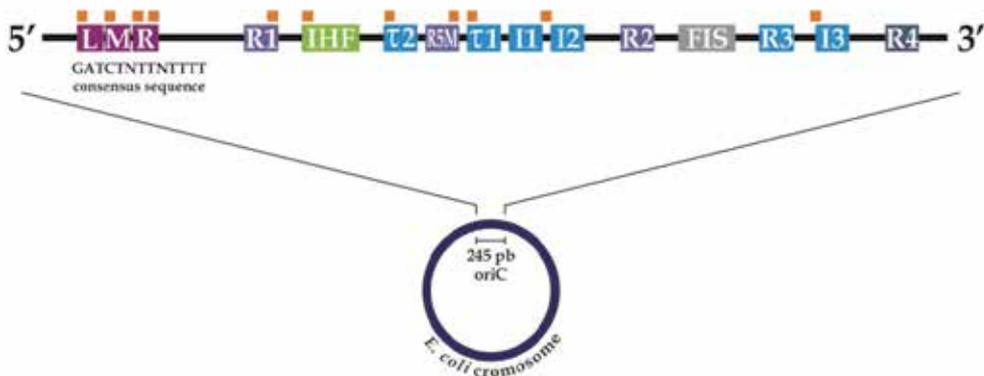


Fig. 1. Description of the origin of replication in the *E. coli* chromosome. The origin of chromosomal replication (*oriC*) contains three AT-rich repeats (L, M and R), each 13 nucleotide residues long and multiple DnaA-binding sites. There are three higher-affinity DnaA-boxes R1, R2 and R4 (dark blue) and seven lower-affinity sites  $\tau$ 1,  $\tau$ 2, I1, I2, I3, R5M and R3 (light blue). All the DnaA-boxes preferentially bind DnaA-ATP rather than DnaA-ADP complexes. *oriC* also contains one site where IHF binds (green), one for Fis (gray) and GATC sites (orange) which are recognized by the Dam enzyme.

associated with a variety of cellular activities). The X-ray structure of crystals of this protein from *Aquifex aeolicus* shows that the protein has four distinctive domains (Erzberger et al., 2002). Domain I serves for the interaction with other proteins, among those identified are: the replicative DnaB helicase and the DnaA-binding assistance protein DiaA (DnaA-initiator association). Domain II is a flexible linker, which provides free rotation for the adjacent domains III and I. Domain III has typical motifs that are characteristic of the AAA<sup>+</sup> protein superfamily of ATPases characterized by a conserved nucleotide phosphate-binding motif, named Walker A (GxxxxGK[S/T]), where x is any amino acid residue). This domain serves in protein binding to either ATP or ADP. When DnaA binds ATP it can form multimeric structures each consisting of 5–7 protomers (DnaA-ATP) by interactions of one subunit with the ATP of the anterior subunit through their “arginine fingers” as shown in Figure 2. It is suggested that the DnaA-*oriC* complex forms a circular pentamer, which is stabilized by interactions between each DnaA unit as mentioned before. The formation of these complexes promotes the unwinding of DNA strands on the initiation of replication. Finally, domain IV of DnaA has a helix-turn-helix motif that allows it to interact with the DnaA-box of *oriC* (Figure 2), (Erzberger et al., 2002; Ozaki & Katayama, 2009).

### 2.1.3 Additional components of the orisome

There are additional components of the orisome that may increase or impede the further unfolding of DNA at the origin of DNA replication. Some of these proteins in *E. coli* include the histone-like DNA-binding proteins IHF and Fis. IHF is a protein that binds to DNA at a poorly defined sequence. It stimulates the initiation of replication *in vivo* and *in vitro*. IHF assists the binding of DnaA to the low-affinity DnaA-boxes during the formation of the pre-replicative-complex. Contrarily, Fis seems to act as a repressor of initiation of DNA replication by inhibiting the binding of DnaA and IHF to their targets sites on DNA. This is achieved because Fis binds to *oriC* in a specific region of 13 nucleotides from position 87 to

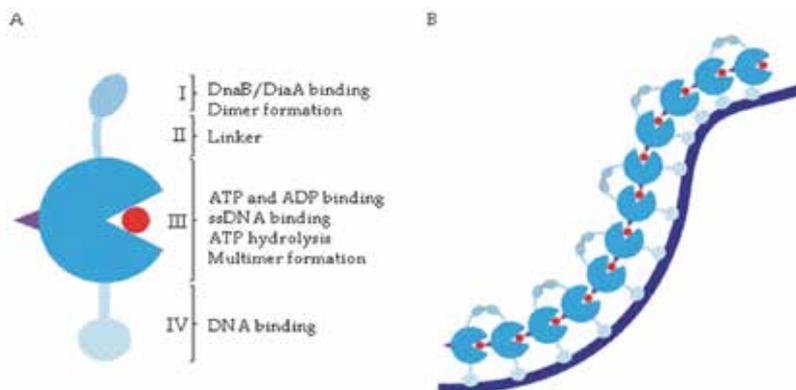


Fig. 2. DnaA as the main protein for the unfolding of DNA strands at *oriC*. A) The DnaA protein family is part of the AAA+ ATPases. In *E. coli* DnaA contains four functional domains as shown in the diagram. The ATP molecule is shown in red, and the arginine finger in purple. B) Domain III binds preferentially ATP over ADP, in addition it has an “arginine finger” which permits the multimerization of these protomers over the DNA.

119 (Figure 1), (Cassler et al., 1995; Ryan et al., 2004). Additional proteins such as DiaA and HU (Heat unstable protein) bind to domain I of DnaA, contributing to the stabilization of the joining of their protomers to *oriC* (Ishida et al., 2004; Chodavarapu et al., 2008). Another protein, ArgP (arginine protein, also called IciA) binds to the AT-rich regions in L, M and R boxes blocking the opening of DNA by DnaA (Hwang et al., 1992), this protein binds in the order of 10-20 monomers per *oriC*. Mutants in this gene however have no a clear defective phenotype of DNA replication and possibly this protein is functioning as an additional mechanism to maintain the robustness of this process. ArgP is also a transcriptional regulator which counts *dnaA* among its target genes. The activity of ArgP is regulated by arginine as its allosteric ligand and the protein is degraded by a specific protease. Another protein that inhibits the binding of DnaA to its target sequences is CNU (*oriC*-binding nucleoid-associated). CNU is a small protein composed of 71 amino acids (8.4 kDa) that binds to a sequence of 26 bp (named *cnb*), which overlaps with the binding sites for DnaA, thereby preventing its binding to *oriC* (Kim et al., 2005). When DnaA-ATP binds to *oriC* it twists the DNA and promotes the separation of DNA-strands in the AT-rich region to produce a single-stranded bubble or “open complex” (Figure 3). The next step is the recruitment of the (DnaBC)<sub>6</sub> complex to DnaA to obtain the pre-Replicative Complex preRC.

Four or five DnaA-ATP molecules interact with the (DnaBC)<sub>6</sub> complex via the N-terminal of the replicative DnaB helicase and their common binding to *oriC* (Seitz et al., 2000). DnaB<sub>6</sub> is a monohexameric helicase with a ring shape. Its function is the unwind of double-stranded DNA employing the hydrolysis of ATP, this activity is maintained as the elongation phase proceeds. DnaB<sub>6</sub> in its inactive form is found associated with the small protein DnaC (also of the AAA+ superfamily) forming a closed complex DnaB<sub>6</sub>-(DnaC-ATP)<sub>6</sub>, (Biswas & Biswas-Fiss, 2006).

The DnaB protein should be loaded onto each of the single-stranded DNA (ssDNA) molecules. For this to happen, the pre-RC needs to release the DnaC from the complex (DnaBC)<sub>6</sub>. It has been suggested that the DNA helicase translocates between parental templates of DNA and interacts via its N-terminal domain with the DnaG primase. The

formation of the DnaB-DnaG complex is known as the “primosome”. Since replication is bidirectional in most bacterial chromosomes, one primosome is loaded on each single stranded parental (Figure 3). DnaB is responsible for the unwinding of the double stranded DNA (dsDNA) in the 5′-3′ direction and the primase synthesizes a small fragment of RNA complementary to the parental DNA-strand, not shorter than 12 and up to 29 ribonucleotides (Figure 3), (Swart & Griep, 1995; Rowen & Kornberg, 1978). The interaction of the primase with DnaB and the use of these primers trigger the release of DnaC. This action defines discrete events in the transition from initiation to the elongation phase of DNA replication (Makowska-Grzyska & Kaguni, 2010).

## 2.2 Elongation of DNA

Since the holoenzyme DNA polymerase III (Pol III, see below for components) cannot initiate DNA polymerization *de novo*, the strands are extended from the RNAs synthesized by the DnaG primase (Figure 4).

Pol III is positioned at the 3′ end of the first RNA primer complementary to the leading strand of DNA and extends it continuously. In contrast on the lagging strand the new DNA-strand is synthesized discontinuously producing Okazaki fragments of about 1 kb in length. The RNA primers are removed and substituted by DNA by DNA polymerase I (Pol I). Pol I uses 5′-3′ exonuclease activity to remove these primers and fill out the gaps with its 3′-5′ DNA polymerase activity. Then DNA-ligase joins adjacent DNA fragments by catalyzing the formation of phosphodiester bonds between the 5′ phosphate of a hydrogen-bonded nucleotide and an adjacent 3′ OH of the nucleotide of the following Okazaki fragment.

The Pol III holoenzyme is composed of three subassemblies: the core polymerase, the  $\beta$ -sliding clamp and the clamp-loader complex. The core DNA polymerase is in turn, composed of three subunits  $\alpha$ ,  $\theta$  and  $\epsilon$ . The  $\alpha$ -subunit is that which really has the activity of DNA polymerase whereas the small subunit  $\epsilon$  has proofreading 3′-5′ exonuclease activity and its function is to remove nucleotides that have been misincorporated by the core-polymerase. The  $\epsilon$ -subunit is stabilized by the  $\theta$ -subunit, which as yet has not been assigned additional functions (Schaeffer et al., 2005).

The clamp-loader or DnaX complex consists of six different subunits ( $\delta'$ ,  $\delta$ ,  $\gamma$ ,  $\tau$ ,  $\psi$ ,  $\chi$ ).  $\gamma$  and  $\tau$  subunits are encoded by the same *dnaX* gene. The full sequence of *dnaX* encodes the protein  $\tau$ . However when the mRNA is being translated the ribosome sometimes undergoes a frame shift and a shorter product (only two-thirds) results. The frameshift occurs in a poly(A) tract and yields a new stop codon immediately following the frameshift signal. This truncated form of  $\tau$  corresponds to the  $\gamma$  protein. In this way, the first three domains of  $\gamma$  and  $\tau$  are identical. These different protein versions bind to the  $\delta$  and  $\delta'$  subunits forming a complex composed of  $\delta'\gamma\delta\tau_2$  subunits. The  $\chi$ - $\psi$  dimer binds either  $\gamma$  or  $\tau$  subunits via the amino-terminus of  $\psi$  constituting the clamp-loader (Gao & McHenry, 2001; Reyes-Lamothe et al., 2010).  $\tau$  proteins have two defined interactions; on one side they attach to the  $\alpha$ -subunit of the core and on the other, interact with the DnaB<sub>6</sub> helicase on the lagging strand, so that this complex forms a bridge between the replicase and helicase proteins (Lee et al., 1996).

The single strands of DNA (ssDNA) are stabilized by a protein called single-stranded DNA-binding protein (SSB). SSB binds to single DNA-strands as a tetramer through its N-terminal domain, which makes contact with the DNA. The clamp-loader recognizes ssDNA coated by SSB<sub>4</sub>, interacting with the  $\chi$  subunit of SSB<sub>4</sub>.  $\chi$  forms a heterodimeric complex with  $\psi$ , which in turn, interacts with the  $\gamma$  and  $\tau$  subunits. In this way  $\chi$  senses the presence (or absence) of ssDNA, facilitating the recognition of the terminal parts of RNA primers by  $\tau$  (Schaeffer, 2005).

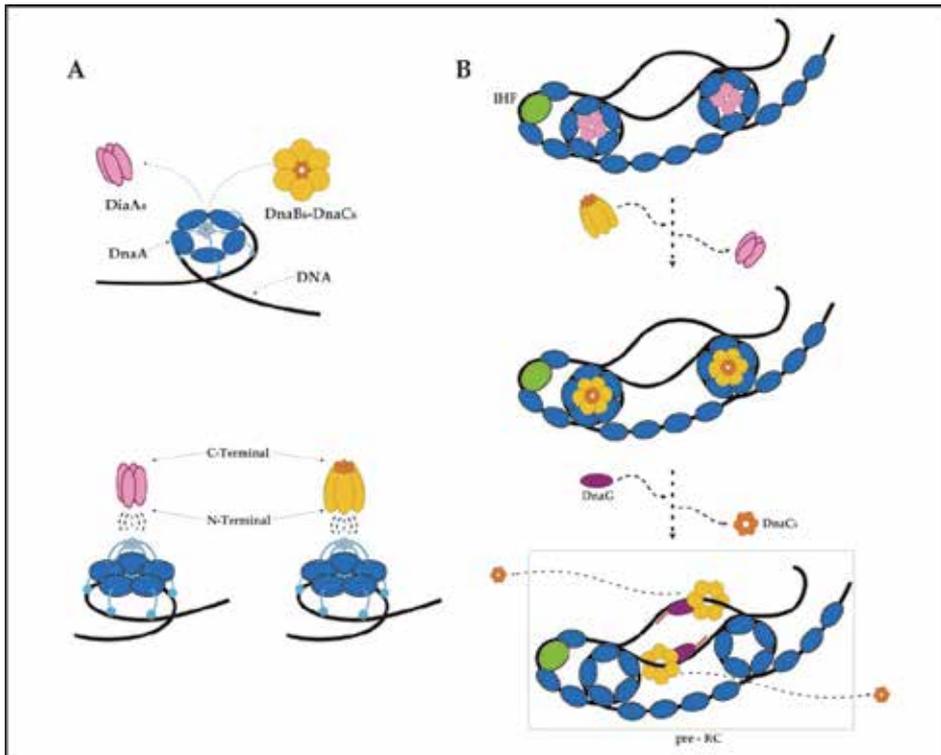


Fig. 3. Formation of the pre-RC. A) Binding of DnaA-ATP to *oriC* to form multimeric structures in conjunction with DiaA and (DnaBC)<sub>6</sub> via domain I, these interactions are important in order to form the pre-RC. B) The binding of DnaA-ATP to this region of DNA is favored when the protein IHf also binds to *oriC*, about 20 molecules of DnaA-ATP bind to *oriC* simultaneously. This DnaA-ATP complex is stabilized by DiaA and finally leads to the unfolding of the DNA at the AT-rich region. At this stage the (DnaBC)<sub>6</sub> complex is attached to domain I of the DnaA-ATP, forming the pre-RC. Subsequently, DnaB releases DnaC and loads to each single stranded DNA in direction 5'-3' with the assistance of the DnaG primase.

The sliding-clamp ( $\beta_2$ ) is a dimer of DnaN proteins, which binds to the hybrid DNA-RNA and serves to direct Pol III to this position for the synthesis of Okasaki fragments. During the elongation phase Pol III can hop from one clamp to another without leaving the replication fork. So Pol III overcomes possible delays due to blockage of DNA by the activity of transcription factors or DNA damage (Georgescu et al., 2010).

### 2.3 Termination of DNA replication

The end of DNA replication takes place when the replisome helicase DnaB<sub>6</sub> on the leading strands collides with a protein called Tus. Tus recognizes and is bound to sites for termination of DNA replication (*ter*). These sites are physically arranged in positions opposite to the *oriC* (Figure 5). In the collision of Tus with the helicase a trap is formed that prevents the further advancement of the replicative machinery in the leading strand and remains arrested until the replicative machinery on the lagging strand reaches this position (Neylon et al., 2005).

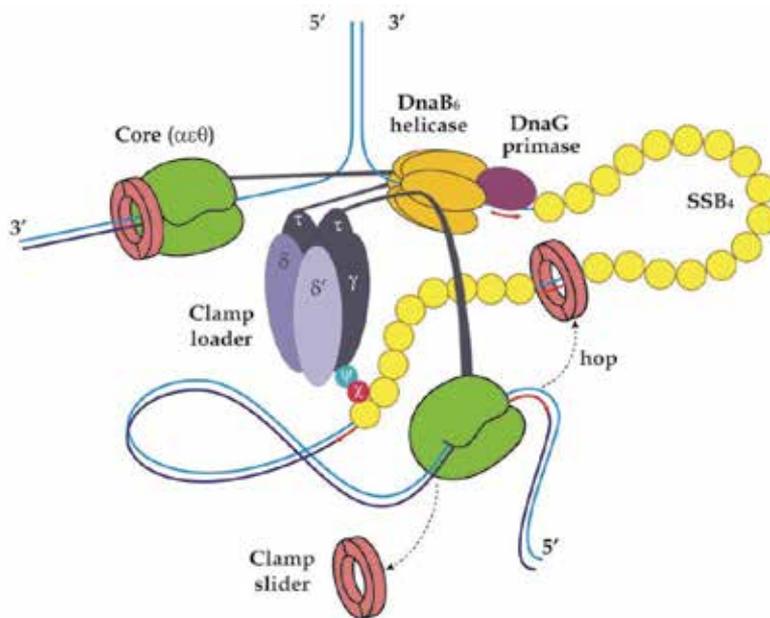


Fig. 4. Elongation of DNA by the replisome machinery.

The elongation of DNA in the *E. coli* chromosome is carried out in both directions of the fork by a multisubunit machinery called the replisome. Each replisome is located in both directions of the fork. The helicase DnaB<sub>6</sub> is loaded on the (3'-5') lagging strand to unfold the DNA duplex in the 5'-3' direction, at this time the primase synthesizes RNA primers complementary to each ssDNA. These primers are extended by Pol III, forming the Okazaki fragments on the lagging strand. When Pol III extends a new Okazaki fragment and reaches a previously synthesized one, it gives a hop, joining to another slider clamp ( $\beta$ -subunit), which recognizes DNA-RNA hybrids. DNA polymerases working on both parent strands are coordinately driven by the clamp-loader, which also binds to the helicase. SSB stabilizes the ssDNA. For the recognition of ssDNA by Pol III, the clamp-loader makes contact with SSB4-DNA via its  $\chi$ -subunit.

The resolution of chromosomes is produced by the activity of several proteins which act together to separate the two daughter chromosomes. In this process the FtsK protein is very important as it acts by coordinating cell division with chromosome segregation through the activities of its N-transmembranal domain (FtsK<sub>N</sub>) and its C-cytosolic domain (FtsK<sub>C</sub>), respectively. FtsK<sub>N</sub> is the target for the division protein that forms the septum FtsZ, which stabilizes the interactions of FtsK with the other components of the divisome FtsQ, FtsI and FtsL (Aussel et al., 2002; Dubarry et al., 2010). FtsK also contains a linker, FtsK<sub>L</sub>, localized between the FtsK<sub>N</sub> and FtsK<sub>C</sub> domains (Bigot et al., 2004). Recently two distinct regions within FtsK<sub>L</sub> have been identified (FtsK179-331 and FtsK332-641), which together with FtsK<sub>N</sub>, are required for normal septation in *E. coli* (Dubarry et al., 2010). FtsK<sub>C</sub> can lead to the dimerization of circular chromosomes, thereby compromising their segregation (Figure 5). FtsK<sub>C</sub> activates events of recombination at the *dif* site (localized beside the replication termination region), which are mediated by two proteins with activities of tyrosine recombinases, XerC and XerD to resolve chromosome dimers to monomers and at the same time promote DNA translocation (Bigot et al., 2004; Kennedy et al., 2008). FtsK<sub>C</sub> is part of the

AAA<sup>+</sup> superfamily and therefore can form a ring-shaped multimer that wraps the DNA and moves along it at the expense of ATP. When a chromosome dimer is present, a site-specific recombination event by XerCD introduces an additional cross over at *dif*, resolving thus the dimer into two monomers, all this is under the control of FtsK (Aussel et al., 2002).

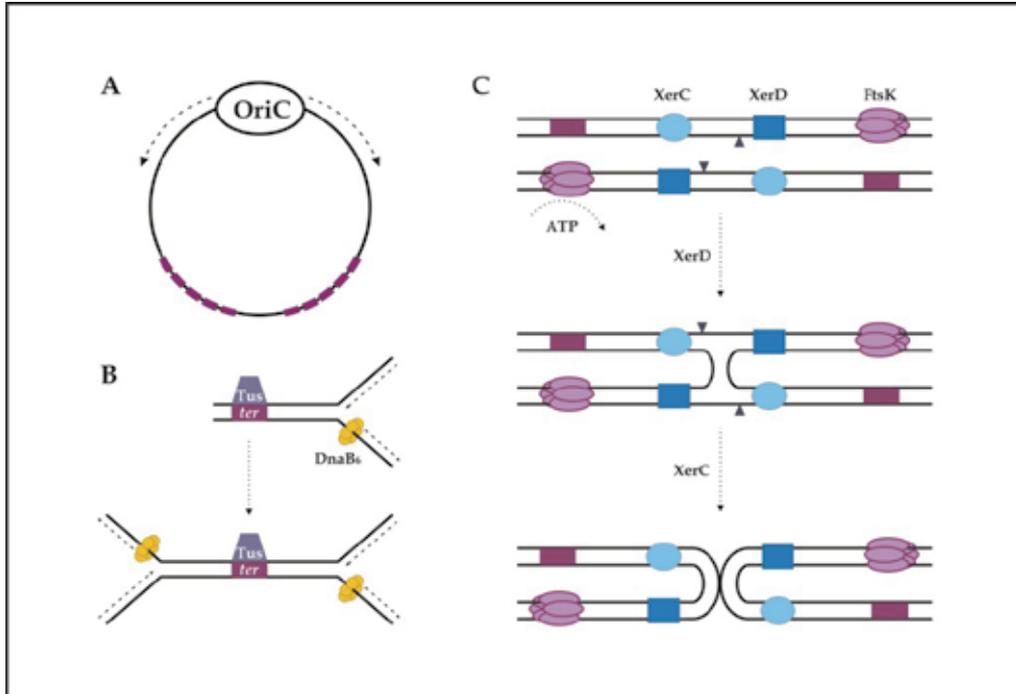


Fig. 5. Termination of DNA replication. A) The site of termination of replication in *E. coli* is opposite to *oriC*, where there are specific *ter* sequences which are recognized by the Tus protein (purple boxes). B) Tus protein-terminator sequence (Tus-*ter*) is a barrier that pauses the leading fork until the lagging fork arrives from the opposite direction and induces termination, which occurs when the helicase touches Tus. The helicase dissociates from DNA and Pol III synthesizes the complementary strand on both sides of the forks. C) Near to the Tus-*ter* sites is found a sequence named *dif*, where site-specific recombination mediated by the XerC and XerD recombinases assisted by the translocase FtsK takes place. Figure taken and modified from Aussel et al. (2010).

A summary of the key enzymes involved in DNA replication known to date in *Escherichia coli*, are shown in table I.

### 3. Regulation of DNA replication

The regulation of DNA replication is a vital cellular process. In a general view, DNA replication is controlled by a series of mechanisms that are centered on the control of cellular DnaA levels, its availability as a free protein and modulation of its activity by binding the small-molecule ligand ATP (Leonard & Grimwade, 2009); the other point of control is by modulating the accessibility of replisome components to the *oriC* region on the DNA. We discuss some aspects of these regulatory mechanisms below.

Protein name	Gene name	Function	Gene length (bp)	MW <sup>a</sup> (kDa)	Essentiality <sup>b</sup>
DnaA	<i>dnaA</i>	Initiator of DNA synthesis by binding to the origin of replication and also acts as a transcriptional regulator. It binds to DnaA boxes, and binds ATP. Around 20 to 30 DnaA monomers bind to the <i>oriC</i> region. It is calculated that around of 1000 molecules per cell are bound reaching up to 70% DnaA-ATP.	1404	52.551	E
DnaB	<i>dnaB</i>	A hexameric DNA helicase, it progressively unwinds DNA strands ahead of replication forks. About 100 DnaB molecules are calculated to be present per cell.	1416	52.39	E
DnaC	<i>dnaC</i>	DnaC is an accessory protein that assists the loading of DnaB onto DNA duplex to initiate replication and onto ssDNA to assist primer formation by the primase. Six DnaC monomers bind to the hexameric DnaB	738	27.935	E
DnaG	<i>dnaG</i>	DNA primase, it catalyzes the synthesis of RNA primers on ssDNA. These primers are necessary for DNA synthesis by DNA polymerase III. A DnaB-DnaG complex was observed by mixing DnaB with a six molar excess of DnaG (hexamers of DnaB and monomers of DnaG). Log-phase cells contain 50 to 100 molecules of primase.	1746	65.565	E
DNA polymerase III holoenzyme (Pol III)		DNA polymerase III holoenzyme is the primary enzyme for DNA synthesis in <i>E. coli</i> . It carries out 5' to 3' DNA polymerization using ssDNA as a template; it also carries out 3'-5' exonuclease edition of mispaired nucleotides. There are estimated to be 10 holoenzymes of DNA polymerase III per cell. Pol III holoenzyme is made up of the following components [(DnaE)(DnaQ)(HolE)] <sub>3</sub> [(DnaX) <sub>3</sub> (HolB)(HolA)][(DnaN) <sub>2</sub> ][(DnaX) <sub>2</sub> ][(HolC)(HolD)] <sub>4</sub> .			
DNA polymerase III (core)		The DNA polymerase III core enzyme can carry out the basic polymerase and exonuclease activities of polymerase III.			
α	<i>dnaE</i>	α subunit catalyzes DNA polymerization from 5' - 3'.	3483	129.9	E
ε	<i>dnaQ</i>	ε subunit catalyzes the 3' - 5' proofreading activity	732	27.099	E
θ	<i>holE</i>	θ subunit allows stabilization of α and ε subunits	231	8.846	NE
β	<i>dnaN</i>	The β subunit dimerizes to form the sliding clamp which positions the core polymerase onto the DNA.	1101	40.587	E
Clamp loader		It catalyzes ATP-driven assembly of the sliding clamp onto primer-template DNA. Clamp loader = δδ'τ <sub>2</sub> γψχ			
δ	<i>holA</i>	δ subunit acts as a wrench to open the sliding clamp probably using ATP. Some δ units exist independently of the preinitiation complex, possibly playing a role in stripping β clamps from DNA in the absence of replication initiation.	1032	38.704	E

$\delta'$	<i>holB</i>	$\delta'$ subunit is part of the clamp loader complex.	1005	36.937	E
$\tau$	<i>dnaX</i>	$\tau$ subunit binds to the alpha subunit dimerizing the core alpha-epsilon-theta polymerase subunits. This is required for synthesis on the lagging strand.	1932	71.138	E
$\gamma$	<i>dnaX</i>	$\gamma$ subunit is part of the clamp loader complex.	1932	47.545	E
$\chi$	<i>holC</i>	$\chi$ subunit allows the binding of the clamp loader to SSB. $\psi$ - $\chi$ also acts in multiple ways improving the binding of DNA polymerase to DNA templates.	444	16.633	E
$\psi$	<i>holD</i>	$\psi$ subunit allows the interactions between $\gamma$ and X subunits	414	15.174	NE
Fis	<i>fis</i>	Fis for "factor for inversion stimulation" allows the organization and maintenance of the nucleoid structure through direct DNA bending and by modulating the production of gyrase and topoisomerase I as well as regulating the expression of other proteins that modulate the nucleoid structure, such as HNS, and HU. It reaches a cell concentration of 40,000-60,000 molecules/cell at the beginning of the exponential phase	297	11.24	E
Dam	<i>dam</i>	The DNA adenine methyltransferase is responsible for methylation of GATC sequences in <i>E. coli</i> . A wild-type, rapidly growing <i>E. coli</i> cell (doubling time = 30 min) was found to contain about 130 molecules of Dam methyltransferase.	837	32.1	NE
DiaA	<i>diaA</i>	DiaA interacts with DnaA, it is required for the timely initiation of chromosomal replication and stimulates the replication of minichromosomes <i>in vitro</i> .	591	21.106	NE
ArgP/IciA	<i>argP</i>	The ArgP transcriptional activator or inhibitor of chromosome initiation (IciA) regulates DNA replication by binding to three 13-mers located in the origin of replication (OriC), blocking the DNA opening by DnaA. It is also a transcriptional repressor of <i>dnaA</i> . There are about 800 molecules/cell of IciA in the exponential phase and the level decreases to about 500 molecules per cell in the early stationary phase.	894	33.472	NE
IHF		"Integration host factor", is a global regulatory protein that helps to maintain the DNA architecture. It binds and bends DNA. IHF plays a role in DNA supercoiling and DNA duplex destabilization and affects processes such as DNA replication, recombination, and the expression of many genes. Consisting of two subunits $\alpha$ and $\beta$ . IHF reaches 6,000-15,000 complexes in the exponential phase and up to 30,000-55,000 in the stationary phase.			
IHF- $\alpha$	<i>ihfA</i>	$\alpha$ subunit of IHF	300	11.354	NE
IHF- $\beta$	<i>ihfB</i>	$\beta$ subunit of IHF	285	10.651	NE

HU		HU for heat unstable protein, is a global regulatory protein and shares properties with histones for nucleoid organization and regulation. It is a heterodimer formed by an $\alpha$ - and a $\beta$ -subunit. HU reaches 30,000-55,000 dimers in the exponential phase and 10,000-17,000 in the stationary phase.			
HU- $\alpha$	<i>hupA</i>	$\alpha$ -subunit of HU	273	9.535	NE
HU- $\beta$	<i>hupB</i>	$\beta$ -subunit of HU	273	9.226	NE
DNA Pol I	<i>polA</i>	In addition to polymerase activity, this DNA polymerase exhibits 3'→5' and 5'→3' exonuclease activities. It is able to utilize nicked circular duplex DNA as a template and can unwind the parental DNA strand from its template. Its cellular abundance is of around 400 molecules per cell.	2787	103.12	NE
LigA	<i>ligA</i>	LigA is one of two known NAD(+)-dependent DNA ligases, it catalyzes the formation of phosphodiester bonds on duplex DNA.	2016	73.606	E
SSB	<i>ssb</i>	Single-stranded DNA-binding protein acts as a tetramer when binding to DNA. Each <i>E. coli</i> cell has about 800 monomers of SSB.	537	18.975	E
Tus	<i>tus</i>	Tus, also known as <i>ter</i> -binding protein (TBP), binds to <i>ter</i> sites, blocking the progress of DNA replication in a polar like form.	930	35.783	E
FtsK	<i>ftsK</i>	FtsK is an essential cell division protein linking cell division with chromosome segregation	3990	146.66	E
Hda	<i>hda</i>	Regulator of DnaA that prevents premature initiation of DNA replication. Around 100 molecules/cell are found.	702	28.37	E
RapA	<i>hepA</i>	A RNA Polymerase-binding ATPase and RNAP recycling factor.	2907	109.77	NE
SeqA	<i>seqA</i>	Sequesters newly replicated hemimethylated <i>oriC</i> to prevent re-initiation; it also binds hemimethylated GATC sequences.	546	20.315	E
Xer site-specific recombination system		Two lambda integrases of the family of recombinases involved in converting chromosome dimers of into monomers so that segregation of the chromosomes can occur during cell division			
XerC	<i>xerC</i>	XerC is part of the Xer site-specific recombination system	897	33.868	E
XerD	<i>xerD</i>	XerD is part of the Xer site-specific recombination system	897	34.246	NE

<sup>a</sup>MW: Molecular weight of the polypeptide product.

<sup>b</sup>Essential gene (E)/ non essential gene (NE).

Table 1. Description of major proteins for replication in *E. coli*

### 3.1 Regulatory mechanisms of DNA replication in *E. coli*

One of the main mechanisms associated with DNA replication is the so-called RIDA system (Regulatory Inactivation of DnaA). The elements of this system are the sliding-clamp of DNA polymerase III and Hda (Homologous to DnaA). This mechanism takes place when DnaA is activated by its binding to ATP. The accumulation of DnaA in this active form leads to the initiation of chromosomal replication since it facilitates its binding to the *oriC* on the DNA. DnaA reverts to its inactive form DnaA-ADP by hydrolysis of ATP (Katayama et al., 1998). Hda-ADP is the monomeric active form for promoting the hydrolysis of DNA-ATP, a process which is mediated by the slider-loader clamp (Su'etsugu et al., 2008). This inactivating regulation of DnaA is key for preventing the over-initiation of replicative events during the cell cycle (Katayama & Sekimizu, 1999). The free-living bacteria *C. crescentus* also presents this regulatory mechanism, as it has HdaA, a protein similar to the *E. coli* Hda. In *C. crescentus* HdaA also inactivates DnaA in a replication-coordinated manner, if DNA replication is successfully initiated then HdaA and the  $\beta$ -sliding clamp promote the hydrolysis of DnaA-ATP to DnaA-ADP and force DnaA to leave the *oriC* (Collier & Shapiro, 2009). A conserved bacterial protein, YabA, has been found in *B. subtilis* and other Gram-positive bacteria where it acts as a repressor for initiation of DNA replication. This is achieved by forming a complex with DnaA and the  $\beta$ -sliding clamp independently of the DNA, a common activity shared between Hda and YabA (Mott & Berger, 2007). Thus the RIDA system is present in *B. subtilis* and is also the primary mechanism for regulation of DNA replication in this bacterium (Noirot-Gros et al., 2006). The formation of the *oriC* and DnaA complex is assisted by the protein DiaA, which forms homo-tetramers and binds various DnaA molecules, especially in the active form of DnaA-ATP but it can also stimulate the formation of the DnaA-ADP-*oriC* complex, this is an inactive complex for initiation of replication (Ishida et al., 2004).

Another mechanism that regulates the initiation of DNA replication is by controlling the availability of free DnaA to bind to DnaA boxes on the *oriC* (Figure 1). Here the role of the 1kb *datA* locus, which is localized near (downstream) from the *oriC* is important. The *datA* locus shows high affinity for DnaA, even more than the DnaA boxes on the *oriC*. Thus the *datA* region is able to bind over 300 DnaA molecules whereas *oriC* binds to 45 DnaA monomers (Kitagawa et al., 1998). The operability of this mechanism is facilitated by the fact that the *oriC* had only few DnaA boxes compared to the *datA* locus and by the close proximity of *datA* in respect to *oriC* on the DNA molecule (Figure 6).

One related control system depends on the property of DnaA to act as a transcription factor and to the presence of DnaA boxes in the promoter regions of several genes. In most cases DnaA represses the expression of the associated gene but in some cases it can activate certain genes (Messer & Weigel, 1997). DnaA regulates around 10 genes in *E. coli* as documented in RegulonDB (Gama-Castro et al., 2010). The transcription of *dnaA* is one of the most important regulatory mechanisms that directly affect the replication of DNA and one of the proteins that negatively regulate the expression of *dnaA* is DnaA itself (Figure 6). At high levels DnaA binds to the DnaA boxes in the promoter region and impedes transcription. This auto-repressive process directly affects the amount of DnaA-ATP available and controls the efficiency of initiation of DNA replication (Mott & Berger, 2007). In *C. crescentus*, it was found that DnaA also auto-represses the transcription of its own gene but additionally DnaA is highly unstable in this organism and gradually degrades after initiating a replication event (Gorbatyuk & Marczyński, 2005).

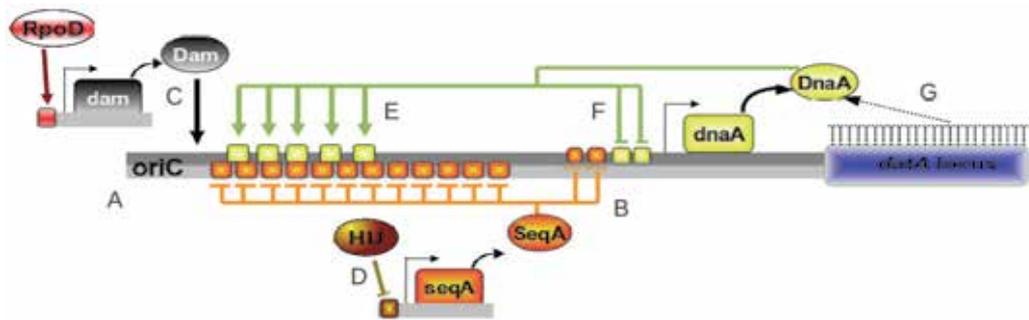


Fig. 6. Mechanisms that regulate DNA replication in *E. coli*. A) The newly replicated DNA duplex is in a hemimethylated state. B) SeqA binds to the hemimethylated GATC sites immediately after they are replicated. C) RpoD activates the transcription of *dam* and Dam methylates GATC sites of the newly synthesized strand. D) HU represses the transcription of SeqA. E) DnaA binds to the DnaA boxes on the *oriC* region. F) when there are many DnaA molecules they repress the transcription of the *dnaA* gene. G) *datA* locus binds many DnaA molecules.

### 3.2 Regulation of DNA replication by DNA methylation

A requirement for initiation of DNA replication is that both DNA strands are methylated, principally the adenine nucleotide in the GATC motifs, this process is mediated by Dam (DNA adenine methyltransferase), (Wion & Casadésus, 2006). Dam binds to the DNA nonspecifically, and methylates the GATC motifs (Figure 6). On DNA strands recently synthesized these motifs are rapidly methylated and exist in the hemimethylated state only during a fraction of the time needed for the replication of the entire DNA (Casadésus & Low, 2006).

The methylation process occurs asynchronously on the newly synthesized strands; i. e. methylation on the lagging arm occurs only after the ligation of the Okazaki fragments. It is postulated that Dam is always present in a complex bound near the replication origin, thus the methylation of nascent DNA strands occurs as soon as polymerization begins. In summary, the presence of hemimethylated GATC sites provides a cue to indicate that DNA replication has just occurred (Stancheva et al., 1999).

Another way to repress the transcription of *dnaA* is that which occurs immediately after the initiation of DNA replication. Here, SeqA binds to the hemimethylated GATC sequences in the regulatory regions of the *dnaA* gene (Lu et al., 1994; Brendler et al., 2000). Similarly, SeqA also represses the replication of DNA by binding to the hemimethylated GATC sequence at the *oriC*, this is possible because SeqA DNA-binding sites overlap with those of low affinity for DnaA (DnaA boxes) on the *oriC*. This overlap impedes the complete access of DnaA-ATP to the *oriC* (Han et al., 2004).

This prevention of replication, dependant of DNA methylation, has been considered as an epigenetic regulatory mechanism because it depends on the chemical modification of the nucleotide residues of the DNA and not in its sequence.

### 3.3 Regulation of DNA replication in *Bacillus subtilis*

*B. subtilis* shares some orthologous genes to the regulators that are involved in DNA replication in *E. coli*, but particular regulatory mechanisms must occur in this organism, as it

lacks some important components of the regulatory machinery found in *E. coli* such as the *seqA* and *dam* genes. In their place other players are present in *B. subtilis* such as Spo0A (Figure 7) and SirA (sporulation inhibitor of replication) (Katayama et al., 2010). Spo0A is the master regulator for sporulation and, at the same time, is an inhibitor of DNA replication. Spo0A is activated by a multicomponent phosphorelay process, this is initiated by a histidine kinase (KinA), that autophosphorylates, and transfers the phosphate to Spo0A through two intermediate phosphotransferases (Spo0F and Spo0B), (Burbulys et al., 1991). Spo0A-P (the active form) binds to specific sites on the *oriC* region and blocks the unwinding of the DNA duplex. Spo0A-P activates SirA, and SirA binds to DnaA in Domain I inhibiting the ability of DnaA to bind to the *oriC* (Wagner et al, 2009). Sda maintains the cellular levels of Spo0A-P low when a new round of replication has initiated (Veening et al., 2009), by inhibiting the accumulation of the autophosphorylated form of KinA (Cunningham & Burkholder, 2008).

Other regulators also implicated in DNA replication in *B. subtilis* are Soj and Spo0J (Figure 7), both components are required for proper chromosome segregation and for the repression of DNA replication. Soj exerts its activity in repressing replication by interacting with DnaA at the *oriC*, thereby preventing DnaA from initiating DNA replication (Murray & Errington, 2008). Otherwise Spo0J produces the complex Soj-Spo0J at the *parS* locus (Autret et al., 2001), promoting the release of Soj from the DNA strands, and allowing DNA replication to be initiated (Lee et al., 2003).

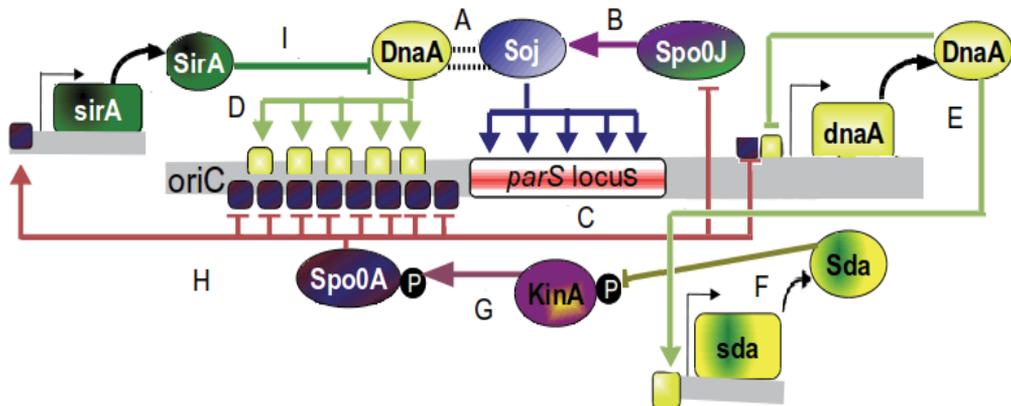


Fig. 7. Mechanisms that regulate DNA replication in *B. subtilis*. A) Soj represses DnaA activity. B) Spo0J stimulates Soj binding to the *parS* locus. C) The complex of Soj at the *parS* locus promotes the separation of Soj from the DNA. D) DnaA binds the DnaA boxes in the *oriC* initiating DNA replication. E) DnaA represses *dnaA* itself and activates the transcription of *sda*. F) Sda inhibits the accumulation of KinA-P. G) KinA activates Spo0A by transferring a phosphate group to Spo0A. H) Spo0A binds to specific sites in the *oriC* and represses replication, it also represses *dnaA* and *spo0J* and activates *sirA*. I) SirA, in turn, binds to DnaA and represses its binding to the *oriC*.

### 3.4 Regulation of DNA replication in *Caulobacter crescentus*

An interesting mechanism for control of DNA replication takes place in the cell cycle of *C. crescentus*, this aquatic, free-living bacteria, divides asymmetrically and this process is

regulated by a complex circuit of master regulatory proteins (Figure 8) coupled to a two-component system.

One of these regulators is the master regulator of cell cycle CtrA (Cell cycle transcriptional regulator), which is transcriptionally regulated by methylation of the GANTC motif on the first of the two of *ctrA* promoters (P1). Transcription initiation at P1 is repressed when the GANTC motifs are fully methylated while in the hemimethylated state transcription takes place. This mechanism ensures that *ctrA* is transcribed only while replication is in progress, producing enough protein to block and prevent the reinitiating of another round of DNA replication during this time (Reisenauer & Shapiro, 2002). In the hemimethylated form the production and accumulation of CtrA occurs, this protein binds to the regulatory region of *ccrM* and activates the transcription of a DNA-methylase encoded by this gene. Once synthesized, this enzyme proceeds to complete the methylation of both DNA strands. CtrA ceases its repressing activity when it is degraded by a Lon-type protease. The transcription of *ccrM* mediated by CtrA is inhibited when the two GANTC regulatory motifs are methylated. This complex machinery determines that when DNA is fully methylated, the transcription of *ctrA* and *ccrM* genes turns off (Stephens et al., 1995). This regulatory mechanism ensures that the synthesis of CcrM remains *off* and takes place only when the replication fork reaches the position of the *ccrM* gene preventing its premature transcription (Reisenauer et al., 1999).

The phosphorylated state of CtrA (CtrA-P) is the active form of this regulatory protein and this process is mediated by a cascade of phosphorylations which start with the activation of DivK, mediated by CtrA. DivK transfers the phosphate group to the CckA intermediate (Cell cycle histidine kinase) and CckA and ChpT finally transfer the phosphate group to CtrA. In the swarmer cell type of *C. crescentus*, CtrA-P binds to five DNA motifs on the *oriC* region, repressing the process of DNA replication (Marczynski & Shapiro, 2002). For the replication process to take place CtrA-P must be degraded by the ClpXP protease, which releases the origin of replication. ClpXP and CtrA are localized to each of the poles in stalked cells. This polar targeting of ClpXP is mediated by CpdR (a two component receiver protein), which is a dephosphorylating protein positioned at the pole where it recruits ClpXP (Jenal, 2009). Sometime after this happens, the proteolysis of CtrA ends and a positive transcriptional feedback loop generates the accumulation of CtrA, blocking again the access of DnaA to the *oriC* (Hung & Shapiro, 2002).

Another regulatory system for DNA replication in *C. crescentus*, is the regulatory circuit of DnaA, CtrA, GcrA and SciP. This genetic circuit regulates the transcription of multiple genes (DnaA alone controls the expression of approximately 40 genes in this bacterium) and many of these genes encode components of the replisome, in particular activating *gcrA*. On the other hand, CtrA regulates about 95 genes principally those involved in flagella biogenesis, cellular division and other regulators, and inhibits *gcrA*. GcrA in turn controls over 50 genes including the activation of *ctrA* and the repression of *dnaA* (Laub et al., 2007). Finally, SciP represses *ctrA*, and it is regulated in a feed forward loop manner; activated by CtrA and repressed by DnaA (Tan et al., 2010).

Some of the regulatory mechanisms concerning DNA replication are conserved in bacteria (as shown throughout this chapter) but specific mechanisms are also characteristic of each organism, table 2 shows the comparison of the regulators present in the three bacterial models described above.

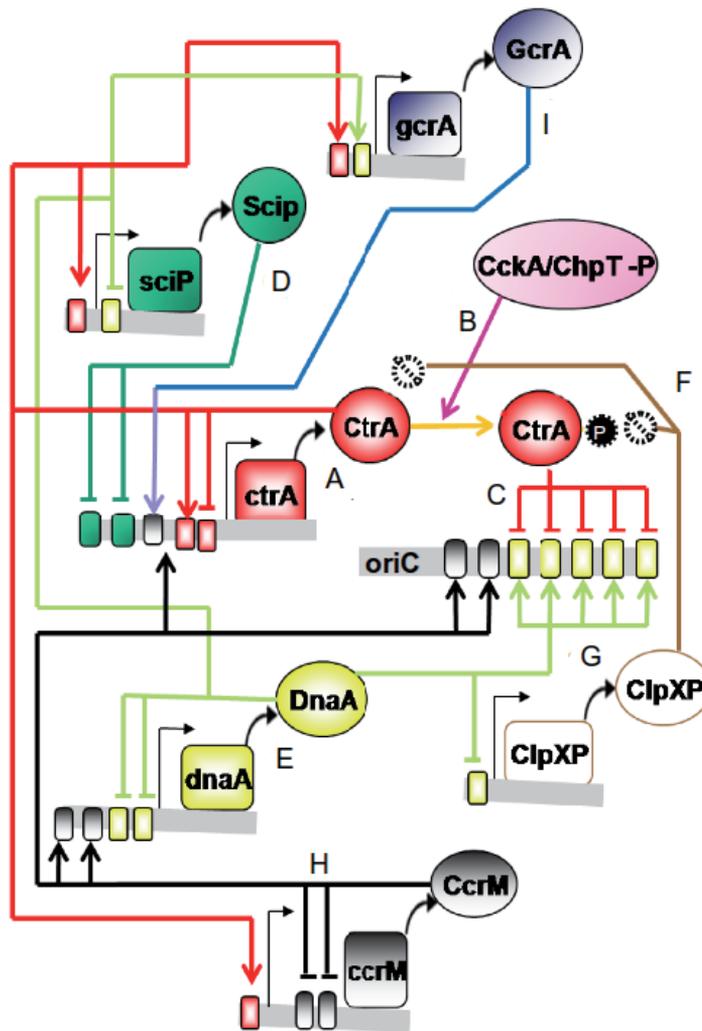


Fig. 8. Regulatory circuits that control the process of DNA-replication in *C. crescentus*. A) CtrA activates and represses the transcription of its own gene, additionally it activates *gcrA*, *ccrM* and *sciP*. B) CckA and ChpT transfer the phosphate group to CtrA. C) CtrA-P binds to *oriC* and inhibits the initiation of DNA replication. D) SciP represses the transcription of *ctrA*. E) DnaA auto-represses its own transcription in addition to the *clpXP* and *sciP* genes, it also activates *gcrA*. F) ClpXP degrades both CtrA and CtrA-P forms. G) DnaA binds to *oriC* to promote the initiation of DNA replication. H) CcrM methylates the GANTC sites on the regulatory regions of *dnaA*, *ctrA* and on its own gene, and also on the *oriC* region. I) GcrA activates the transcription of *ctrA*. Figure taken and modified from Tan et al. (2010).

#### 4. The stringent response arrests DNA replication in bacteria

When bacteria are under metabolic stress, mainly in starvation conditions, they activate a regulatory mechanism called the stringent response. This response usually corresponds to

Regulatory systems of DNA replication	Regulatory genes present in the organisms		
	<i>E. coli</i>	<i>B. subtilis</i>	<i>C. crescentus</i>
RIDA	Hda	YbaA *	HdaA*
<i>dnaA</i> gene regulation	<i>DnaA autoregulation, promoter methylation</i>	DnaA, Soj, SirA	DnaA, CcrM
DnaA regulation	DnaA-ATP/ADP, <i>datA</i> sequestration	DnaA-ATP/ADP	DnaA-ATP/ADP
<i>oriC</i> blocking	SeqA	Spo0A	CtrA
DNA methylation	Dam	-	Ccrm
Phosphorylation cascade	-	KinA, Spo0F, Spo0B	DivK, ChpT, CckA
Proteolysis	-	-	ClpXP

\* Orthologous to the *E. coli* components. - unidentified.

Table 2. Comparison of the controls that regulate DNA replication in *E. coli*, *B. subtilis* and *C. crescentus*.

the deprivation of amino acids, carbon, and limitations of nitrogen and phosphate. Under these conditions the cells suffer a reduction in size and restrict the content of their genetic information to only one nucleoid per cell (Schreiber et al., 1995).

The signal which triggers the stringent response is mediated by the accumulation of small-molecule nucleotides. These are guanosine tetra- and penta-phosphates; ppGpp and pppGpp (Ferullo & Lovett, 2008). These alarmones are synthesized as a response to the nutritional limitations by the proteins ReIA (synthetase I) and SpoT (synthetase II), (Bernardo et al., 2006). During the stringent conditions the elongation phase of DNA replication is inhibited because ppGpp and pppGpp specifically block the activity of the primase enzyme (DnaG). This is caused by the binding of a phosphate group of ppGpp to the primase resulting in an allosteric inhibition of the replication complex, the primase cannot therefore bind to the helicase. High cellular levels (up to millimolar concentrations) of these small nucleotides completely arrest DNA replication whereas lower levels only diminish the rate of replication (Wang et al., 2007).

Another path of regulation of DNA replication under a stringent condition is produced by the fact that the promoter of *dnaA* is also subject to the stringent response and the transcription of *dnaA* is also repressed under these conditions (Chiaramello & Zyskind, 1990; Levine et al., 1995).

## 5. DNA replication and asymmetrical bacterial cell-division

In *B. subtilis* the arrest of DNA replication takes place around the *oriC*, from the *gnt* gene on the left arm over an equal distance to the *gerD* gene of right arm, covering at least 190 kpb on both sides of the *oriC* (Levine et al., 1991). During this process the stages of chromosomal segregation in cell division differ between prespores and vegetative cells. First, the newly replicated chromosomes are attached at each of the cell poles (one pole will become the spore and the other pole the mother cell). Upon the asymmetric septation, under stress, about 30% of one of the replicated chromosomes is trapped in the prespore (Wu & Errington, 1994). The protein SpoIIIE forms a pore in the invaginating septum around the trapped DNA and permits the transfer of the remaining chromosome through the septum into the prespore (Lewis, 2001). All this produces an imbalance among regulators in the forespore and vegetative cell that results in an asymmetrical cell division in *B. subtilis*.

Another example of asymmetrical cell-division happens in *C. crescentus*, this bacterium differentiates into two different progeny: a flagellated swarmer cell and a stalked cell. The swarmer cells are incapable of replicating their DNA (prevented by the mechanisms previously mentioned in this chapter), until they differentiate into a stalked cell, this cell-type immediately enters into a new period of chromosome replication and cell division, and generates again the two cell types (Ryan & Shapiro, 2003). When *C. crescentus* is starved of carbon sources, its DnaA protein is degraded in a manner that depends on the stringent response mediated by the protein Spo, a ppGpp synthetase (Lesley & Shapiro, 2008). Additionally starvation increases the degradation of DnaA leading to the stabilization of CtrA resulting in the inhibition of DNA replication (Gorbatyuk & Marczyński, 2005).

## 6. DNA replication in bacteria with two chromosomes

Until now, in this chapter we have discussed replication focusing on bacteria with one chromosome, but some bacteria have more than one chromosome, one example of this is *Vibrio cholerae*, a human pathogen, which possesses two chromosomes, chrI and chrII (Heidelberg et al., 2000). The components and regulation of DNA replication for chrI in *V. cholera* are similar to the *oriC* of *E. coli* whereas the *oriC* of chrII shares some characteristics with plasmid replicons. Both cases (chrI and chrII) also require a specific repeated sequence for the replicative machinery (Zakrzewska-Czerwińska, et al., 2007). One of the specific requirements is that chrI initiates replication assisted only by DnaA whereas chrII requires the activity of the RctB protein that binds specifically to its *oriC* (Duigou et al., 2006), and an untranslated trans-acting RNA (*rctA*) (Egan et al., 2005). However the two chromosomes replicate synchronously although each has requirements for specific components which reduces the competition between both origins of replication for the replicative machineries (Duigou et al., 2006).

The proper regulation of DNA replication in bacteria with multiple chromosomes must involve interesting strategies to control the replication of both chromosomes. Unfortunately our knowledge about the regulation of DNA replication in these cases is poorly understood. It has been suggested that organisms with two chromosomes have an advantage for regulation of replication in some environmental conditions such as in free-living aquatic conditions or in association with a host, since faster replication of all DNA content is facilitated (Egan & Waldor, 2003).

## 7. Bacteria with multiple nucleoids

Another interesting phenomenon associated with DNA replication is endoreduplication (duplication of DNA in the absence of cell-division) as happens in the differentiation of *Rhizobium etli*, when these bacteria form a nodule and enter on it, in an endosymbiotic association with roots of leguminous plants. Irreversible cell differentiation occurs in these bacteria, which generates a nitrogen fixing bacteroid that is metabolically and morphologically different from the original pre-nodule cell. The differences between these types of cells result from cellular elongation and endoreduplication, without cell division. These bacteroid cells result from normal cells suffering repeated rounds of DNA replication and since the cell division is blocked they have multiple nucleoids (Mergaert et al., 2006). Interestingly this endoreplicative process is controlled by factors that are nodule-specific cysteine-rich (NCR) peptides generated from the host plant and targeted to the bacterial periplasm, with the ability to penetrate the bacteria membrane and function in its cytoplasm (Van de Velde et al., 2009).

## 8. Future perspectives

There are many details pending even in the best studied bacterial models. Some of the advantages of knowing in detail the replication process and its regulation are the possibilities for controlling the replication rates in bacteria, for example, to block the DNA replication of a pathogen or achieve cell-synchronization in bacterial cultures. Using this last premise, Ferullo et al. (2009), developed a method for synchronizing *E. coli* cultures, by treating the bacteria with DL-serine hydroxamate, a structural analogue of the amino acid serine, this treatment induces a natural stringent response, causing the arrest of the initiation of DNA replication, once the stringent signal is released, cells initiate a synchronized round of DNA replication.

Another advantage of knowing the details of the replication process and its regulation is to allow us to control and use it as a clock in some bio-engineered systems, an example of this is the ON and OF switch, generated by the methylated or hemimethylated state of DNA in *E. coli* (Low & Casadesús, 2008), specially at the GATC sites of the regulatory regions of many genes and the possibility of timing the replication rate in this organism.

Some organisms with reduced genomes such as the obligate endosymbionts *Baumannia cicadellinicola* and *Carsonella ruddii*, have lost most of the relevant components of the replicative machinery, such as DnaA. It is suggested that, the lack of DnaA allows the host to control DNA replication of the symbiont avoiding over-reproduction of the bacteria in its cytosol (Akman et al., 2002). Another possibility is that DNA replication happens at a low basal- rate in these stable conditions, in an unrepressed manner. It is postulated that the association between different organisms leads to adaptation in the rate of DNA replication of the bacteria in balance with the developmental status of their hosts (Gil et al., 2003).

## 9. Conclusions

The replication of DNA is a complex process in which a great number of regulators and mechanisms are involved, one of the most important is the DnaA protein. Replication normally begins by the formation of a complex of DnaA at the *oriC* region, with the assistance of DiaA, and the incorporation of some proteins that form the replisome, subsequently the formation of the open complex takes place, followed by a complex

interaction of the proteins needed to execute and complete the DNA replication. The process finalizes with the recognition of the *ter* site and disassembly of the replisome. Many of the proteins are broadly conserved within the bacteria but some special factors are required in bacteria which undergo particular processes such as asymmetrical cell division.

In general these processes are controlled by a series of circuits, which usually center on the *oriC* and affect the activity of DnaA. The result is regulation of the initiation step of DNA replication. Some of the regulatory mechanisms are time-dependent allowing only one DNA replication event per cell cycle. The methylation state of the DNA-strands is another important condition that not only controls the possibility of starting DNA replication but also regulates the transcription of many genes important for the execution of this function. All or certain of these mechanisms are adjusted under some special conditions, such as when the stringent response is triggered by amino acid starvation. In some bacteria with extremely reduced genomes it is still a mystery as to how DNA replication takes place and how it is controlled. Many of these latter organisms lack several important proteins implicated in the control and execution of DNA replication, and these bacteria can be useful as models for generating a system with the minimal components necessary for DNA replication.

## 10. Acknowledgments

Authors thank June Simpson for critical comments to the ms and David Velázquez Ramírez for assistance with Figures 1-5. This work was supported by CONACYT grant (102854) for young researchers given to AM-A. CQ-V thanks CINESTAV-IPN and UAA for a PhD scholarship and LE-S thanks CONACYT for a PhD scholarship (208153).

## 11. References

- Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M. & Aksoy, S. (2002). Genome sequence of the endocellular obligate symbiont of tsetse flies *Wigglesworthia glossinidia*. *Nat Genet*, 32, 402–407.
- Aussel, L., Barre, F. X., Aroyo, M., Stasiak, A., Stasiak, A. Z. & Sherratt, D. (2002). FtsK Is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell*, 108, 195–205.
- Autret, S., Nair, R., & Errington, J. (2001) Genetic analysis of the chromosome segregation protein Spo0J of *Bacillus subtilis*: evidence for separate domains involved in DNA binding and interactions with Soj protein. *Mol Microbiol*, 41, 743–755.
- Berlatzky, I. A., Rouvinski, A. & Ben-Yehuda, S. (2008). Spatial organization of a replicating bacterial chromosome. *Proc Natl Acad Sci U S A*, 105, 14136–40.
- Bernardo, L. M. D., Johansson, L. U. M., Solera, D., Skärfstad, E., & Shingler, V. (2006). The guanosine tetraphosphate (ppGpp) alarmone, DksA and promoter affinity for RNA polymerase in regulation of s54-dependent transcription. *Mol Microbiol*, 60, 749–764.
- Bigot, S., Corre, J., Louarn, J. M., Cornet, F. & Barre, F. X. (2004). FtsK activities in Xer recombination, DNA mobilization and cell division involve overlapping and separate domains of the protein. *Mol Microbiol*, 54, 876–86.

- Biswas, S. B. & Biswas-Fiss, E. E. (2006). Quantitative analysis of binding of single-stranded DNA by *Escherichia coli* DnaB helicase and the DnaB x DnaC complex. *Biochemistry*, 45, 11505-13.
- Bramhill, D. & Kornberg, A. (1988). A model for initiation at origins of DNA replication. *Cell*, 54, 915-8.
- Brendler, T., Sawitzke, J., Sergueev, K. & Austin, S. (2000). A case for sliding SeqA tracts at anchored replication forks during *Escherichia coli* chromosome replication and segregation. *EMBO J*, 19, 6249-6258.
- Burbuly, D., Trach, K. A., Hoch, J. A. (1991). Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. *Cell*, 64, 545-552.
- Casadésus, J. & Low, D. (2006). Epigenetic Gene Regulation in the Bacterial World. *MMBR*, 70, 830-856.
- Cassler, M. R., Grimwade, J. E. & Leonard, A. C. (1995). Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. *EMBO J.*, 14, 5833-5841.
- Chiaramello, A. E. & Zyskind, J. (1990). Coupling of DNA replication to growth rate in *Escherichia coli*: a possible role for guanosine tetraphosphate. *J. Bacteriol*, 171, 4272-4280.
- Chodavarapu, S., Felczak, M. M., Yaniv, J. R. & Kaguni, J. M. (2008). *Escherichia coli* DnaA interacts with HU in initiation at the *E. coli* replication origin. *Mol Microbiol*, 67, 781-792.
- Collier, J. & Shapiro, L. (2009). Feedback control of DnaA-mediated replication initiation by replisome-associated HdaA protein in *Caulobacter*. *J Bacteriol*, 191, 5706-5716.
- Cunningham, K. A. & Burkholder, W. F. (2008). The histidine kinase inhibitor Sda binds near the site of autophosphorylation and may sterically hinder autophosphorylation and phosphotransfer to Spo0F. *Mol Microbiol*, 71, 659-677.
- Dubarry, N., Possoz, C., Barre, F. X. (2010). Multiple regions along the *Escherichia coli* FtsK protein are implicated in cell division. *Mol Microbiol*, 78, 1088-1100.
- Duigou, S., Knudsen, K., Skovgaard, O., Egan, E., L'bnér-Olesen, A. & Waldor, M. (2006). Independent control of replication initiation of the two *Vibrio cholera* chromosomes by DnaA and *rctB*. *J Bacteriol*, 108, 6419-6424.
- Egan, E. S. Fogel, M. A. & Waldor, M.K. (2005). Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol Microbiol*, 56, 1129-1138.
- Egan, E. S. & Waldor, M. K. (2003). Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell*, 114, 521-530.
- Erzberger, J., Pirruccello, M. M. & Berger, J. M. (2002). The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation. *EMBO J*, 21, 4763-4773.
- Ferullo, D. J, Cooper, D. L., Moore, H. R. & Lovett, S. T. (2009). Cell cycle synchronization of *E. coli* using the stringent response, with fluorescence labeling assays for DNA content and replication. *Methods*, 48, 8-13.
- Ferullo, D. J. & Lovett, S. T. (2008). The Stringent Response and Cell Cycle Arrest in *Escherichia coli*. *PLoS Genet*, 4, e1000300.

- Gama-Castro, S., Salgado, H., Peralta-Gil, M., Santos-Zavaleta, A., Muñiz-Rascado, L., Solano-Lira, H., Jimenez-Jacinto, V., Weiss, V., García-Sotelo, J., López-Fuentes, A., Porrón-Sotelo, L., Alquiciría-Hernández, S., Medina-Rivera, A., Martínez-Flores, I., Alquiciría-Hernández, K., Martínez-Adame, R., Bonavides-Martínez, C., Miranda-Ríos, J., Huerta, A., Mendoza-Vargas, A., Collado-Torres, L., Taboada, B., Vega-Alvarado, L., Olvera, M., Olvera, L., Grande, R., Morett, E. & Collado-Vides, J. (2010). RegulonDB version 7.0: transcriptional regulation of *Escherichia coli* K-12 integrated within genetic sensory response units (Sensor Units). *Nucleic Acids Res.*, 39, D98–D105.
- Gao, D. & McHenry, C. S. (2001). *Tau* binds and organizes *Escherichia coli* replication proteins through distinct domains. Domain III, shared by gamma and tau, binds delta, delta' and chi psi. *J Biol Chem*, 276, 4447–53.
- Gao, F. & Zhang, C. T. (2007). DoriC: a database of oriC regions in bacterial genomes. *Bioinformatics*, 23, 866–867.
- Georgescu, R. E., Yao, N. Y. & O'Donnell, M. (2010). Single-molecule analysis of the *Escherichia coli* replisome and use of clamps to bypass replication barriers. *FEBS Lett*, 584, 2596–605.
- Gil, R., Silva, F. J., Zientz, E., Delmotte, F., González-Candelas, F., Latorre, A., Rausell, C., Kamerbeek, J., Gadau, J. I., Hölldobler, B., van Ham, R. C., Gross, R. & Moya, A. (2003). The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. *Proc Natl Acad Sci USA*, 100, 9388–9393.
- Gorbatyuk, B. & Marczyński, G. T. (2005). Regulated degradation of chromosome replication proteins DnaA and CtrA in *Caulobacter crescentus*. *Mol. Microbiol.*, 55, 1233–1245.
- Han, J. S., Kang, S., Kim, S. H., Ko, M. J. & Hwang, S. D. (2004). Binding of SeqA protein to hemi-methylated GATC sequences enhances their interaction and aggregation properties. *J. Biol. Chem*, 279, 30236–30243.
- Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., Haft, D., Hickey, E., Peterson, J., Umayam, L., Gill, S., Nelson, K., Read, T., Tettelin, H., Richardson, D., Ermolaeva, M., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R., Nierman, W.I, White, O., Salzberg, S., Smith, H., Colwell, R., Mekalanos, J., Venter, C. & Frase, C. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*, 406, 477–83.
- Hung, D. Y., & Shapiro, L. (2002). A signal transduction protein cues proteolytic events critical to *Caulobacter* cell cycle progression. *Proc. Natl. Acad. Sci. USA*, 99, 13160–13165.
- Hwang, D. S., Thöny, B. & Kornberg, A. (1992). IciA protein, a specific inhibitor of initiation of *Escherichia coli* chromosomal replication. *J Biol Chem.*, 267, 2209–2213.
- Ishida, T., Akimitsu, N., Kashioka, T., Hatano, M., Kubota, T., Ogata, Y., Sekimizu, K. & Katayama, T. (2004). DiaA, a novel DnaA-binding protein, ensures the timely initiation of *Escherichia coli* chromosome replication. *J Biol Chem*, 279, 45546–45555.

- Erzberger, J. P., Pirruccello, M. M. & Berger, J. M. (2002). The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation *EMBO J.*, 21, 4763–4773.
- Jenal, U. (2009). The role of proteolysis in the *Caulobacter crescentus* cell cycle and development. *Res Microbiol*, 160, 687–695.
- Katayama, T., Ozaki, S., Keyamura, K. & Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and *oriC*. *Nat Rev Microbiol.*, 8, 163–170.
- Katayama, T. & Sekimizu, K. (1999). Inactivation of *Escherichia coli* DnaA protein by DNA polymerase III and negative regulations for initiation of chromosomal replication. *Biochimie*, 81, 835–40.
- Katayama, T., Kubota, T., Kurokawa, T., Crooke, E. & Sekimizu, K. (1998). The initiator function of DnaA protein is negatively regulated by the sliding clamp of the *E. coli* chromosomal replicase. *Cell*, 94, 61–71.
- Kennedy, S. P., Chevalier, F. & Barre, F. X. (2008). Delayed activation of Xer recombination at *dif* by FtsK during septum assembly in *Escherichia coli*. *Mol Microbiol*, 68, 1018–1028.
- Kim, M. S., Bae, S. H., Yun, S. H., Lee, H.J., Ji, S. C., Lee, J.H., Srivastava, P., Lee, S. H., Chae, H., Lee, Y., Choi, B. S., Chattoraj, D. K. & Lim, H. M. (2005). Cnu, a novel *oriC*-binding protein of *Escherichia coli*. *J Bacteriol.* 187, 6998–7008.
- Kitagawa, R., Ozaki, T., Moriya, S. & Ogawa, T. (1998). Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for *Escherichia coli* DnaA protein. *Genes Dev.*, 12, 3032–3043.
- Laub, M. T., Shapiro, L. & McAdams, H. H. (2007). Systems Biology of *Caulobacter*. *Annu. Rev. Genet*, 41, 429–441.
- Lee, P.S., Lin, D.C.-H., Moriya, S., & Grossman, A.D. (2003) Effects of the chromosome partitioning protein Spo0J (ParB) on *oriC* positioning and replication initiation in *Bacillus subtilis*. *J Bacteriol*, 185, 1326–1337.
- Lee, Y. S., Kim, H. & Hwang, D. S. (1996). Transcriptional activation of the *dnaA* gene encoding the initiator for *oriC* replication by IciA protein, an inhibitor of in vitro *oriC* replication in *Escherichia coli*. *Mol Microbiol.*, 19, 389–396.
- Leonard, A. C. & Grimwade, J. E. (2009). Initiating chromosome replication in *E. coli*: It makes sense to recycle. *Genes Dev*, 23, 1145–115.
- Lesley, J. A. & Shapiro, L. (2008). SpoT regulates DnaA stability and initiation of DNA replication in carbon-starved *Caulobacter crescentus*. *J. Bacteriol.*, 190, 6867–6880.
- Levine, A., Autret, A. & Seror, S. J. (1995). A checkpoint involving RTP, the replication terminator protein, arrests replication downstream of the origin during the Stringent Response in *Bacillus subtilis*. *Mol Microbiol*, 15, 287–295.
- Levine, L., Vannier, F., Dehbi, M., Henckes, G. & Séror, S. J. (1991). The Stringent Response Blocks DNA Replication Outside the *ori* Region in *Bacillus subtilis* and at the Origin in *Escherichia coli*. *J. Mol. Biol.*, 219, 605–613.
- Lewis, P. J. (2001). Bacterial chromosome segregation. *Microbiology*, 147, 519–526.

- Low, D. & Casadesús, J. (2008). Clocks and switches: bacterial gene regulation by DNA adenine methylation. *Current opinion in microbiology*, 11, 106-12.
- Lu, M., Campbell, J. L., Boye, E. & Kleckner, N. (1994). SeqA: a negative modulator of replication initiation in *E. coli*. *Cell*, 77, 413-426.
- Makowska-Grzyska, M. & Kaguni, J. M. (2010). Primase directs the release of DnaC from DnaB. *Mol Cell*, 37, 90-101.
- Marczynski, G. T. & Shapiro, L. (2002). Control of chromosome replication in *Caulobacter crescentus*. *Annu Rev Microbiol*, 56, 625-656.
- Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A. E., Barloy-Hubler, F., Galibert, F., Kondorosi, A. & Kondorosi, E. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. *Proc Natl Acad Sci USA*, 103, 5230-5235.
- Messer, W. & Weigel, C. (1997). DnaA initiator – also a transcription factor. *Mol. Microbiol.*, 24, 1-6.
- Mott, M. L. & Berger, J. M. (2007). DNA replication initiation: mechanisms and regulation in bacteria. *Nature Rev Microbiol.*, 5, 343-54.
- Murray, H. & Errington, J. (2008). Dynamic control of the DNA replication initiation protein DnaA by Soj/ParA. *Cell*, 135, 74-84.
- Neylon, C., Kralicek, A. V., Hill, T. M. & Dixon, N. E. (2005) Replication termination in *Escherichia coli*: structure and antihelicase activity of the Tus-ter complex. *Microbiol Mol Biol Rev.*, 69, 501-26.
- Noirot-Gros, M. F., Velten, M., Yoshimura, M., McGovern, S., Morimoto, T. & Ehrlich, S. D. (2006). Functional dissection of YabA, a negative regulator of DNA replication initiation in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA*, 103, 2368- 2373.
- Ogura, Y., Ogasawara, N., Harry, E. J. & Moriya, S. (2003). Increasing the ratio of Soj to Spo0J promotes replication initiation in *Bacillus subtilis*. *J Bacteriol*, 185, 6316-6324.
- Ozaki, S., Katayama, T., (2009). DnaA structure, function and dynamics in the initiation at the chromosomal origin. *Plasmid*, 62, 71-82.
- Reisenauer, A. & Shapiro, L. (2002). DNA methylation affects the cell cycle transcription of the CtrA global regulator in *Caulobacter*. *EMBO J.*, 21, 4969-4977.
- Reisenauer, A., Kahng, L. S., McCollum, S. & Shapiro, L. (1999). Bacterial DNA methylation: a cell cycle regulator? *J Bacteriol*, 181, 5135-5139.
- Reyes-Lamothe, R., Sherratt, D. J. & Leake, M. C. (2010). Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science*. 328, 498-501.
- Rowen, L. & Kornberg, A. (1978). Primase, the dnaG protein of *Escherichia coli* an enzyme which starts DNA chains. *J Biol Chem*, 253, 758-764.
- Ryan, V. T., Grimwade, J. E., Camara, J.E., Crooke, E. & Leonard, A.C. (2004). *Escherichia coli* prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. *Mol Microbiol.*, 51, 1347-59.
- Ryan, K. R. & Shapiro, L. (2003). Temporal and spatial regulation in prokaryotic cell cycle progression. *Review Literature And Arts Of The Americas*, 72, 367-394.
- Schaeffer, P. M., Headlam, M. J. & Dixon, N. E. (2005). Protein-protein interactions in the eubacterial replisome. *IUBMB Life*. 57, 5-12.

- Schreiber, G., Ron, E. Z. & Glaser, G. (1995). ppGpp-mediated regulation of DNA replication and cell division in *Escherichia coli*. *Curr Microbiol.*, 30, 27-32.
- Shamoo, Y. & Steitz, T. A. (1999). Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell*, 15, 155-66.
- Stancheva, I., Koller, T. & Sogo, J. M. (1999). Asymmetry of Dam remethylation on the leading and lagging arms of plasmid replicative intermediates, *EMBO J*, 18, 6542-6551.
- Stephens, C. M., Zweiger, G. & Shapiro, L. (1995). Coordinate cell cycle control of a *Caulobacter* DNA methyltransferase and the flagellar genetic hierarchy. *J Bacteriol*, 177, 1662-1669.
- Su'etsugu, M., Nakamura, K., Keyamura, K., Kudo, Y. & Katayama, T. (2008). Hda Monomerization by ADP Binding Promotes Replicase Clamp-mediated DnaA-ATP Hydrolysis. *J Biol Chem*, 283, 36118-36131.
- Swart, J. R. & Griep, M.A. (1995). Primer synthesis kinetics by *Escherichia coli* primase on single-stranded DNA templates. *Biochemistry*, 34, 16097-16106.
- Tan, M. H., Kozdon, J. B., Shen, X., Shapiro, L., & McAdams, H. H. (2010). An essential transcription factor, SciP, enhances robustness of *Caulobacter* cell cycle regulation. *Proc. Natl. Acad. Sci. USA*, 107, 18985-18990.
- Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Wagner, J. K., Marquis, K. A. & Rudner, D. Z. (2009). SirA enforces diploidy by inhibiting the replication initiator DnaA during spore formation in *Bacillus subtilis*. *Mol Microbiol*, 73, 963-974.
- Veening, J., Murray, H. & Errington, J. (2009). A mechanism for cell cycle regulation of sporulation initiation in *Bacillus subtilis*. *Genes Dev*, 23, 1959-1970.
- Wang, J. D. & Levin, P. A. (2009). Metabolism, cell growth and the bacterial cell cycle. *Nat Rev Microbiol*, 7, 822-827.
- Wang, J., Sanders, G. M. & Grossman, A. D. (2007). Nutritional Control of Elongation of DNA Replication by (p)ppGpp. *Cell*, 128, 865-875.
- Wang, X., Possoz, C. & Sherratt, D. J. (2005). Dancing around the divisome: asymmetric chromosome segregation in *Escherichia coli*. *Genes Dev.*, 19, 2367-2377.
- Wion, D. & Casadésus, J. (2006). N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. *Nature Rev Microbiol*, 4, 183-192.
- Wu, L. J. & Errington, J. (1994). *Bacillus subtilis* SpoIIIE protein required for DNA segregation during asymmetric cell division. *Science*, 264, 572-575.
- Zakrzewska-Czerwińska, J., Jakimowicz, D., Zawilak-Pawlik, A. & Messer, W. (2007). Regulation of the initiation of chromosomal replication in bacteria. *FEMS Microbiol Rev*, 31, 378-87.

Relevant Links:

<http://biocyc.org/>

<http://ecocyc.org/>

[http://ecoliwiki.net/colipedia/index.php/Welcome\\_to\\_EcoliWiki](http://ecoliwiki.net/colipedia/index.php/Welcome_to_EcoliWiki)

<http://regulondb.ccg.unam.mx/>

<http://www.ebi.ac.uk/uniprot/>

<http://www.ecogene.org/index.php>

<http://www.ncbi.nlm.nih.gov/>

<http://www.york.ac.uk/res/thomas/index.cfm>

# Propagating Epigenetic States During DNA Replication

Jennifer L. Jacobi and Ann L. Kirchmaier  
*Purdue University*  
USA

## 1. Introduction

The propagation of epigenetic states during DNA replication is critical for maintaining gene expression patterns across cell generations. Phenotypically diverse, but genetically identical, cells within a multicellular organism originate from a single cell, the zygote. During development, the cells derived from this zygote will divide and differentiate along multiple developmental pathways until reaching their final cell fates, with each cell-type expressing a different subset of their common genetic information. Differential gene expression in alternate cell-types in both unicellular and multi-cellular organisms is dependent, in part, upon post-translational modifications to DNA and histones. When the DNA itself is replicated, these post-translational modification patterns must also be replicated in order for epigenetic states to be inherited. The replication of epigenetic states involves many proteins acting in a concerted manner. Failure of one protein to act may have devastating effects on the cell and organism including loss of cell identity, inviability or disease. In this review, we will explore how multiple chromatin and gene expression states are transmitted epigenetically from mother cell to daughter cell during DNA replication and the contribution of replication factors to this process, taking examples from *S. cerevisiae*, *Drosophila* and mammals.

## 2. Nucleosomes as a minimal unit for carrying epigenetic information

The minimal repeating unit of chromatin in eukaryotes is the nucleosome, which is composed of approximately 147 bp of DNA wrapped 1.7 times around an octamer of histones containing two each of H2A, H2B, H3, and H4 (Luger et al., 1997). In chromatin, nucleosomes are separated from each other by differing lengths of linker DNA that can be bound by linker histones, such as H1 and H5 in mammals (Kornberg, 1977). This minimal, or “beads on a string”, order of chromatin conformation can then be folded into several higher order structures containing numerous other chromatin-associated proteins and RNAs. Some of the most dramatic of these structures include centromeres, highly condensed metaphase chromosomes and Barr bodies in mammals. Information integral to epigenetic processes and written onto these nucleosomes is found in the form of post-translational modifications to histones and chemical modifications to DNA. These modification patterns are critical for regulating diverse cellular processes ranging from gene expression and DNA repair to chromatin compaction. When DNA is replicated to pass on

genetic information to daughter cells, these nucleosomal modifications must also be duplicated to ensure regulatory and structural information related to the accessibility of that genetic information is also inherited.

Distinct histone modification patterns are found on newly synthesized histones relative to parental histones, which display altered modifications reflecting their locus-specific functions. H4 found within newly synthesized histone H3.1-H4 dimers are diacetylated at K5 and K12 (Sobel et al., 1995, Loyola et al., 2006). This H4 modification pattern is initiated in the cytoplasm, is highly conserved across evolution and is mediated by the acetyltransferase HAT1 (Sobel et al., 1995, Parthun, 2007). However, once H4 is incorporated into chromatin, these marks are removed within ~20 min and other patterns are then created (Taddei et al., 1999). In budding yeast, acetylation of H4 K91 also occurs prior to incorporation into chromatin (Ye et al., 2005). H4 K91 lies at the site of interaction between H3/H4 dimers and H2A/H2B dimers (English et al., 2006). The role of this modification in regulating nucleosome formation is not yet understood. However, nucleosomes containing H4 K91A mutants are more easily digested by micrococcal nuclease (Ye et al., 2005), raising the possibility that this modification regulates nucleosome stability.

In contrast to H4, the modifications present on newly synthesized H3 vary somewhat across organisms. In *Drosophila*, H3 is primarily acetylated on H3 K9 and/or K23 (Sobel et al., 1995). In contrast, in mammals, newly synthesized H3 is monomethylated on K9 (Loyola et al., 2006) and may also be acetylated at K56 prior to assembly (Das et al., 2009, Xie et al., 2009). In *S. cerevisiae*, newly synthesized H3 is acetylated at numerous sites including K9, K27 and K56 (Masumoto et al., 2005, Kuo et al., 1996, Adkins et al., 2007, Burgess et al., 2010). These modifications may aid in distinguishing parental from newly synthesized histones and facilitate nucleosome assembly by the chromatin assembly machinery (See Sec. 3). Revision of the modification patterns on newly synthesized histones to match pre-existing locus-specific patterns must occur during or shortly after DNA replication. Such changes mediated by enzymes targeted to the replication fork or those loci ensure successful propagation of epigenetic states from mother to daughter cell.

### 3. Chromatin disassembly and assembly during DNA replication

#### 3.1 Nucleosome organization and dynamics at the replication fork

During DNA replication,  $4 \pm 1$  nucleosomes are transiently destabilized at the replication fork by the DNA replication machinery. These nucleosomes must be disassembled in front of the fork for DNA polymerase to gain access to its template and then, after replication, be reassembled behind the fork to repackage the newly synthesized DNA into chromatin. During SV40 replication *in vitro*, this process destabilizes a region of ~650-1100 bp of DNA involving approximately two nucleosomes in front of the replication fork and a short region behind the fork on the daughter strands (Gasser et al., 1996). In SV40 minichromosomes, 0 - ~380 bp of unpackaged DNA lies ahead of the branchpoint signifying the replication fork and ~260 - ~440 bp of unpackaged DNA is present in the daughter strands behind the fork. The region comprising the nucleosomes adjacent to either side of the fork likely lacks linker histones. The nucleosome immediately upstream of the fork is partially disassembled, and may exist as a H3/H4 tetramer. After replication, nucleosome assembly occurs on the daughter strands once the length of DNA needed to wrap around the histone octamer has passed through the replication machinery (Sogo et al., 1986, Gasser et al., 1996).

This chromatin disassembly and reassembly process during DNA replication requires the removal of nucleosomes containing parental histones and the assembly of both parental and newly synthesized histones into nucleosomes. Parental (pre-replicative) H3/H4 have been shown to be removed from in front of the replication fork as H3/H4 dimers and then reassembled together with other parental H3/H4 dimers to reform parental (H3/H4)<sub>2</sub> tetramers on both daughter strands through the use of density labeling of parental H3/H4 and sedimentation analysis (Jackson, 1990, Yamasu & Senshu, 1990, Gruss et al., 1993) or stable isotope labeling of “old” H3.1 variants and mass spectrometry (Xu et al., 2010). In contrast, H2A/H2B dimers appear to be removed from in front of the fork prior to H3/H4, and then, along with newly synthesized H2A/H2B dimers, to be randomly incorporated onto nucleosomes containing either parental or newly synthesized (H3/H4)<sub>2</sub> tetramers behind the fork. The assembly of the basic structure of chromatin is then completed by the loading of a linker histone, e.g. H1 (Jackson, 1990). This pattern of nucleosome disassembly and assembly during DNA replication implicates parental H3/H4, and the modifications on these histones, as being the direct and critical instructions for the reformation of preexisting epigenetic states after passage of a replication fork through a chromosomal locus. These patterns on the parental nucleosomes may act as a guide for generating similar patterns on adjacent nucleosomes that contain newly synthesized histones, but direct demonstration of how this occurs has yet to be accomplished. Other key instructions will be found on the parental DNA strands themselves in the form of methylated and hydroxymethylated cytosines (See Sec. 4.2).

### 3.2 Chromatin disassembly at the replication fork

Several proteins have been implicated in nucleosome disassembly in front of the replication fork, including FACT and Asf1p (Fig. 1). FACT, or Facilitator of Chromatin Transcription, consists of Spt16p and Pob3p in yeast and SPT6 and SSRP1 in mammals. FACT is best understood for its role in transcription elongation where FACT replaces H2A/H2B dimers upon passage of the transcription machinery. FACT can also bind H3/H4 through Spt16p to promote nucleosome disassembly and reassembly during both transcription elongation and DNA replication (Stuwe et al., 2008, Belotserkovskaya et al., 2003). FACT localizes to replication foci in mammals (Hertel et al., 1999) and is required for DNA replication in *Xenopus* egg extracts (Okuhara et al., 1999). FACT is thought to remove H2A/H2B from in front of the replication fork, thereby facilitating DNA replication. Consistent with this function, Spt16p of FACT interacts with the Mcm4p subunit of the replicative MCM helicase (Tan et al., 2006). FACT also co-purifies with DNA pol $\alpha$  (Wittmeyer & Formosa, 1997) and promotes replication fork progression (Gambus et al., 2006). Also, Pob3p interacts with Rfa1p, a subunit of RPA that binds ssDNA during DNA replication (VanDemark et al., 2006). These interactions implicate FACT in promoting nucleosome disassembly and deposition, respectively, on either side of the replication fork.

Anti-Silencing Factor 1, Asf1p (mammalian Asf1a & Asf1b), is an evolutionarily conserved chromatin assembly factor that was discovered in a screen for genes, which when overexpressed, led to silencing defects in *S. cerevisiae* (Le et al., 1997). Asf1 binds H3/H4 dimers through a surface on H3/H4 that associates with a second H3/H4 dimer in the context of a nucleosome. This Asf1-H3/H4 interaction may prevent premature (H3/H4)<sub>2</sub> tetramer formation prior to nucleosome assembly (English et al., 2006, Tagami et al., 2004). Asf1p participates in chromatin assembly during both transcription and replication (Green

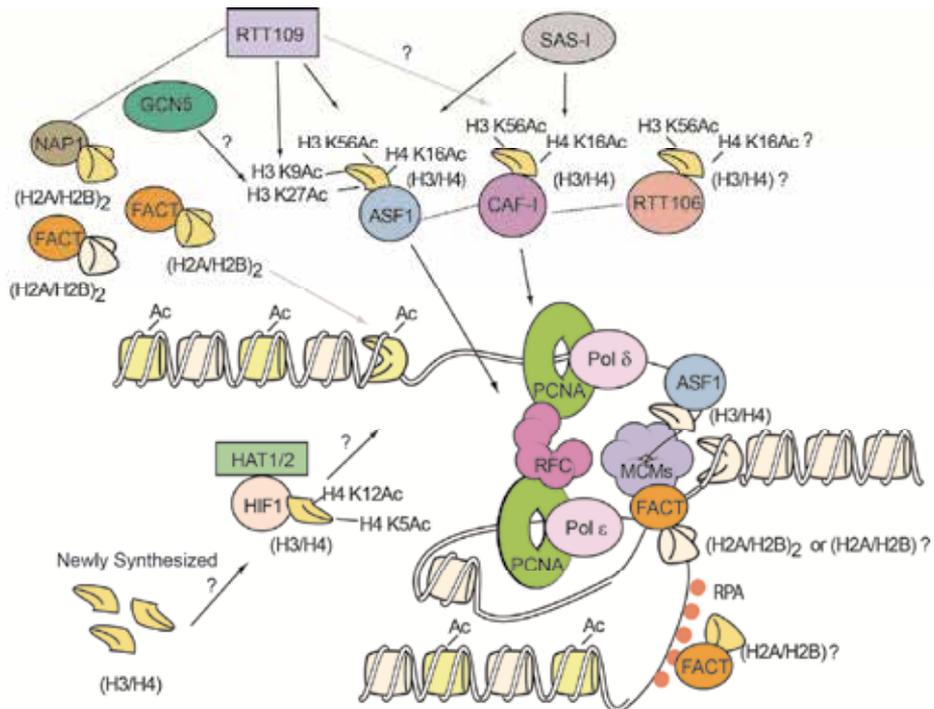


Fig. 1. Replication-coupled chromatin assembly in *S. cerevisiae*. Parental histones (beige) are removed from in front of the fork by assembly factors FACT (orange) and Asf1p (blue), which both interact directly or indirectly with the MCM helicase (blue-grey). FACT also binds to single stranded DNA binding protein Rpa1p of the RPA complex (coral). Whether FACT binds H2A/H2B dimers or tetramers is unclear. Newly synthesized H3/H4 dimers (yellow) are acetylated at H4 K5 and 12 by Hat1p in complex with Hat2p (dull green) and assembly factor Hif1p (peach). Rtt109p (light purple) binds to Asf1p bound to newly synthesized H3/H4 dimers (yellow) and acetylates H3 K56. If, where, and how newly synthesized H3/H4 are transferred from Hat1p/Hat2p/Hif1p to Asf1p or other assembly factors is unknown. CAF-1 (magenta) binds H3/H4 containing H3 K56ac, and possibly H3 K9ac and H4 K16ac, and interacts with Asf1p through the Cac2p subunit (dotted line). Assembly factor, Rtt106p (light red), binds newly synthesized H3/H4 containing H3 K56ac, and possibly H3 K9ac and H4 K16ac, and also interacts with CAF-1 through the Cac1p subunit. SAS-I (grey) associates with CAF-1 or Asf1p bound to H3/H4. Sas2p of SAS-I acetylates H4 K16. PCNA (bright green) tethers Pol $\delta$  (light pink) and Pol $\epsilon$  (light pink) to the replication fork. PCNA is loaded onto DNA by RFC (dark pink). CAF-1 and Asf1p associate with the replication fork through interactions between Cac1p and PCNA and Asf1p and the Rfc2-4p subunits of RFC. Gcn5p (turquoise) and Rtt109p acetylate residues including H3 K9 and K27 and influence chromatin assembly during replication. Assembly of new and parental histones behind the fork is facilitated by FACT, and likely Nap1p (brown), which bind H2A/H2B, as well as CAF-1, Rtt106p, Asf1p, and possibly Hif1p, which bind H3/H4. Current models predict that Asf1p transfers H3/H4 dimers to CAF-1 for assembly. However, Asf1p may also directly assemble H3/H4 dimers. Question marks indicate where mechanisms are unclear.

et al., 2005, Sanematsu et al., 2006). Asf1p also functions in chromatin disassembly; Asf1p increases the rate of histone eviction at the yeast *PHO5* and *PHO8* promoters during transcription (Korber et al., 2006) and globally removes H3/H4 from chromatin (Adkins & Tyler, 2004). In *Drosophila*, Asf1 localizes to replication foci during S phase and depletion of Asf1 from *Drosophila* or mammalian cells results in delayed progression through S phase and inefficient DNA replication (Tyler et al., 1999, Groth et al., 2005, Sanematsu et al., 2006, Schulz & Tyler, 2006). These defects are related, in part, to a proposed role of Asf1p in facilitating nucleosome disassembly in front of the replication fork in conjunction with the FACT complex. Consistent with this model, Asf1 associates with the MCM helicase via H3/H4, and depletion of Asf1p slows DNA duplex unwinding by MCMs (Groth et al., 2007a, Groth et al., 2007b). Moreover, when fork progression is inhibited by treating cells with hydroxyurea, MCMs continue to unwind DNA and complexes of Asf1-H3/H4-MCMs accumulate. Consistent with disassembly, H3/H4 in these complexes carry marks associated with parental histones, H4 K16ac and H3 K9me3, rather than newly synthesized H3/H4 (Groth et al., 2007a). Other factors must also participate in disassembly of H3/H4 during replication as *ASF1* is not essential in *S. cerevisiae*.

### 3.3 Chromatin assembly at the replication fork

Several chromatin assembly factors have been linked to replication-coupled assembly behind the fork including Asf1p, CAF-1, Rtt106p and FACT (Fig. 1). In *Drosophila* and other organisms, the histone variant H3.1 is assembled into nucleosomes during replication-coupled chromatin assembly in S phase whereas the variant H3.3 is incorporated into chromatin throughout the cell cycle (Ahmad & Henikoff, 2002). The human Asf1 homologs, Asf1a and Asf1b, associate with both H3.1 and H3.3 (Tagami et al., 2004), consistent with their dual roles in transcription and replication-coupled chromatin assembly. In contrast, Chromatin Assembly Factor 1, CAF-1, comprised of p150, p60, and p48 in mammals and Cac1p, Cac2p, and Cac3p in *S. cerevisiae*, associates with newly synthesized H3/H4 (Kaufman et al., 1995) and mediates their incorporation into chromatin during DNA replication (Smith & Stillman, 1989, Verreault et al., 1996). In mammals, CAF-1 associates with H3.1, but not H3.3 (Tagami et al., 2004), and is required for progression through S phase (Hoek & Stillman, 2003, Ye et al., 2003), consistent with CAF-1 playing a critical role in assembly during DNA replication and repair (Smith & Stillman, 1989, Moggs et al., 2000, Kamakaka et al., 1996, Gaillard et al., 1996). CAF-1 is recruited to replication forks through binding of the Cac1p subunit of CAF-1 to PCNA (Zhang et al., 2000, Shibahara & Stillman, 1999). Cac1p contains a PCNA-binding motif and mutations in this region disrupt CAF-1-PCNA interactions in pull down experiments as well as result in silencing defects in budding yeast (Krawitz et al., 2002).

In *S. cerevisiae*, *Drosophila*, and human cells, Asf1p functions with CAF-1 to promote rapid nucleosome assembly during DNA replication and repair (Sharp et al., 2001, Tyler et al., 1999, Mello et al., 2002). However, CAF-1 does not enhance replication-independent histone deposition, implying CAF-1 may be exclusively involved in replication-coupled chromatin assembly (Sharp et al., 2001). Asf1p is proposed to transport H3/H4 dimers to CAF-1 for deposition onto DNA. Consistent with this model, CAF-1-dependent nucleosome assembly is stimulated by Asf1p. In the absence of Asf1p, H3/H4 are not readily transferred to CAF-1 (Tyler et al., 1999, Sharp et al., 2001, Mello et al., 2002, Groth et al., 2005). Asf1 interacts with CAF-1 via the Cac2p subunit (Krawitz et al., 2002, Mello et al., 2002, Tyler et al., 2001), and H3 mutants that do not bind to human Asf1 can associate with CAF-1 (Galvani et al., 2008). In

addition, Asf1p binds to Replication Factor-C, RFC, which loads PCNA onto DNA and this interaction may localize Asf1p behind the fork. RFC loaded onto nicked templates is sufficient to target Asf1p to DNA, and the Rfc2-5p subunits of RFC co-precipitate with Asf1p (Franco et al., 2005). *In vivo*, *asf1Δ rfc1-1* mutants exhibit synthetic growth defects and accumulate between S phase and the metaphase-to-anaphase transition. Similar slow growth phenotypes are not observed in *cac1Δ rfc1-1* mutants (Kaufman et al., 1998, Kaufman et al., 1997, Franco et al., 2005).

In *S. cerevisiae*, CAF-1 also interacts with the assembly factor Rtt106p through Cac1p (Huang et al., 2005), and Rtt106p and Asf1p co-purify from *in vivo* extracts (Lambert et al., 2010). Rtt106p binds H3/H4 dimers through PH domains, similar to Pob3p, and exhibits chromatin assembly activity *in vitro* (Huang et al., 2005, Li et al., 2008). Chromatin assembly by CAF-1 and Rtt106p is stimulated by H3 K56ac (Li et al., 2008), a modification catalyzed by the acetyltransferase Rtt109p while H3/H4 dimers are bound by Asf1p (Tsubota et al., 2007, Driscoll et al., 2007). Cells lacking CAF-1, Asf1p or Rtt106p exhibit defects in Sir-mediated silencing in *S. cerevisiae* (Huang et al., 2005, Huang et al., 2007, Kaufman et al., 1997, Tyler et al., 1999). Silencing defects are more severe in *cac1 rtt106*, and *asf1 cac1* mutants relative to single mutants, indicating that multiple assembly pathways impact epigenetically silenced chromatin (Huang et al., 2005, Tyler et al., 1999).

Several additional factors participate in nucleosome assembly during DNA replication, including FACT (and potentially Nap1p), the INO80 complex, and ACF1-SNF2H. As outlined above, FACT-RPA interactions via binding of Pob3p to Rfa1p may promote H2A/H2B deposition behind the replication fork (VanDemark et al., 2006), but the understanding of H2A/H2B disassembly and assembly during DNA replication lags behind that of H3/H4. Also, the ATP-dependent chromatin remodeler Ino80p localizes to origins of replication and replication forks during entry into S phase, is required continuously for fork progression under replication stress (Papamichos-Chronakis & Peterson, 2008) and functions in DNA repair (Morrison et al., 2004, van Attikum et al., 2004). How these proteins interact with other chromatin assembly factors during DNA replication remains to be explored. Currently, it is unclear whether certain factors are targeted to replication forks as they pass through some regions of the genome but not others, or the extent to which specialized factors to promote certain epigenetic processes. In support of some factors being critical for replication through epigenetically silenced loci, ACF1-SNF2H facilitates replication through heterochromatin in mammals (Collins et al., 2002). And, CAF-1 and Rtt106p contribute to the recruitment and spreading of Sirs in silent chromatin in budding yeast (Huang et al., 2007).

## 4. Propagation of chromatin modifications and epigenetic states

### 4.1 Histone modifications and replication factors in heterochromatin formation

The integrity of silent chromatin is influenced by the composition of nucleosomes at silenced loci as well as elsewhere throughout the genome. In budding yeast, transcriptionally active loci are enriched in acetylated histones and certain methylated forms of histones (e.g. methylated H3 K4 and H3 K36), whereas histones in silenced loci (rDNA locus, telomeres and silent mating-type loci *HML* and *HMR*) are hypoacetylated and hypomethylated (Bernstein et al., 2002, Braunstein et al., 1996, Katan-Khaykovich & Struhl, 2005, Rusche et al., 2002, Suka et al., 2001). Overexpression or loss of histone modifying enzymes often results in silencing defects. During silent chromatin formation, the Sirs, Sir1-4p, are

recruited to silencers flanking the *HM* loci. Sir2-4p then spread across *HMR* as the deacetylase Sir2p removes acetyl groups from H3 and H4 to facilitate nucleosomal binding by the Sirs. Once formed, silent chromatin is inherited efficiently as the genome is duplicated each S phase (Rusche et al., 2002, Hoppe et al., 2002) (see also Luo et al., 2002). Inappropriately modified histones can prevent Sirs from interacting stably with silent loci. This can occur either through disrupting Sir binding to nucleosomes or, upon global loss of histone modifications, through redistribution of Sirs to other genomic regions containing hypoacetylated and/or hypomethylated nucleosomes. Re-localization can deplete the pool of Sirs available for forming silent chromatin at appropriate loci as well as result in silent chromatin formation at inappropriate sites in the genome (Singer et al., 1998, van Leeuwen et al., 2002). Once formed, this “off target” silent chromatin can be propagated epigenetically during DNA replication.

Defects several factors involved in DNA replication and replication-coupled nucleosome assembly affect silencing, including Pol30p (PCNA), Rfc1p, Dna2p, Orc1p, Orc2p, Orc5p, Cdc7p, Cdc45p, Pol $\epsilon$ , Hif1p, CAF-1, Asf1p and Rtt106p (Axelrod & Rine, 1991, Foss et al., 1993, Kaufman et al., 1997, Loo et al., 1995, Smith et al., 1999, Zhang et al., 2000, Singer et al., 1998, Ehrenhofer-Murray et al., 1999, Huang et al., 2005, Poveda et al., 2004, Triolo & Sternglanz, 1996). These factors impact silent chromatin in multiple ways. For example, Orc1p binds Sir1p and facilitates Sir recruitment to silencers containing ARS elements adjacent to the *HM* loci (Gardner et al., 1999, Zhang et al., 2002, Rusche et al., 2002). Other ORC subunits also affect silencer function, but the role of ORC in silencing and replication initiation can be genetically separated (Dillin & Rine, 1997). The mechanisms by which some factors, including Cdc7p, Cdc45p and Pol $\epsilon$ , contribute to silencing have yet to be elucidated.

Silencing defects associated with CAF-1, Asf1p, PCNA, Rfc1p, Hif1p and Rtt106p, are linked to replication-coupled chromatin assembly and the misregulation of replication-coupled histone modifications (Fig. 1). Yeast *pol30* mutants with silencing defects have defects in CAF-1 and *ASF1*-dependent pathways (Zhang et al., 2000, Sharp et al., 2001). The silencing defects in *pol30*, *cac1*, *asf1* and *rfc1-1* mutants reflect, in part, misregulation of histone acetylation by SAS-I, Rtt109p and/or Gcn5p leading to hypoacetylation of at least H3 K9, H3 K56 and H4 K16 throughout the genome (Miller et al., 2010, Miller et al., 2008).

CAF-1 and Asf1p both bind to the H4 K16 acetyltransferase complex SAS-I (Meijsing & Ehrenhofer-Murray, 2001, Osada et al., 2001) and loss of the catalytic subunit of SAS-I, Sas2p, alters the chromosomal distribution of Sirs and results in silencing defects (Kimura et al., 2002, Meijsing and Ehrenhofer-Murray, 2001, Osada et al., 2001, Reifsnnyder et al., 1996, Suka et al., 2002). Asf1p also binds the H3 K56 acetyltransferase Rtt109p, stimulates H3 K56ac *in vitro* and is required, along with *RTT109*, for H3 K56ac in S phase *in vivo* (Driscoll et al., 2007, Recht et al., 2006, Schneider et al., 2006, Tsubota et al., 2007, Han et al., 2007). Misregulation of H3 K56ac leads to silencing defects (Hyland et al., 2005, Miller et al., 2008, Xu et al., 2007, Sharp et al., 2001) and SAS-I and *rtt109* mutants have silencing phenotypes similar to those of *cac1*, *asf1* and *pol30* mutants (Ehrenhofer-Murray et al., 1999, Meijsing & Ehrenhofer-Murray, 2001, Miller et al., 2010, Miller et al., 2008, Osada et al., 2001). PCNA interacts with Rtt109p and SAS-I *in vivo*, but this interaction is lost in *pol30* mutants with defects in CAF-1- and Asf1p-dependent pathways (Miller et al., 2010), implying that acetylation of H3 K56 and H4 K16 are coupled to DNA replication. In addition, Asf1p binds the histone chaperone complex Hif1p/Hat1p/Hat2p in a Hat2p-dependent manner

(Fillingham et al., 2008). Cells lacking the chromatin assembly factor Hif1p or the acetyltransferase Hat1p have telomeric silencing defects, implicating this complex and Hat1p-dependent modifications in regulating silencing (Kelly et al., 2000, Poveda et al., 2004). Whether SAS-I, Rtt109p and/or Hat1p associate with Asf1p independently, simultaneously or sequentially, and how these complexes interact with Rtt106p and CAF-1 remains to be clarified.

The acetyltransferase Gcn5p has been proposed to facilitate replication-coupled chromatin assembly through modifying the N-terminal tail of H3. Consistent with this model, co-precipitation of H3 with the Cac2p subunit of CAF-1 is dramatically reduced in *gcn5* or H3 K5R mutants (Burgess et al., 2010). And, in *gcn5* or H3 K5R mutants, reduced levels of H3 containing modifications of newly synthesized histones, K9ac, K27ac and K56ac, are incorporated into chromatin adjacent to an early firing replication origin in cells arrested in early S phase (Burgess et al., 2010). *RTT109* and *GCN5*-dependent H3 K9ac (Fillingham et al., 2008, Adkins et al., 2007, Berndsen et al., 2008, Kuo et al., 1996) is also defective in *po30* mutants (Miller et al., 2010). This loss of H3 K9ac is consistent with loss of interactions between *pol30p* mutants and Rtt109p (Miller et al., 2010), but could also indicate the activity of Gcn5p during chromatin assembly was compromised in *pol30* mutants.

PCNA localizes numerous factors to the replication fork to propagate epigenetic states in mammals as well (Fig. 2). PCNA binds the maintenance DNA methyltransferase Dnmt1 (Chuang et al., 1997, Iida et al., 2002) (See Sec. 4.2). PCNA also recruits CAF-1 to DNA and promotes CAF-dependent chromatin assembly *in vitro* (Moggs et al., 2000, Shibahara & Stillman, 1999). MBD1, a methyl CpG binding protein and SETDB1, a H3 K9 methyltransferase, are, in turn, targeted to replication foci by CAF-1 and together with 5-mC DNA, MBD1 and SETDB1 promote stable heterochromatin formation (Sarraf & Stancheva, 2004, Moldovan et al., 2007). In addition, a mammalian H4 K20 methyltransferase, SET8, binds PCNA and co-localizes with PCNA at replication foci *in vivo* (Huen et al., 2008). Monomethylation of H4 K20 by SET8 is important for progression through S phase (Huen et al., 2008) and methylated H4 K20 is enriched in heterochromatic regions in multiple species (e.g. Schotta et al., 2004). In *Drosophila*, mutants of *mus209*, a PCNA ortholog, also suppress position-effect variegation (Henderson et al., 1994), but why this occurs is unclear.

Another illustrative example of silent chromatin formation can be found during development in *Drosophila*. Polycomb group (PcG) proteins maintain transcriptional repression patterns of the homeotic (*Hox*) genes, which control segmental identities and body patterning. PcG proteins are recruited to *cis*-acting PcG response elements, PRE, at target loci. There, Polycomb Repressive Complexes 1 and 2, PRC1 and PRC2, establish silent chromatin, which is then propagated over multiple cell generations to maintain *Hox* genes in a silenced state during development. Variants of PRC2 complexes contain a H3 K27 methyltransferase Ezh2, ESC, Suz12 and the histone binding protein Nurf55 (p55 of dCAF-1). The Polycomb, Pc, subunit of PRC1 can bind H3 K27me3, implying the catalytic function of PRC2 reinforces the association of PRC1 with chromatin (Min et al., 2003, Fischle et al., 2003b, Francis, 2009). Consistent with chromatin structure containing PcG proteins being inherited during DNA replication as opposed to re-established *de novo* after replication, recent studies using an *in vitro* SV40 replication system have demonstrated PRC1 remains associated with both naked DNA and chromatin templates upon DNA replication (Francis et al., 2009), implying PRC1 was passed from in front of to behind the fork along with parental H3/H4.

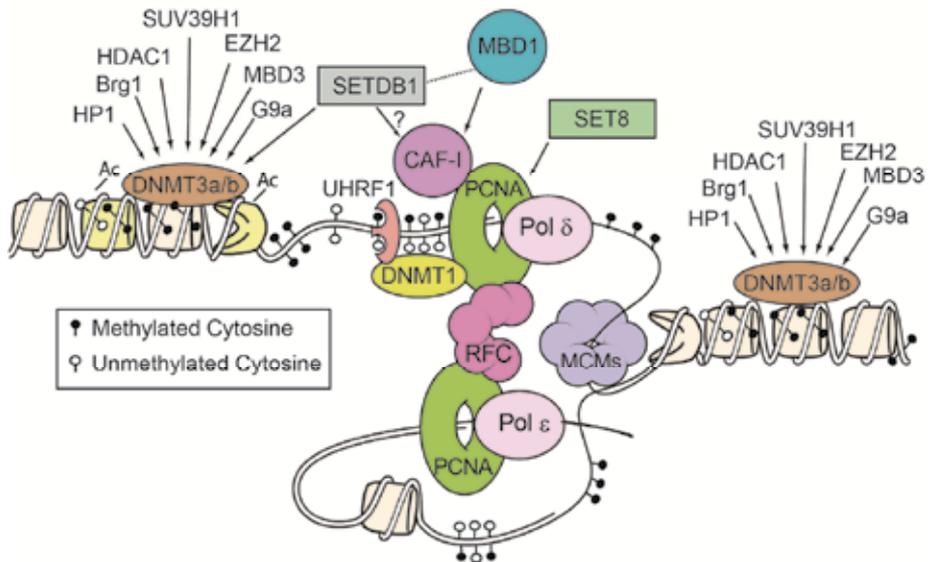


Fig. 2. Propagation of DNA methylation. Hemimethylated cytosines on newly synthesized DNA are recognized by UHRF1 (pink). The unmethylated daughter strand is methylated by the maintenance methyltransferase, Dnmt1, (yellow) which is localized to the replication fork through interactions with PCNA (bright green) and UHRF1. Dnmt3a/3b (brown) are *de novo* methyltransferases that localize to nucleosome-bound DNA and also interact with UHRF1 (not shown). Factors interacting with Dnmt3a/3b, include heterochromatin protein HP1, chromatin remodeler Brg1, histone deacetylase HDAC1, H3 K9 methyltransferases SUV39H1 and G9a, H3 K27 methyltransferase EZH2 and methyl DNA binding protein MBD3. The H3 K9 methyltransferase SETDB1 (grey) also interacts with Dnmt3a/3b and is localized to the replication fork through the chromatin assembly factor CAF-1 (magenta). CAF-1, which binds PCNA, also recruits the methyl DNA binding protein MBD1 (teal) that directly interacts with SETDB1. The mechanism for interaction between SETDB1 and CAF-1 is unclear and denoted by the question mark. The H4 K20 methyltransferase SET8 (dark green) is also targeted to the replication fork through interactions with PCNA.

#### 4.2 Insights into mechanisms of epigenetic inheritance: DNA methylation

DNA methylation plays an important role in epigenetic processes during development by impacting a range of biological functions including gene expression, genome integrity, imprinting, and aging, as well as by contributing to diseases ranging from neuronal defects to cancer when misregulated. 5-methylcytosine, 5-mC, constitutes ~2-8% of the cytosines in human genomic DNA and occurs primarily within CpG dinucleotides, although non-CpG methylation can also occur (Gowher and Jeltsch, 2001, Ramsahoye et al., 2000). Maintaining the average methylation state of a locus during DNA replication, rather than the individual sites of DNA methylation is generally more important for maintaining the proper function, or expression state, of that locus. Consistent with this model, different regions of the genome tend to be either hypo- or hypermethylated. And, in methylated regions, slight variations in methylation patterns are commonly found at individual loci in both cell lines and tissues (Meissner et al., 2008, Zhang et al., 2009). Methylation events contributing to

these varied patterns have been quantified for the CpG island at the *FRM1* locus. At *FRM1*, the fidelity of maintenance methylation is  $\sim 0.96$  and the probability of *de novo* methylation events having occurred per site per round of replication is  $\sim 0.17$  (Laird et al., 2004). Thus, DNA methylation occurs stochastically (Riggs & Xiong, 2004) and can vary slightly from cell to cell without altering an epigenetic state.

DNA methylation is mediated by a family of DNA methyltransferases, Dnmts, which are classified as *de novo* (Dnmt3a/3b) or maintenance (Dnmt1) methyltransferases, according to their primary role in establishing new methylation patterns or copying existing patterns onto newly synthesized DNA upon DNA replication. While methylating cytosines, Dnmts flip the target base out of the DNA helix and into a hydrophobic pocket to catalyze the transfer of the methyl group from S-adenosyl-L-methionine to the C5 position of cytosine to create 5-mC plus S-adenosyl-L-homocysteine, but their specificity for unmethylated versus hemimethylated DNA varies. The *de novo* methyltransferases Dnmt3a/3b readily methylate both unmethylated and hemimethylated DNA (Okano et al., 1998, Gowher & Jeltsch, 2001) and are critical for establishing proper DNA methylation patterns in early development in mammals (Okano et al., 1999). Dnmt3a/3b also help maintain DNA methylation within heterochromatin (Chen et al., 2003, Jeong et al., 2009, Liang et al., 2002). Dnmt3a/3b can interact with an additional Dnmt family member, Dnmt3L, which stimulates their catalytic activity *in vitro* and *in vivo*, despite Dnmt3L itself being catalytically inactive (Chedin et al., 2002, Chen et al., 2005, Gowher et al., 2005). In this role, Dnmt3L acts as a regulatory factor and is critical for proper methylation of imprinted genes and male germ cell development (Bourc'his et al., 2001, Hata et al., 2002, Webster et al., 2005).

Dnmt3a/3b also interact with several proteins to promote silenced epigenetic states (Fig. 2). Dnmt3a/3b binding partners include the histone deacetylase HDAC1 (Fuks et al., 2001), the histone methyltransferases SUV39H1 (Fuks et al., 2003), SETDB1 (Li et al., 2006), G9a (Epsztejn-Litman et al., 2008, Feldman et al., 2006) and EZH2 (Vire et al., 2006), the heterochromatin protein HP1 (Fuks et al., 2003, Smallwood et al., 2007), the 5-mC binding protein MBD3 and the chromatin remodeling factor Brg1 (Datta et al., 2005). These interactions all contribute to silent chromatin formation. For example, the H3 K9 methyltransferase G9a facilitates *de novo* DNA methylation and gene inactivation through recruiting Dnmt3a/3b and HP1 to multiple early embryonic genes to drive heterochromatin formation (Epsztejn-Litman et al., 2008, Feldman et al., 2006). Similarly, the H3 K27 methyltransferase Ezh2 recruits Dnmt3a/3b to chromosomal loci to promote DNA methylation and heterochromatin formation (Vire et al., 2006). Once formed, this heterochromatin and the associated modification patterns will be propagated epigenetically during DNA replication.

In addition to interacting with chromatin-modifying enzymes and structural components of heterochromatin, Dnmt3a/3b binds histones to promote DNA methylation. Dnmt3a/3b bind to the N-terminal tail of H3 lacking methylated K4 (Otani et al., 2009, Zhang et al., 2010, Ooi et al., 2007) and DNA methylation tends to be low at active promoters, which are enriched for H3 K4me3 (Hodges et al., 2009, Zhang et al., 2009). Dnmt3a/3b also preferentially bind H3 K36me3 and enhances DNA methylation of a nucleosomal substrate by Dnmt3a (Zhang et al., 2010). Like H3 K36me3, DNA methylation is enriched in bodies of active genes, especially in exons (Weber et al., 2007, Hodges et al., 2009, Kolasinska-Zwierz et al., 2009). Thus, Dnmt3a/3b-H3 interactions contribute to genome-wide chromatin modification patterns in transcriptionally active regions as well.

DNA methylation patterns established during development must be faithfully propagated throughout the lifespan of an organism via maintenance methylation during DNA

replication. The maintenance methyltransferase Dnmt1 localizes to replication foci in S phase (Leonhardt et al., 1992) and associates with the replication fork (Easwaran et al., 2004). Dnmt is loaded onto DNA through transient interactions with PCNA and binding to PCNA promotes the activity of Dnmt1 (Chuang et al., 1997, Iida et al., 2002, Schermelleh et al., 2007). By preferentially methylating hemimethylated DNA over unmethylated DNA (Fatemi et al., 2001, Goyal et al., 2006), Dnmt1 copies parental DNA methylation patterns onto newly synthesized daughter strands (Fig. 2). *In vitro*, Dnmt1 and PCNA form a complex with HDAC1, and Dnmt1/PCNA/HDAC1 co-localize *in vivo* (Chuang et al., 1997, Fuks et al., 2000, Milutinovic et al., 2002).

Dnmt1 also associates with newly synthesized DNA through binding UHRF1. UHRF1 is a multifunctional protein that reads hemimethylated DNA and histone modifications to facilitate inheritance of epigenetic states. UHRF1 targets Dnmt1 to newly synthesized DNA by specifically binding hemimethylated DNA (Arita et al., 2008, Avvakumov et al., 2008, Bostick et al., 2007, Hashimoto et al., 2008). UHRF1 co-localizes with PCNA and Dnmt1 at replicating heterochromatin and is required for maintaining DNA methylation in mammals (Bostick et al., 2007, Sharif et al., 2007). Disruption of Dnmt1/PCNA/UHRF1 interactions leads to global DNA hypomethylation and promotes tumor formation (Hervouet et al., 2010).

UHRF1 also binds H3 K9me2 and H3 K9me3, through a PHD domain. A second domain in UHRF1, SRA, also contributes to binding affinity (Karagianni et al., 2008). In addition, this SRA domain facilitates binding 5-mC DNA (Arita et al., 2008, Avvakumov et al., 2008, Hashimoto et al., 2008) and both domains are required to localize UHRF1 to pericentric heterochromatin. Consistent with H3 K9 methylation reinforcing DNA methylation during replication, localization of UHRF1 is reduced in cells overexpressing the H3 K9 demethylase JMJD2A (Karagianni et al., 2008). UHRF1 plays a second role in epigenetic processes by influencing histone modification states. UHRF1 contains a C3HC4 RING finger motif and acts as an E3 ubiquitin ligase targeting H3 *in vitro* and *in vivo*. This ubiquitination activity is important for maintaining higher order chromatin structure *in vivo* (Citterio et al., 2004, Karagianni et al., 2008). UHRF1/Dnmt1 also form a complex with the deacetylase HDAC1 and the H2A K5 acetyltransferase Tip60. Depletion of UHRF1 results in hypoacetylation of H2A K5 (Achour et al., 2009). Analyses how Dnmt1/PCNA/UHRF1 and UHRF1-hemimethylated DNA interactions regulate these histone modifications should clarify the extent to which deacetylation by HDAC1, ubiquitination by UHRF1 and acetylation of H2A K5 by TIP60 are coupled to DNA replication.

Although Dnmt1 and Dnmt3a/3b display preferences for different substrates, maintenance methyltransferases also participate in *de novo* methylation and *de novo* methyltransferases function in maintaining methylation patterns. Dnmt1 plays a secondary role in *de novo* DNA methylation at unmethylated loci through the conversion of hemimethylated DNA created by Dnmt3a/3b to fully methylated DNA (Fatemi et al., 2002, Feltus et al., 2003). Likewise, Dnmt3a/3b facilitate maintaining DNA methylation states upon DNA replication, particularly in chromosomal regions that are highly methylated or repetitive (Jones & Liang, 2009). Consistent with this, proliferating mouse embryonic stem cells lacking Dnmt3a/3b lose DNA methylation over time, despite the continued presence of Dnmt1 (Chen et al., 2003, Liang et al., 2002). Thus, although necessary, Dnmt1 alone is insufficient to maintain normal methylation levels on newly replicated DNA. Dnmt3a/3b likely aid in maintenance methylation through interacting directly with UHRF1 (Meilinger et al., 2009) as well as through catalyzing *de novo* methylation events within highly methylated regions, especially

in heterochromatic regions (Jeong et al., 2009). Thus, crosstalk between replication factors and chromatin modifying machinery reinforces propagation of DNA methylation, histone modifications and structural components of silent chromatin to maintain epigenetic states.

## 5. Switching of epigenetic states

### 5.1 Establishment, maintenance and inheritance of epigenetic states

Extensive reprogramming of epigenetic states occurs during primordial germ cell development, in mammalian early embryonic development and upon cell-type differentiation throughout an individual's lifespan (Sasaki & Matsui, 2008). Concepts for understanding the formation and stability of silent chromatin have been developed from the analysis of repression in the bacteriophage lambda (Ptashne, 1992). Studies in lambda showed a protein could facilitate the initial inactivation of a target gene, but once repressed, the maintenance of that the gene in its inactive state no longer required that protein. During silent chromatin formation, such proteins are said to be important for establishing, but not maintaining, the silenced state. In contrast, proteins that are required constantly to keep a repressed gene inactive are considered to be necessary for maintaining the silenced state. Proteins that are involved in the inheritance of silenced states facilitate the propagation of that state to subsequent cell generations.

An example of a protein important for establishing, but not maintaining, epigenetic states is Sir1p from budding yeast. Sir1p facilitates establishment by increasing the probability of Sir proteins being recruited to the *HM* loci (Rusche et al., 2002). Cells lacking *SIR1* can exist in two populations; transcriptionally active or silenced. Each population is stable over multiple generations, indicating that Sir1p is not required for maintaining or inheriting the different expression states (Pillus & Rine, 1989, Xu et al., 2006). These silenced and transcriptionally active cells will occasionally switch states, and often these switching events demonstrate a "grandmother effect". In these instances, all progeny ("granddaughters") derived from a single derepressed "grandmother" cell from two generations earlier will switch to a silenced state simultaneously, raising the possibility that an event linked to DNA replication in the grandmother was propagated to subsequent generations (Pillus & Rine, 1989). When mother and daughter cells switch epigenetic states simultaneously upon cell division, the daughter cell usually silences more rapidly than the mother. In instances where the switching event does not occur in both cells of a mother/daughter pair, the daughter cell is more likely to switch to a silenced state than the mother (Osborne et al., 2011). The mechanism behind this difference is unknown but could be linked to asymmetric inheritance of soluble proteins (e.g. Sirs) or the sister chromatids during cell division. Alternatively, asymmetric expression of proteins that inhibit (in the mother cell) or promote (in the daughter cell) silent chromatin formation could also contribute to this process (Osborne et al., 2011).

### 5.2 Switching histone modification patterns

Switching of histone modification patterns to promote different epigenetic states can occur several ways. Histone modifications can be actively removed by enzymes such as deacetylases or demethylases, exchange of histones via chromatin remodeling (Lu et al., 2009), or proteolytic cleavage of histone tails (Jenuwein & Allis, 2001, Bannister et al., 2002). Histone modifications at chromosomal loci can also be passively removed by dilution upon DNA replication. Alternatively, certain histone modifications can remain present at loci in cells with different epigenetic states, but the function of those modifications may be altered

through adding or removing modifications at neighboring residues (Bannister et al., 2002, Fischle et al., 2003a). Together, these changes can affect chromatin structure by altering interactions between histones and DNA or chromatin-associated proteins that bind nucleosomes, ultimately leading to switching of epigenetic states.

Several examples of histone modifications influencing the probability of establishing silencing have been observed in studies of yeast mating-type silencing. In the absence of the H4 K16 acetyltransferase Sas2p, the probability of establishing silencing in a given cell cycle decreases, whereas loss of the H3 K4 and H3 K79 methyltransferases encoded by *SET1* and *DOT1* increases the probability of establishment (Osborne et al., 2009). In the case of *DOT1*, H3 K79 methylation status has been demonstrated to influence this switching event (Osborne et al., 2011). The rate and mechanism of removal of different histone modifications can also vary during establishment. At *HMR* and telomeres, acetyl groups from lysine residues on the N terminal tail of H3 are rapidly removed by Sir2p, a structural component of silent chromatin. The rapid rate of loss of H3 K4me2 during silent chromatin formation also implicates an active process mediating removal of this mark. In contrast, H3 K79me2 appears to be removed passively via dilution during DNA replication over 3 to 5 cell generations (Katan-Khaykovich & Struhl, 2005).

### 5.3 Switching DNA methylation patterns

In mammals, high 5-mC levels are commonly found in transcriptionally silent loci and promote the formation and maintenance of heterochromatin. In contrast, loss of 5-mC can disrupt binding sites for methyl binding proteins, destabilize heterochromatin, and lead to gene reactivation. Like histone modifications, DNA methylation can be removed by both passive and active processes. Passive demethylation of the maternal genome during pre-implantation development is thought to reflect loss of 5-mC via dilution upon DNA replication (Howell et al., 2001, Rougier et al., 1998). 5-mC can also be actively removed by DNA glycosylases in plants and base excision repair and nucleotide excision repair have been implicated in DNA demethylation in multiple organisms (Chen & Riggs, 2011, Zhu, 2009).

Passive and active removal of DNA methylation and gene reactivation may also involve 5-hydroxymethylcytosine, 5-hmC, a modification recently identified in mammalian DNA (Tahiliani et al., 2009, Kriaucionis & Heintz, 2009, Munzel et al., 2010). 5-hmC is created through the conversion of 5-mC to 5-hmC by the oxygenase TET1 (Tahiliani et al., 2009). TET1 is a homolog of the trypanosome proteins JBP1 and JBP2 that oxidize 5-methyl thymine (Tahiliani et al., 2009). Two other mammalian TET family members, TET2 and TET3, are also predicted to convert 5-mC to 5-hmC. Passive demethylation via dilution during DNA replication may occur upon conversion of 5-mC to 5-hmC if Dnmt1 cannot use 5-hmC-containing DNA as a substrate, although evidence for this remains to be established. Alternatively, 5-hmC may function in replication-independent, or "active", demethylation through spontaneous conversion to cytosine or processing via base-excision repair, analogous to that catalyzed by DEMETER in plants during the removal of 5-mC (Gehring & Henikoff, 2007). Rapid demethylation of the paternal genome following fertilization involves conversion of 5-mC to 5-hmC, but how 5-hmC is later processed is unclear (Iqbal et al., 2011, Santos et al., 2002).

Currently, the genomic locations of 5-hmC are largely unknown, although 5-hmC is likely less prevalent than 5-mC in most cell types. 5-hmC is indistinguishable from 5-mC in bisulfite sequencing studies used to map genomic methylated CpGs sites as bisulfite

treatment converts 5-hmC into cytosine 5-methylenesulfonate, which is resistant to deamination and conversion to uracil (Huang et al., 2010). Thus, many chromosomal sites previously thought to contain 5-mC may actually contain 5-hmC. Identification of genomic regions enriched for 5-hmC will require alternative approaches such as chromatin immunoprecipitation coupled to next generation sequencing.

The impact of 5-hmC on gene expression and chromatin structure is also poorly understood. However, TET1 has been mapped to CpG rich sequences within hypomethylated promoters (Wu et al., 2011). TET1 association at these loci correlates with histone modification patterns signifying either repressed (e.g. H3 K27me3) or active (e.g. H3 K4me3 and H3 K36me3) chromatin, implying a role for TET1 in regulating multiple epigenetic states (Wu et al., 2011). Conversion of 5-mC to 5-hmC by TET oxygenases disrupts interactions between methyl binding domain (MBD) proteins and DNA, and thus, likely epigenetic states. MBD1, MBD2b, MBD2b/MBD3L and MBD4 bind to 5-mC DNA but not 5-hmC DNA *in vitro* (Jin et al., 2010). The Rett's syndrome protein MeCP2 is also predicted not to bind to 5-hmC DNA (Valinluck et al., 2004). Together, these observations imply TET oxygenases may function in maintaining hypomethylated chromosomal regions, but whether TET proteins can be directly coupled to DNA replication is not yet known.

## 6. Summary

Propagating epigenetic states of gene expression involves the concerted effort of a complex network of histone and DNA-modifying enzymes, histone chaperones and replication proteins to ensure efficient duplication and inheritance of chromatin modification patterns and chromatin associated proteins. In recent years, numerous chromatin modifications and corresponding enzymes have been identified and their links to epigenetic processes have been demonstrated. However, understanding their roles in regulating epigenetic processes upon DNA replication remains to be developed. Many questions still must be addressed to clarify how chromatin assembly is regulated by chromatin modifying enzymes, how histones and structural components of heterochromatin are removed from in front of the replication fork, and then rebuilt behind the fork during replication. Future progress will reveal how histone and DNA modification patterns contribute to locus-specific epigenetic states across the genome.

## 7. Acknowledgements

Support for this project from the National Science Foundation (ALK), the Purdue University Center for Cancer Research Small Grants Program (ALK) and the Purdue Department of Biochemistry is gratefully acknowledged. We apologize to authors whose work we could not cite due to space limitations.

## 8. References

- Achour, M., Fuhrmann, G., Alhosin, M., Ronde, P., Chataigneau, T., Mousli, M., Schini-Kerth, V. B. & Bronner, C. 2009. UHRF1 recruits the histone acetyltransferase Tip60 and controls its expression and activity. *Biochem Biophys Res Commun*, 390, 523-8.

- Adkins, M. W., Carson, J. J., English, C. M., Ramey, C. J. & Tyler, J. K. 2007. The histone chaperone anti-silencing function 1 stimulates the acetylation of newly synthesized histone H3 in S-phase. *J Biol Chem*, 282, 1334-40.
- Adkins, M. W. & Tyler, J. K. 2004. The histone chaperone Asf1p mediates global chromatin disassembly in vivo. *J Biol Chem*, 279, 52069-74.
- Ahmad, K. & Henikoff, S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell*, 9, 1191-200.
- Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. & Shirakawa, M. 2008. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature*, 455, 818-21.
- Avvakumov, G. V., Walker, J. R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C. H. & Dhe-Paganon, S. 2008. Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature*, 455, 822-5.
- Axelrod, A. & Rine, J. 1991. A role for CDC7 in repression of transcription at the silent mating-type locus HMR in *S. cerevisiae*. *Mol Cell Biol*, 11, 1080-91.
- Bannister, A. J., Schneider, R. & Kouzarides, T. 2002. Histone methylation: dynamic or static? *Cell*, 109, 801-6.
- Belotserkovskaya, R., Oh, S., Bondarenko, V. A., Orphanides, G., Studitsky, V. M. & Reinberg, D. 2003. FACT facilitates transcription-dependent nucleosome alteration. *Science*, 301, 1090-3.
- Berndsen, C. E., Tsubota, T., Lindner, S. E., Lee, S., Holton, J. M., Kaufman, P. D., Keck, J. L. & Denu, J. M. 2008. Molecular functions of the histone acetyltransferase chaperone complex Rtt109-Vps75. *Nat Struct Mol Biol*, 15, 948-56.
- Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T. & Schreiber, S. L. 2002. Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci U S A*, 99, 8695-700.
- Bostick, M., Kim, J. K., Esteve, P. O., Clark, A., Pradhan, S. & Jacobsen, S. E. 2007. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science*, 317, 1760-4.
- Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B. & Bestor, T. H. 2001. Dnmt3L and the establishment of maternal genomic imprints. *Science*, 294, 2536-9.
- Braunstein, M., Sobel, R. E., Allis, C. D., Turner, B. M. & Broach, J. R. 1996. Efficient transcriptional silencing in *S. cerevisiae* requires a heterochromatin histone acetylation pattern. *Mol Cell Biol*, 16, 4349-56.
- Burgess, R. J., Zhou, H., Han, J. & Zhang, Z. 2010. A role for Gcn5 in replication-coupled nucleosome assembly. *Mol Cell*, 37, 469-80.
- Chedin, F., Lieber, M. R. & Hsieh, C. L. 2002. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc Natl Acad Sci U S A*, 99, 16916-21.
- Chen, T., Ueda, Y., Dodge, J. E., Wang, Z. & Li, E. 2003. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol*, 23, 5594-605.
- Chen, Z. X., Mann, J. R., Hsieh, C. L., Riggs, A. D. & Chedin, F. 2005. Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. *J Cell Biochem*, 95, 902-17.

- Chen, Z. X. & Riggs, A. D. 2011. DNA methylation and demethylation in mammals. *J Biol Chem*.
- Chuang, L. S., Ian, H. I., Koh, T. W., Ng, H. H., Xu, G. & Li, B. F. 1997. Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science*, 277, 1996-2000.
- Citterio, E., Papait, R., Nicassio, F., Vecchi, M., Gomiero, P., Mantovani, R., Di Fiore, P. P. & Bonapace, I. M. 2004. Np95 is a histone-binding protein endowed with ubiquitin ligase activity. *Mol Cell Biol*, 24, 2526-35.
- Collins, N., Poot, R. A., Kukimoto, I., Garcia-Jimenez, C., Dellaire, G. & Varga-Weisz, P. D. 2002. An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. *Nat Genet*, 32, 627-32.
- Das, C., Lucia, M. S., Hansen, K. C. & Tyler, J. K. 2009. CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature*, 459, 113-7.
- Datta, J., Majumder, S., Bai, S., Ghoshal, K., Kutay, H., Smith, D. S., Crabb, J. W. & Jacob, S. T. 2005. Physical and functional interaction of DNA methyltransferase 3A with Mbd3 and Brg1 in mouse lymphosarcoma cells. *Cancer Res*, 65, 10891-900.
- Dillin, A. & Rine, J. 1997. Separable functions of ORC5 in replication initiation and silencing in *S. cerevisiae*. *Genetics*, 147, 1053-62.
- Driscoll, R., Hudson, A. & Jackson, S. P. 2007. Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. *Science*, 315, 649-52.
- Easwaran, H. P., Schermelleh, L., Leonhardt, H. & Cardoso, M. C. 2004. Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO Rep*, 5, 1181-6.
- Ehrenhofer-Murray, A. E., Kamakaka, R. T. & Rine, J. 1999. A role for the replication proteins PCNA, RF-C, polymerase epsilon and Cdc45 in transcriptional silencing in *S. cerevisiae*. *Genetics*, 153, 1171-82.
- English, C. M., Adkins, M. W., Carson, J. J., Churchill, M. E. & Tyler, J. K. 2006. Structural basis for the histone chaperone activity of Asf1. *Cell*, 127, 495-508.
- Epsztejn-Litman, S., Feldman, N., Abu-Remaileh, M., Shufaro, Y., Gerson, A., Ueda, J., Deplus, R., Fuks, F., Shinkai, Y., Cedar, H. & Bergman, Y. 2008. De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol*, 15, 1176-83.
- Fatemi, M., Hermann, A., Gowher, H. & Jeltsch, A. 2002. Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *Eur J Biochem*, 269, 4981-4.
- Fatemi, M., Hermann, A., Pradhan, S. & Jeltsch, A. 2001. The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *J Mol Biol*, 309, 1189-99.
- Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H. & Bergman, Y. 2006. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol*, 8, 188-94.
- Feltus, F. A., Lee, E. K., Costello, J. F., Plass, C. & Vertino, P. M. 2003. Predicting aberrant CpG island methylation. *Proc Natl Acad Sci U S A*, 100, 12253-8.
- Fillingham, J., Recht, J., Silva, A. C., Suter, B., Emili, A., Stagljar, I., Krogan, N. J., Allis, C. D., Keogh, M. C. & Greenblatt, J. F. 2008. Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. *Mol Cell Biol*, 28, 4342-53.

- Fischle, W., Wang, Y. & Allis, C. D. 2003a. Binary switches and modification cassettes in histone biology and beyond. *Nature*, 425, 475-9.
- Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D. & Khorasanizadeh, S. 2003b. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev*, 17, 1870-81.
- Foss, M., McNally, F. J., Laurenson, P. & Rine, J. 1993. Origin recognition complex (ORC) in transcriptional silencing and DNA replication in *S. cerevisiae*. *Science*, 262, 1838-44.
- Francis, N. J. 2009. Mechanisms of epigenetic inheritance: copying of polycomb repressed chromatin. *Cell Cycle*, 8, 3513-8.
- Francis, N. J., Follmer, N. E., Simon, M. D., Aghia, G. & Butler, J. D. 2009. Polycomb proteins remain bound to chromatin and DNA during DNA replication in vitro. *Cell*, 137, 110-22.
- Franco, A. A., Lam, W. M., Burgers, P. M. & Kaufman, P. D. 2005. Histone deposition protein Asf1 maintains DNA replisome integrity and interacts with replication factor C. *Genes Dev*, 19, 1365-75.
- Fuks, F., Burgers, W. A., Brehm, A., Hughes-Davies, L. & Kouzarides, T. 2000. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat Genet*, 24, 88-91.
- Fuks, F., Burgers, W. A., Godin, N., Kasai, M. & Kouzarides, T. 2001. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J*, 20, 2536-44.
- Fuks, F., Hurd, P. J., Deplus, R. & Kouzarides, T. 2003. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res*, 31, 2305-12.
- Gaillard, P. H., Martini, E. M., Kaufman, P. D., Stillman, B., Moustacchi, E. & Almouzni, G. 1996. Chromatin assembly coupled to DNA repair: a new role for chromatin assembly factor I. *Cell*, 86, 887-96.
- Galvani, A., Courbeyrette, R., Agez, M., Ochsenbein, F., Mann, C. & Thuret, J. Y. 2008. In vivo study of the nucleosome assembly functions of ASF1 histone chaperones in human cells. *Mol Cell Biol*, 28, 3672-85.
- Gambus, A., Jones, R. C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R. D. & Labib, K. 2006. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol*, 8, 358-66.
- Gardner, K. A., Rine, J. & Fox, C. A. 1999. A region of the Sir1 protein dedicated to recognition of a silencer and required for interaction with the Orc1 protein in *S. cerevisiae*. *Genetics*, 151, 31-44.
- Gasser, R., Koller, T. & Sogo, J. M. 1996. The stability of nucleosomes at the replication fork. *J Mol Biol*, 258, 224-39.
- Gehring, M. & Henikoff, S. 2007. DNA methylation dynamics in plant genomes. *Biochim Biophys Acta*, 1769, 276-86.
- Gowher, H. & Jeltsch, A. 2001. Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *J Mol Biol*, 309, 1201-8.

- Gowher, H., Liebert, K., Hermann, A., Xu, G. & Jeltsch, A. 2005. Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *J Biol Chem*, 280, 13341-8.
- Goyal, R., Reinhardt, R. & Jeltsch, A. 2006. Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic Acids Res*, 34, 1182-8.
- Green, E. M., Antczak, A. J., Bailey, A. O., Franco, A. A., Wu, K. J., Yates, J. R., 3rd & Kaufman, P. D. 2005. Replication-independent histone deposition by the HIR complex and Asf1. *Curr Biol*, 15, 2044-9.
- Groth, A., Corpet, A., Cook, A. J., Roche, D., Bartek, J., Lukas, J. & Almouzni, G. 2007a. Regulation of replication fork progression through histone supply and demand. *Science*, 318, 1928-31.
- Groth, A., Ray-Gallet, D., Quivy, J. P., Lukas, J., Bartek, J. & Almouzni, G. 2005. Human Asf1 regulates the flow of S phase histones during replicational stress. *Mol Cell*, 17, 301-11.
- Groth, A., Rocha, W., Verreault, A. & Almouzni, G. 2007b. Chromatin challenges during DNA replication and repair. *Cell*, 128, 721-33.
- Gruss, C., Wu, J., Koller, T. & Sogo, J. M. 1993. Disruption of the nucleosomes at the replication fork. *EMBO J*, 12, 4533-45.
- Han, J., Zhou, H., Horazdovsky, B., Zhang, K., Xu, R. M. & Zhang, Z. 2007. Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. *Science*, 315, 653-5.
- Hashimoto, H., Horton, J. R., Zhang, X., Bostick, M., Jacobsen, S. E. & Cheng, X. 2008. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature*, 455, 826-9.
- Hata, K., Okano, M., Lei, H. & Li, E. 2002. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development*, 129, 1983-93.
- Henderson, D. S., Banga, S. S., Grigliatti, T. A. & Boyd, J. B. 1994. Mutagen sensitivity and suppression of position-effect variegation result from mutations in mus209, the Drosophila gene encoding PCNA. *EMBO J*, 13, 1450-9.
- Hertel, L., De Andrea, M., Bellomo, G., Santoro, P., Landolfo, S. & Gariglio, M. 1999. The HMG protein T160 colocalizes with DNA replication foci and is down-regulated during cell differentiation. *Exp Cell Res*, 250, 313-28.
- Hervouet, E., Laliere, L., Debien, E., Cheray, M., Geairon, A., Rogniaux, H., Loussouarn, D., Martin, S. A., Vallette, F. M. & Cartron, P. F. 2010. Disruption of Dnmt1/PCNA/UHRF1 interactions promotes tumorigenesis from human and mice glial cells. *PLoS One*, 5, e11333.
- Hodges, E., Smith, A. D., Kendall, J., Xuan, Z., Ravi, K., Rooks, M., Zhang, M. Q., Ye, K., Bhattacharjee, A., Brizuela, L., McCombie, W. R., Wigler, M., Hannon, G. J. & Hicks, J. B. 2009. High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome Res*, 19, 1593-605.
- Hoek, M. & Stillman, B. 2003. Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. *Proc Natl Acad Sci U S A*, 100, 12183-8.
- Hoppe, G. J., Tanny, J. C., Rudner, A. D., Gerber, S. A., Danaie, S., Gygi, S. P. & Moazed, D. 2002. Steps in assembly of silent chromatin in yeast: Sir3-independent binding of a

- Sir2/Sir4 complex to silencers and role for Sir2-dependent deacetylation. *Mol Cell Biol*, 22, 4167-80.
- Howell, C. Y., Bestor, T. H., Ding, F., Latham, K. E., Mertineit, C., Trasler, J. M. & Chaillet, J. R. 2001. Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell*, 104, 829-38.
- Huang, S., Zhou, H., Katzmann, D., Hochstrasser, M., Atanasova, E. & Zhang, Z. 2005. Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. *Proc Natl Acad Sci U S A*, 102, 13410-5.
- Huang, S., Zhou, H., Tarara, J. & Zhang, Z. 2007. A novel role for histone chaperones CAF-1 and Rtt106p in heterochromatin silencing. *EMBO J*, 26, 2274-83.
- Huang, Y., Pastor, W. A., Shen, Y., Tahiliani, M., Liu, D. R. & Rao, A. 2010. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One*, 5, e8888.
- Huen, M. S., Sy, S. M., van Deursen, J. M. & Chen, J. 2008. Direct interaction between SET8 and proliferating cell nuclear antigen couples H4-K20 methylation with DNA replication. *J Biol Chem*, 283, 11073-7.
- Hyland, E. M., Cosgrove, M. S., Molina, H., Wang, D., Pandey, A., Cottee, R. J. & Boeke, J. D. 2005. Insights into the role of histone H3 and histone H4 core modifiable residues in *S. cerevisiae*. *Mol Cell Biol*, 25, 10060-70.
- Iida, T., Suetake, I., Tajima, S., Morioka, H., Ohta, S., Obuse, C. & Tsurimoto, T. 2002. PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. *Genes Cells*, 7, 997-1007.
- Iqbal, K., Jin, S. G., Pfeifer, G. P. & Szabo, P. E. 2011. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci U S A*, 108, 3642-7.
- Jackson, V. 1990. In vivo studies on the dynamics of histone-DNA interaction: evidence for nucleosome dissolution during replication and transcription and a low level of dissolution independent of both. *Biochemistry*, 29, 719-31.
- Jenuwein, T. & Allis, C. D. 2001. Translating the histone code. *Science*, 293, 1074-80.
- Jeong, S., Liang, G., Sharma, S., Lin, J. C., Choi, S. H., Han, H., Yoo, C. B., Egger, G., Yang, A. S. & Jones, P. A. 2009. Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol Cell Biol*, 29, 5366-76.
- Jin, S. G., Kadam, S. & Pfeifer, G. P. 2010. Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res*, 38, e125.
- Jones, P. A. & Liang, G. 2009. Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet*, 10, 805-11.
- Kamakaka, R. T., Bulger, M., Kaufman, P. D., Stillman, B. & Kadonaga, J. T. 1996. Postreplicative chromatin assembly by *Drosophila* and human chromatin assembly factor 1. *Mol Cell Biol*, 16, 810-7.
- Karagianni, P., Amazit, L., Qin, J. & Wong, J. 2008. ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. *Mol Cell Biol*, 28, 705-17.
- Katan-Khaykovich, Y. & Struhl, K. 2005. Heterochromatin formation involves changes in histone modifications over multiple cell generations. *EMBO J*, 24, 2138-49.

- Kaufman, P. D., Cohen, J. L. & Osley, M. A. 1998. Hir proteins are required for position-dependent gene silencing in *S. cerevisiae* in the absence of chromatin assembly factor I. *Mol Cell Biol*, 18, 4793-806.
- Kaufman, P. D., Kobayashi, R., Kessler, N. & Stillman, B. 1995. The p150 and p60 subunits of chromatin assembly factor I: a molecular link between newly synthesized histones and DNA replication. *Cell*, 81, 1105-14.
- Kaufman, P. D., Kobayashi, R. & Stillman, B. 1997. Ultraviolet radiation sensitivity and reduction of telomeric silencing in *S. cerevisiae* cells lacking chromatin assembly factor-I. *Genes Dev*, 11, 345-57.
- Kelly, T. J., Qin, S., Gottschling, D. E. & Parthun, M. R. 2000. Type B histone acetyltransferase Hat1p participates in telomeric silencing. *Mol Cell Biol*, 20, 7051-8.
- Kimura, A., Umehara, T. & Horikoshi, M. 2002. Chromosomal gradient of histone acetylation established by Sas2p and Sir2p functions as a shield against gene silencing. *Nat Genet*, 32, 370-7.
- Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X. S. & Ahringer, J. 2009. Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet*, 41, 376-81.
- Korber, P., Barbaric, S., Luckenbach, T., Schmid, A., Schermer, U. J., Blaschke, D. & Horz, W. 2006. The histone chaperone Asf1 increases the rate of histone eviction at the yeast PHO5 and PHO8 promoters. *J Biol Chem*, 281, 5539-45.
- Kornberg, R. D. 1977. Structure of chromatin. *Annu Rev Biochem*, 46, 931-54.
- Krawitz, D. C., Kama, T. & Kaufman, P. D. 2002. Chromatin assembly factor I mutants defective for PCNA binding require Asf1/Hir proteins for silencing. *Mol Cell Biol*, 22, 614-25.
- Kriaucionis, S. & Heintz, N. 2009. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*, 324, 929-30.
- Kuo, M. H., Brownell, J. E., Sobel, R. E., Ranalli, T. A., Cook, R. G., Edmondson, D. G., Roth, S. Y. & Allis, C. D. 1996. Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature*, 383, 269-72.
- Laird, C. D., Pleasant, N. D., Clark, A. D., Sneed, J. L., Hassan, K. M., Manley, N. C., Vary, J. C., Jr., Morgan, T., Hansen, R. S. & Stoger, R. 2004. Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc Natl Acad Sci U S A*, 101, 204-9.
- Lambert, J. P., Fillingham, J., Siahbazi, M., Greenblatt, J., Baetz, K. & Figeys, D. 2010. Defining the budding yeast chromatin-associated interactome. *Mol Syst Biol*, 6, 448.
- Le, S., Davis, C., Konopka, J. B. & Sternglanz, R. 1997. Two new S-phase-specific genes from *S. cerevisiae*. *Yeast*, 13, 1029-42.
- Leonhardt, H., Page, A. W., Weier, H. U. & Bestor, T. H. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell*, 71, 865-73.
- Li, H., Rauch, T., Chen, Z. X., Szabo, P. E., Riggs, A. D. & Pfeifer, G. P. 2006. The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. *J Biol Chem*, 281, 19489-500.

- Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A. & Zhang, Z. 2008. Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell*, 134, 244-55.
- Liang, G., Chan, M. F., Tomigahara, Y., Tsai, Y. C., Gonzales, F. A., Li, E., Laird, P. W. & Jones, P. A. 2002. Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol Cell Biol*, 22, 480-91.
- Loo, S., Fox, C. A., Rine, J., Kobayashi, R., Stillman, B. & Bell, S. 1995. The origin recognition complex in silencing, cell cycle progression, and DNA replication. *Mol Biol Cell*, 6, 741-56.
- Loyola, A., Bonaldi, T., Roche, D., Imhof, A. & Almouzni, G. 2006. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol Cell*, 24, 309-16.
- Lu, P. Y., Levesque, N. & Kobor, M. S. 2009. NuA4 and SWR1-C: two chromatin-modifying complexes with overlapping functions and components. *Biochem Cell Biol*, 87, 799-815.
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389, 251-60.
- Luo, K., Vega-Palas, M. A. & Grunstein, M. 2002. Rap1-Sir4 binding independent of other Sir, yKu, or histone interactions initiates the assembly of telomeric heterochromatin in yeast. *Genes Dev*, 16, 1528-39.
- Masumoto, H., Hawke, D., Kobayashi, R. & Verreault, A. 2005. A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature*, 436, 294-8.
- Meijsing, S. H. & Ehrenhofer-Murray, A. E. 2001. The silencing complex SAS-I links histone acetylation to the assembly of repressed chromatin by CAF-I and Asf1 in *S. cerevisiae*. *Genes Dev*, 15, 3169-82.
- Meilinger, D., Fellingner, K., Bultmann, S., Rothbauer, U., Bonapace, I. M., Klinkert, W. E., Spada, F. & Leonhardt, H. 2009. Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. *EMBO Rep*, 10, 1259-64.
- Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B. E., Nusbaum, C., Jaffe, D. B., Gnirke, A., Jaenisch, R. & Lander, E. S. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, 454, 766-70.
- Mello, J. A., Sillje, H. H., Roche, D. M., Kirschner, D. B., Nigg, E. A. & Almouzni, G. 2002. Human Asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway. *EMBO Rep*, 3, 329-34.
- Miller, A., Chen, J., Takasuka, T. E., Jacobi, J. L., Kaufman, P. D., Irudayaraj, J. M. & Kirchmaier, A. L. 2010. Proliferating cell nuclear antigen (PCNA) is required for cell cycle-regulated silent chromatin on replicated and nonreplicated genes. *J Biol Chem*, 285, 35142-54.
- Miller, A., Yang, B., Foster, T. & Kirchmaier, A. L. 2008. Proliferating cell nuclear antigen and ASF1 modulate silent chromatin in *S. cerevisiae* via lysine 56 on histone H3. *Genetics*, 179, 793-809.

- Milutinovic, S., Zhuang, Q. & Szyf, M. 2002. Proliferating cell nuclear antigen associates with histone deacetylase activity, integrating DNA replication and chromatin modification. *J Biol Chem*, 277, 20974-8.
- Min, J., Zhang, Y. & Xu, R. M. 2003. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev*, 17, 1823-8.
- Moggs, J. G., Grandi, P., Quivy, J. P., Jonsson, Z. O., Hubscher, U., Becker, P. B. & Almouzni, G. 2000. A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing DNA damage. *Mol Cell Biol*, 20, 1206-18.
- Moldovan, G. L., Pfander, B. & Jentsch, S. 2007. PCNA, the maestro of the replication fork. *Cell*, 129, 665-79.
- Morrison, A. J., Highland, J., Krogan, N. J., Arbel-Eden, A., Greenblatt, J. F., Haber, J. E. & Shen, X. 2004. INO80 and gamma-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair. *Cell*, 119, 767-75.
- Munzel, M., Globisch, D., Bruckl, T., Wagner, M., Welzmueller, V., Michalakakis, S., Muller, M., Biel, M. & Carell, T. 2010. Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angew Chem Int Ed Engl*, 49, 5375-7.
- Okano, M., Bell, D. W., Haber, D. A. & Li, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99, 247-57.
- Okano, M., Xie, S. & Li, E. 1998. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res*, 26, 2536-40.
- Okuhara, K., Ohta, K., Seo, H., Shioda, M., Yamada, T., Tanaka, Y., Dohmae, N., Seyama, Y., Shibata, T. & Murofushi, H. 1999. A DNA unwinding factor involved in DNA replication in cell-free extracts of *Xenopus* eggs. *Curr Biol*, 9, 341-50.
- Ooi, S. K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S. P., Allis, C. D., Cheng, X. & Bestor, T. H. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*, 448, 714-7.
- Osada, S., Sutton, A., Muster, N., Brown, C. E., Yates, J. R., 3rd, Sternglanz, R. & Workman, J. L. 2001. The yeast SAS (something about silencing) protein complex contains a MYST-type putative acetyltransferase and functions with chromatin assembly factor ASF1. *Genes Dev*, 15, 3155-68.
- Osborne, E. A., Dudoit, S. & Rine, J. 2009. The establishment of gene silencing at single-cell resolution. *Nat Genet*, 41, 800-6.
- Osborne, E. A., Hiraoka, Y. & Rine, J. 2011. Symmetry, asymmetry, and kinetics of silencing establishment in *S. cerevisiae* revealed by single-cell optical assays. *Proc Natl Acad Sci U S A*, 108, 1209-16.
- Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M. & Shirakawa, M. 2009. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep*, 10, 1235-41.
- Papamichos-Chronakis, M. & Peterson, C. L. 2008. The Ino80 chromatin-remodeling enzyme regulates replisome function and stability. *Nat Struct Mol Biol*, 15, 338-45.
- Parthun, M. R. 2007. Hat1: the emerging cellular roles of a type B histone acetyltransferase. *Oncogene*, 26, 5319-28.
- Pillus, L. & Rine, J. 1989. Epigenetic inheritance of transcriptional states in *S. cerevisiae*. *Cell*, 59, 637-47.

- Poveda, A., Pamblanco, M., Tafrov, S., Tordera, V., Sternglanz, R. & Sendra, R. 2004. Hif1 is a component of yeast histone acetyltransferase B, a complex mainly localized in the nucleus. *J Biol Chem*, 279, 16033-43.
- Ptashne, M. 1992. *A Genetic Switch*, Cambridge, MA, Blackwell Sci.
- Ramsahoye, B. H., Biniszkiwicz, D., Lyko, F., Clark, V., Bird, A. P. & Jaenisch, R. 2000. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A*, 97, 5237-42.
- Recht, J., Tsubota, T., Tanny, J. C., Diaz, R. L., Berger, J. M., Zhang, X., Garcia, B. A., Shabanowitz, J., Burlingame, A. L., Hunt, D. F., Kaufman, P. D. & Allis, C. D. 2006. Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc Natl Acad Sci U S A*, 103, 6988-93.
- Reifsnyder, C., Lowell, J., Clarke, A. & Pillus, L. 1996. Yeast SAS silencing genes and human genes associated with AML and HIV-1 Tat interactions are homologous with acetyltransferases. *Nat Genet*, 14, 42-9.
- Riggs, A. D. & Xiong, Z. 2004. Methylation and epigenetic fidelity. *Proc Natl Acad Sci U S A*, 101, 4-5.
- Rougier, N., Bourc'his, D., Gomes, D. M., Niveleau, A., Plachot, M., Paldi, A. & Viegas-Pequignot, E. 1998. Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev*, 12, 2108-13.
- Rusche, L. N., Kirchmaier, A. L. & Rine, J. 2002. Ordered nucleation and spreading of silenced chromatin in *S. cerevisiae*. *Mol Biol Cell*, 13, 2207-22.
- Sanematsu, F., Takami, Y., Barman, H. K., Fukagawa, T., Ono, T., Shibahara, K. & Nakayama, T. 2006. Asf1 is required for viability and chromatin assembly during DNA replication in vertebrate cells. *J Biol Chem*, 281, 13817-27.
- Santos, F., Hendrich, B., Reik, W. & Dean, W. 2002. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol*, 241, 172-82.
- Sarraf, S. A. & Stancheva, I. 2004. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Mol Cell*, 15, 595-605.
- Sasaki, H. & Matsui, Y. 2008. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat Rev Genet*, 9, 129-40.
- Schermelleh, L., Haemmer, A., Spada, F., Rosing, N., Meilinger, D., Rothbauer, U., Cardoso, M. C. & Leonhardt, H. 2007. Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res*, 35, 4301-12.
- Schneider, J., Bajwa, P., Johnson, F. C., Bhaumik, S. R. & Shilatifard, A. 2006. Rtt109 is required for proper H3K56 acetylation: a chromatin mark associated with the elongating RNA polymerase II. *J Biol Chem*, 281, 37270-4.
- Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D. & Jenuwein, T. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev*, 18, 1251-62.
- Schulz, L. L. & Tyler, J. K. 2006. The histone chaperone ASF1 localizes to active DNA replication forks to mediate efficient DNA replication. *FASEB J*, 20, 488-90.
- Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T. A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., Tajima, S., Mitsuya, K., Okano, M. &

- Koseki, H. 2007. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature*, 450, 908-12.
- Sharp, J. A., Fouts, E. T., Krawitz, D. C. & Kaufman, P. D. 2001. Yeast histone deposition protein Asf1p requires Hir proteins and PCNA for heterochromatic silencing. *Curr Biol*, 11, 463-73.
- Shibahara, K. & Stillman, B. 1999. Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell*, 96, 575-85.
- Singer, M. S., Kahana, A., Wolf, A. J., Meisinger, L. L., Peterson, S. E., Goggin, C., Mahowald, M. & Gottschling, D. E. 1998. Identification of high-copy disruptors of telomeric silencing in *S. cerevisiae*. *Genetics*, 150, 613-32.
- Smallwood, A., Esteve, P. O., Pradhan, S. & Carey, M. 2007. Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes Dev*, 21, 1169-78.
- Smith, J. S., Caputo, E. & Boeke, J. D. 1999. A genetic screen for ribosomal DNA silencing defects identifies multiple DNA replication and chromatin-modulating factors. *Mol Cell Biol*, 19, 3184-97.
- Smith, S. & Stillman, B. 1989. Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell*, 58, 15-25.
- Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T. & Allis, C. D. 1995. Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc Natl Acad Sci U S A*, 92, 1237-41.
- Sogo, J. M., Stahl, H., Koller, T. & Knippers, R. 1986. Structure of replicating simian virus 40 minichromosomes. The replication fork, core histone segregation and terminal structures. *J Mol Biol*, 189, 189-204.
- Stuwe, T., Hothorn, M., Lejeune, E., Rybin, V., Bortfeld, M., Scheffzek, K. & Ladurner, A. G. 2008. The FACT Spt16 "peptidase" domain is a histone H3-H4 binding module. *Proc Natl Acad Sci U S A*, 105, 8884-9.
- Suka, N., Luo, K. & Grunstein, M. 2002. Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. *Nat Genet*, 32, 378-83.
- Suka, N., Suka, Y., Carmen, A. A., Wu, J. & Grunstein, M. 2001. Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. *Mol Cell*, 8, 473-9.
- Taddei, A., Roche, D., Sibarita, J. B., Turner, B. M. & Almouzni, G. 1999. Duplication and maintenance of heterochromatin domains. *J Cell Biol*, 147, 1153-66.
- Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell*, 116, 51-61.
- Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L. & Rao, A. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 324, 930-5.
- Tan, B. C., Chien, C. T., Hirose, S. & Lee, S. C. 2006. Functional cooperation between FACT and MCM helicase facilitates initiation of chromatin DNA replication. *EMBO J*, 25, 3975-85.
- Triolo, T. & Sternglanz, R. 1996. Role of interactions between the origin recognition complex and SIR1 in transcriptional silencing. *Nature*, 381, 251-3.

- Tsubota, T., Berndsen, C. E., Erkmann, J. A., Smith, C. L., Yang, L., Freitas, M. A., Denu, J. M. & Kaufman, P. D. 2007. Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes. *Mol Cell*, 25, 703-12.
- Tyler, J. K., Adams, C. R., Chen, S. R., Kobayashi, R., Kamakaka, R. T. & Kadonaga, J. T. 1999. The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature*, 402, 555-60.
- Tyler, J. K., Collins, K. A., Prasad-Sinha, J., Amiot, E., Bulger, M., Harte, P. J., Kobayashi, R. & Kadonaga, J. T. 2001. Interaction between the Drosophila CAF-1 and ASF1 chromatin assembly factors. *Mol Cell Biol*, 21, 6574-84.
- Valinluck, V., Tsai, H. H., Rogstad, D. K., Burdzy, A., Bird, A. & Sowers, L. C. 2004. Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res*, 32, 4100-8.
- van Attikum, H., Fritsch, O., Hohn, B. & Gasser, S. M. 2004. Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair. *Cell*, 119, 777-88.
- van Leeuwen, F., Gafken, P. R. & Gottschling, D. E. 2002. Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell*, 109, 745-56.
- VanDemark, A. P., Blanksma, M., Ferris, E., Heroux, A., Hill, C. P. & Formosa, T. 2006. The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition. *Mol Cell*, 22, 363-74.
- Verreault, A., Kaufman, P. D., Kobayashi, R. & Stillman, B. 1996. Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell*, 87, 95-104.
- Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M., Bollen, M., Esteller, M., Di Croce, L., de Launoit, Y. & Fuks, F. 2006. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature*, 439, 871-4.
- Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Paabo, S., Rebhan, M. & Schubeler, D. 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet*, 39, 457-66.
- Webster, K. E., O'Bryan, M. K., Fletcher, S., Crewther, P. E., Aapola, U., Craig, J., Harrison, D. K., Aung, H., Phutikanit, N., Lyle, R., Meachem, S. J., Antonarakis, S. E., de Kretser, D. M., Hedger, M. P., Peterson, P., Carroll, B. J., *et al.* 2005. Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. *Proc Natl Acad Sci U S A*, 102, 4068-73.
- Wittmeyer, J. & Formosa, T. 1997. The *S. cerevisiae* DNA polymerase alpha catalytic subunit interacts with Cdc68/Spt16 and with Pob3, a protein similar to an HMG1-like protein. *Mol Cell Biol*, 17, 4178-90.
- Wu, H., D'Alessio, A. C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Eve Sun, Y. & Zhang, Y. 2011. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature*.
- Xie, W., Song, C., Young, N. L., Sperling, A. S., Xu, F., Sridharan, R., Conway, A. E., Garcia, B. A., Plath, K., Clark, A. T. & Grunstein, M. 2009. Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. *Mol Cell*, 33, 417-27.

- Xu, E. Y., Zawadzki, K. A. & Broach, J. R. 2006. Single-cell observations reveal intermediate transcriptional silencing states. *Mol Cell*, 23, 219-29.
- Xu, F., Zhang, Q., Zhang, K., Xie, W. & Grunstein, M. 2007. Sir2 deacetylates histone H3 lysine 56 to regulate telomeric heterochromatin structure in yeast. *Mol Cell*, 27, 890-900.
- Xu, M., Long, C., Chen, X., Huang, C., Chen, S. & Zhu, B. 2010. Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. *Science*, 328, 94-8.
- Yamasu, K. & Senshu, T. 1990. Conservative segregation of tetrameric units of H3 and H4 histones during nucleosome replication. *J Biochem*, 107, 15-20.
- Ye, J., Ai, X., Eugeni, E. E., Zhang, L., Carpenter, L. R., Jelinek, M. A., Freitas, M. A. & Parthun, M. R. 2005. Histone H4 lysine 91 acetylation a core domain modification associated with chromatin assembly. *Mol Cell*, 18, 123-30.
- Ye, X., Franco, A. A., Santos, H., Nelson, D. M., Kaufman, P. D. & Adams, P. D. 2003. Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. *Mol Cell*, 11, 341-51.
- Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W. & Jeltsch, A. 2010. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res*, 38, 4246-53.
- Zhang, Y., Rohde, C., Tierling, S., Jurkowski, T. P., Bock, C., Santacruz, D., Ragozin, S., Reinhardt, R., Groth, M., Walter, J. & Jeltsch, A. 2009. DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS Genet*, 5, e1000438.
- Zhang, Z., Hayashi, M. K., Merkel, O., Stillman, B. & Xu, R. M. 2002. Structure and function of the BAH-containing domain of Orc1p in epigenetic silencing. *EMBO J*, 21, 4600-11.
- Zhang, Z., Shibahara, K. & Stillman, B. 2000. PCNA connects DNA replication to epigenetic inheritance in yeast. *Nature*, 408, 221-5.
- Zhu, J. K. 2009. Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet*, 43, 143-66.

# Epigenetic Modifications: Genetic Basis of Environmental Stress Response

Takeo Kubota, Kunio Miyake and Takae Hirasawa

*Department of Epigenetic Medicine, Faculty of Medicine, University of Yamanashi  
Japan*

## 1. Introduction

Genomic DNA is faithfully replicated and divided between two daughter cells in the course of each cell cycle. In order to maintain the inheritance of gene expression patterns, the cell must not only replicate the DNA, but also duplicate its chromatin structure (McNairn & Gilbert, 2003). Following replication, DNA is methylated and packaged into nucleosomes by the binding of histone octamers to form chromatin. DNA methyltransferases (DNMTs), the enzymes that transfer methyl (CH<sub>3</sub>) residue to CpG dinucleotides, are coordinated with DNA replication to maintain the DNA methylation pattern (Fig. 1). DNMTs recognize methylated CpG dinucleotides on the parent strand and methylate correlating CpG dinucleotides on the daughter strand (Bestor et al., 1996). This heritability of the DNA methylation pattern, as well as histone modification patterns, is mediated by epigenetic machinery.

Epigenetics was first used by Conrad Waddington in 1939 to describe “the causal interactions between genes and their products, which bring the phenotype into being” (Waddington, 1942). The current definition is “the study of heritable changes in gene expression that occur independent of changes in the primary DNA sequence” (Sharma et al., 2010). Waddington’s definition initially referred to the role of the epigenetics in embryonic development, in which cells develop distinct identities despite having the same genetic information; however, the definition of epigenetics has evolved over time as it is implicated in a wide variety of biological processes, including maintenance of the normal gene expression, carcinogenesis and genomic response to environmental stresses.

In this chapter, we take a look at the current understanding of epigenetic status in human cells, describe human diseases associated with congenital epigenetic errors, and also discuss how human diseases may be caused by acquired epigenetic errors as a result of environmental factors. We also discuss epigenetic therapies that take advantage of the fact that epigenetic changes are reversible.

## 2. Epigenetic status in human cells

The DNA methylation pattern is established during tissue development (Sakashita et al., 2001). Once the pattern is established in a cell, it is stably maintained through DNA replications at each cycle of cell division. Therefore, cells keep distinct identities while containing the same genetic information.

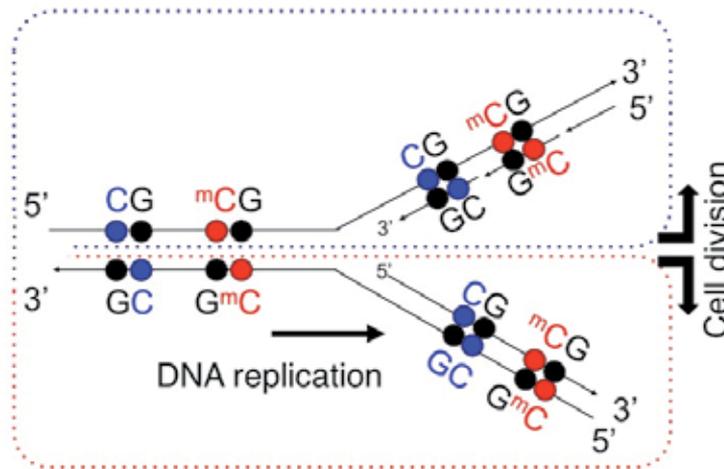


Fig. 1. Maintenance of the DNA methylation pattern during DNA replication and cell division

## 2.1 DNA methylation

During evolution, the CpG dinucleotide, the principal site of DNA methylation, has been selectively depleted through conversion of methylated cytosines to thymidines via a deamination process. Therefore, the human genome has only 10% of the expected frequency of CpGs, and 70 to 80% of these are heavily methylated. Small regions of DNA (1 to 2%), termed CpG islands, are not CpG-depleted. CpG islands are strongly protected from methylation and are associated with the transcription start sites in almost half of human genes. The genome organization facilitated by an epigenetic pattern is only present in higher order eukaryotes including mammals and humans. It is absent in *Drosophila*, *Caenorhabditis elegans*, and yeast (Baylin, 1997).

DNA methylation patterns closely correlate with patterns of gene expression (Fig. 2). Heavily methylated genomic regions are generally associated with chromatin organization that is inhibitory to transcription. In humans, such methylated genomic regions often contain highly repeated sequences; methylation may help guard against transcriptional

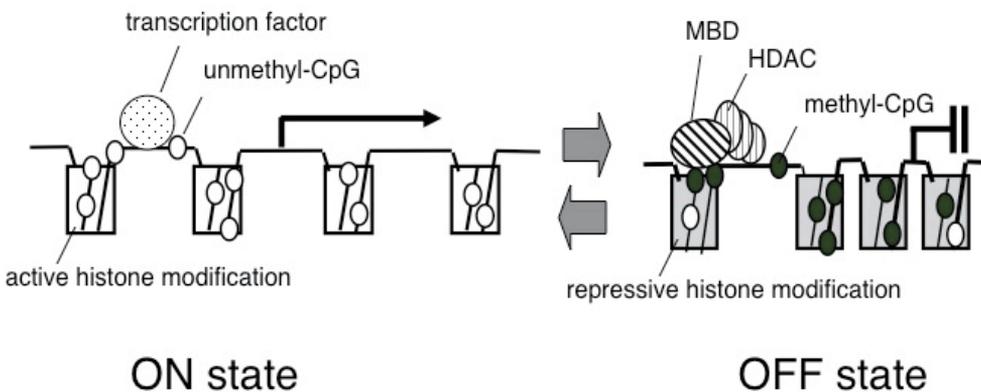


Fig. 2. Epigenetic gene regulation via DNA methylation and histone modifications

expression of parasitic sequences, which were introduced into the genome over evolution by transposable elements and DNA viruses (Bester et al., 1996). In contrast, the unmethylated CpG islands of genes are associated with chromatin containing highly transcribed DNA.

## 2.2 Histone modification

Histone proteins, which comprise the nucleosomes core, contain a globular C-terminal domain and an unstructured N-terminal tail (Lugar et al., 1997). The N-terminal tails of histones can undergo a variety of posttranslational covalent modifications including acetylation, methylation, phosphorylation, ubiquitylation, and sumoylation (Kouzarides, 2007). The complement of modifications is proposed to store the epigenetic memory inside a cell in the form of a “histone code” that determines the structure and activity of different chromatin regions (Jeniwein et al., 2001) (Fig. 2).

Unlike DNA methylation, histone modifications can lead to either activation or repression depending upon which residues are modified and the type of modifications present. For example, lysine acetylation correlates with transcriptional activation (Kouzarides, 2007; Hebbes et al., 1988), whereas lysine methylation leads to transcriptional activation or repression depending upon which residue is modified and the degree of methylation. For examples, trimethylation of lysine 4 on histone H3 (H3K4me3) is enriched at transcriptionally active gene promoters (Liang et al., 2004), and trimethylation of lysine 9 on histone H3 (H3K9me3) and trimethylation of lysine 27 on histone H3 (H3K27me3) is present at gene promoters that are transcriptionally repressed (Kouzarides, 2007). A vast array of active and repressive histone modifications have been identified, which constitute a complex gene regulatory network essential for the physiological activities of cells (Sharma et al., 2009).

## 2.3 Interplay of these epigenetic modifications

DNA methylation and histone modifications, not only perform individually, but also interact with each other at multiple levels to determine expression status, chromatin organization and cellular identity (Cedar et al., 2009). The two histone modifications (H3K9me3 and H3K27me3) that constitute the silencing mechanism in mammalian cells work in concert with DNA methylation. Furthermore, a histone methyltransferase (HMT) can direct DNA methylation to specific genomic targets by recruiting DNMTs to stably silence genes (Tachibana et al., 2008), and a histone demethylase (HDM) stabilizes DNMT1 protein to maintain DNA methylation (Wang et al., 2009).

DNMTs can in turn recruit methyl-binding domain proteins (MBDs) and histone deacetylases (HDACs) to achieve gene silencing and chromatin condensation (Jones et al., 1998; Nan et al., 1998) (Fig. 2). DNA methylation can also induce histone H3K9 methylation through an MBD (MeCP2), thereby establishing a repressive chromatin state (Fuks et al., 2003).

The interplay of these modifications creates an epigenetic landscape that regulates the way the mammalian genome manifests itself in different cell types, developmental stages and disease states. The distinct patterns of these modifications present in different cellular states serve as a guardian of cellular identity (Sharma et al., 2009).

## 2.4 Aberrant epigenetic modifications

A normal epigenetic landscape is known to be disturbed in specific disease conditions. For example, the cancer epigenome (the whole genomic epigenetic state) is marked by genome-

wide hypomethylation and site-specific CpG island promoter hypermethylation (Jones & Baylin, 2002). Global DNA hypomethylation plays a significant role in tumorigenesis and occurs at various genomic sequences including repetitive elements, retrotransposons, CpG poor promoters, introns and gene deserts (Rodriguez et al., 2006). Activation of the retrotransposons due to hypomethylation lead to increased genomic instability by promoting chromosomal rearrangements (Jones et al., 2002; Eden et al., 2003; Howard et al., 2008). Furthermore, methylation is known to stabilize various repetitive sequences. Thus, hypomethylation CAG trinucleotide repeats in a DNMT1-deficient mouse display increased repeat instability (Dion et al., 2008).

In contrast to hypomethylation, which increases genomic instability and activates proto-oncogenes, region-specific hypermethylation contributes to tumorigenesis by silencing tumor suppressor genes, such as *Rb*, *p16* and *BRCA1* (Sharma et al., 2009). These genes are involved in cellular processes integral to cancer development and progression, including DNA repair, cell cycle regulation, cell adhesion, apoptosis and angiogenesis. Silencing of DNA repair genes enables cells to accumulate further genetic lesions leading to the rapid progression of cancer. Hypermethylation at the binding site for CCCTC-binding factor (CTCF), a chromatin barrier by preventing the spread of heterochromatin structures, inhibits CTCF binding, and leads to instability of repetitive sequences, which is a causing-mechanism for various neurodegenerative diseases (López Caste et al., 2010). However, how genes are targeted for this aberrant DNA methylation is still unclear.

Both aberrant histone acetylation and histone methylation are found in cancer cells. These changes associated with overexpression of HDACs and dysregulation of HMTs (Halkidou et al., 2004; Song et al., 2005). Alterations in H3K9 and H3K27 methylation patterns are associated with aberrant gene silencing in cancers. It has recently demonstrated that aberrant nucleosome positioning is created by a co-repressor Nerd (nucleosome remodelling and deacetylase) complex that recruits PML-Para (an oncogenic transcription factor), polycomb repressor complex 2, DNMT3A, and MBD2, resulting in abnormal gene silencing in leukemia (Feng et al 2001; Morey et al., 2008).

Re-establishing normal histone acetylation patterns through treatment with HDAC inhibitors have been shown to have anti-tumorigenic effects, via reactivation of silenced tumor suppressor genes (Carew et al., 2008). Suberoylanilide hydroxamic acid (SAHA), which is an HDAC inhibitor, has now been approved for use in the clinic for treatment of lymphoma (Sharma et al., 2009).

## 2.5 Understanding of the global epigenetic landscape

The global epigenetic landscape that is correlated with important biological processes and disease state has not been comprehensively investigated for most cell types. However, recent advances in genomic technology, in particular high-throughput sequencing, have enabled genome-wide analysis of histone modifications and DNA methylation at nucleotide resolution (Beck, 2010). Large-scale epigenomic mapping studies have the potential to enhance three major areas of science: basic gene regulatory processes, cellular differentiation and reprogramming and the role of epigenetic regulation in disease (Satterlee, 2010). Understanding how the epigenomic state of human embryonic stem (ES) cells changes during the differentiation process is crucial for understanding normal development and establishing epigenomic maps of induced pluripotent stem (iPS) cells will be essential to enable regenerative medicine to reach its full potential for treating diseases (Deng et al., 2009; Ball et al., 2009; Doi et al., 2009). Genome-wide association studies have been

successful in identifying genetic variants associated with many different diseases. In the case of diseases that have a strong environmental component, epigenome-wide association studies based on the epigenomic maps of specific cell types that statistically correlate epigenetic variation with phenotypes, could be of great value (Kong et al., 2009).

To generate epigenomic maps for cell types, large-scale epigenomics effort have already been initiated. The NIH Roadmap Epigenomics Program (<http://www.roadmapepigenomics.org>) will permanently archive data in the GEO database ([Http://www.ncbi.nlm.nih.gov/epigenomics](http://www.ncbi.nlm.nih.gov/epigenomics)) at the US National Center for Biotechnology Information (NCBI), and the International Human Epigenome Consortium (IHEC) (<http://ihec-epigenomes.org>) aims to expand the number of cell types and generate additional 1,000 reference epigenomes (Beck, 2010) that are not being characterized in the NIH Roadmap Program.

### **3. Human diseases associated with congenital epigenetic errors**

Epigenetic gene control is an intrinsic mechanism for normal tissue development and abnormalities in the molecules associated with this mechanism are known to cause various congenital diseases.

#### **3.1 Genomic imprinting diseases**

Genomic imprinting is the epigenetic phenomenon initially discovered in human diseases. In an imprinted gene, out of the two parental alleles, one allele is active and the other allele is inactive due to epigenetic mechanism such as DNA methylation (Fig. 3C). Therefore, defect in the active allele of the imprinted gene results in the loss of expression. This has been found in neurodevelopmental diseases, Prader-Willi syndrome and Angelman syndrome (Kubota et al., 1997).

#### **3.2 X-chromosome inactivation disorders**

The X chromosome has a large number of genes, whereas the Y chromosome has relatively few genes. Thus, females (XX) have more genes than males (XY). To minimize this sex imbalance, one of the two X chromosomes in females is inactivated by epigenetic mechanism (Kubota et al., 1998). Improper X inactivation is thought to be an embryonic lethal condition. This hypothesis is supported by the recent findings in cloned animals produced by somatic nuclear transfer in which failure of X-chromosome inactivation was observed in the clones with embryonic abortion (Xue et al., 2002; Nolen et al., 2005). Even if one of the X chromosomes is extremely small due to a large terminal deletion, so that over dosage effect of X-linked genes is minimized, the affected female show a severe congenital neurodevelopmental delay (Kubota et al., 2002), indicating that proper gene suppression by epigenetic mechanism is essential for normal development (Fig. 3D).

#### **3.3 DNA methylation-associated protein diseases**

DNA methylation is a fundamental step in epigenetic gene control, and it is achieved by an addition of the methyl group (CH<sub>3</sub>) to CpG dinucleotides mediated by DNMTs. Defect in a DNMT (e.g., DNMT3B) can cause an ICF syndrome that is characterized by Immunodeficiency, centromere instability, facial abnormalities, and mild mental retardation (Fig. 3A) (Okano et al, 1999; Shirohzu et al., 2002; Kubota et al., 2004).

MBDs are also important molecules in the control of gene expression. Mutations in a MBD (e.g., MeCP2) can cause Rett syndrome, which is characterized by seizures, ataxic gait, language dysfunction and autistic behavior (Amir et al., 1999; Chunshu et al., 2006). Therefore, it has been thought that MeCP2 dysfunction leads to aberrant gene expression in the brain associated with neurological features of the disease. Recent studies have shown that MeCP2 controls a subset of neuronal genes (Chen et al., 2003; Martinowich et al., 2003; Horike et al., 2005; Itoh et al., 2007), suggesting that epigenetic dysregulation of the neuronal genes may cause neurological features of the disease (Fig. 3B).

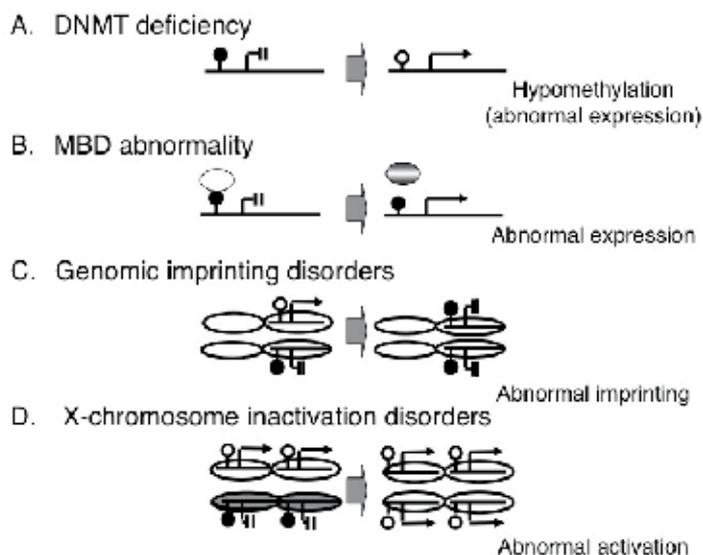


Fig. 3. Abnormal epigenetic patterns in human congenital diseases

#### 4. Proposed human diseases associated with acquired epigenetic errors caused by environmental factors

Health or diseases is shaped for all individuals by interactions between their genes and the environment. How the environment changes gene expression and how this can lead to a disease are being explored in a fruitful new approach to environmental health research. If these causal relations become clear, they offer new avenues for risk assessment for diseases (Edwards & Myers, 2007).

##### 4.1 Social background

The Ministry of Health, Welfare, and Labor in Japan has recently reported that the number of children with mild neurodevelopmental disorders, such as autism, is increasing by 10,000 cases per year (Basic report, 2005). Similar trends are found in other countries, including the US (Yeargin-Allsopp et al., 2003; Holoden, 2009; Fombonne, 2009), in which the increase is partly attributed to social factors, such as diagnostic substitution in which children formerly diagnosed with mental retardation or learning disabilities are now diagnosed with autism. However, the increase in cases cannot be fully attributed to such diagnostic substitutions,

and it is possible that biological changes in the brains of the children are also involved in this increase. Thanks to advances in genomic research, several genetic factors for autism have been identified. Mutations in genes encoding synaptic molecules have been identified in a subset of autistic children (Zoghbi, 2003; Persco & Bourgeron, 2006). However, the increase in autism cannot be solely attributed to genetic factors, because it is unlikely that mutation rates suddenly increased in recent years. Therefore, environmental factors are more likely to be involved in this increase. Epigenetic modifications represent one mechanism by which environmental factors can lead to health effects (Qiu, 2006).

#### 4.2 Acquired neurodevelopmental diseases

It is known that either a mutation, deletion or a duplication of a specific-neuronal gene causes a neurological disease. In other words, loss-of function, deficiency, or over-dosage can result in the same disease phenotype. For examples, Pelizaeus-Merzbacher disease, a severe child onset disorder, is caused by either a mutation, deletion or a duplication of the *PLP1* gene (Inoue et al., 2001), lissencephaly syndrome, a child-onset migration disorder, is caused by either a mutation, deletion or a duplication of *LIS1* (Reiner et al., 1993; Bi et al., 2009), Charcot-Marie-Tooth disease, an adult-onset neuromuscular disorder, is caused by either a mutation, deletion or a duplication *PMP22* (Roa & Lupski, 1993), and Parkinson disease is caused by either a mutation, deletion or a duplication the  *$\alpha$ -synuclein* gene (Obi et al., 2008). This suggests that the brain is sensitive to the dosage of gene products that requires a strict control system for gene expression. In fact, congenital diseases with defects in epigenetic gene regulation usually show neurological features and mental retardation.

It has recently been reported that short-term mental stress after birth can alter the epigenetic status in the brain, resulting in abnormal behaviour (Weaver et al., 2004). In rats, when the offspring is separated from the mother for a couple of weeks, DNA methylation at the *GR* (glucocorticoid receptor) gene is increased in the hippocampus in the brain, and this change suppresses gene expression. This study is now considered as an animal model for cruelty in childhood in human, because hypermethylation of the neuron-specific glucocorticoid receptor promoter, in combination with decreased levels of its expression, have been found in human postmortem hippocampus obtained from suicide victims with a history of childhood abuse (MacGowan et al., 2009), suggesting that adverse effects of early-life stress last life-time long on the DNA methylation programs (Margatroyd et al., 2009). It raises the question of whether neurodevelopmental problems may be the result of epigenetic dysregulation caused by environmental factors in the early life.

#### 4.3 Environmental factors in fetal period

Another social issue in Japan is that birth weight has decreased during the past 20 years. This trend is thought to be caused by the popularity of dieting among young women and obstetric physicians' recommendations to minimize pregnancy weight gain in order to reduce the risk of diabetes mellitus (Gluckman et al., 2007). Based on current epidemiological studies for famines in the Netherlands and China (St Clair et al., 2005; Painter et al., 2006), the generation with lower birth weight is expected to have increased risk for obesity and adult diseases in the future in Japan. This hypothesis is referred to as "Developmental Origin of Health and Diseases (DOHaD)" (Gillman et al., 2007; Silveira et al., 2007). Recent animal experiments suggest that the developmental basis of adult diseases is due to a change of DNA methylation status of the *PPAR $\alpha$*  gene, a thrifty gene, in the liver due to malnutrition

in the fetal period (Lillycrop et al., 2005; Lillycrop et al., 2008) (Fig. 4). Such DNA methylation alterations have been confirmed in individuals who suffered malnutrition during a period of famine (Tobi et al., 2009).

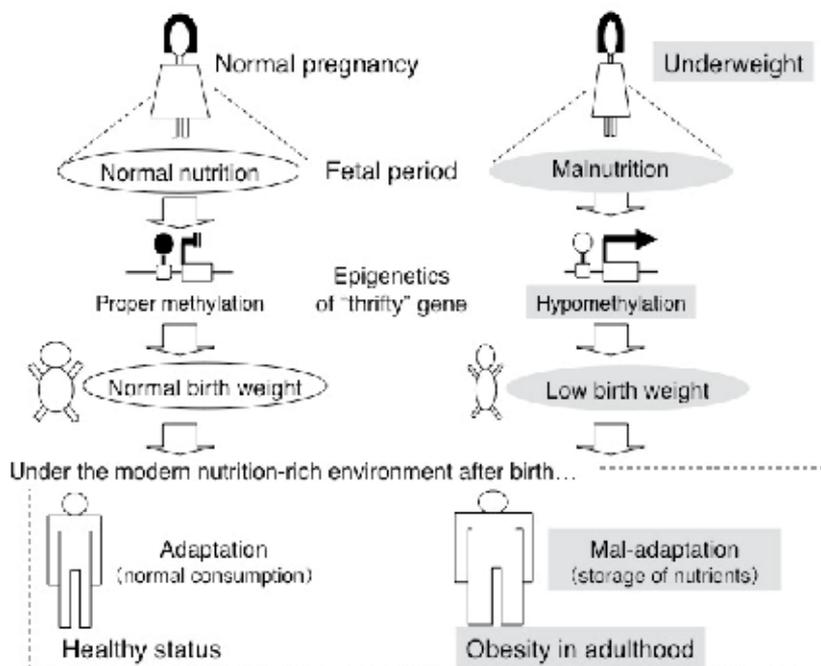


Fig. 4. Epigenetic mechanism proposed in the “Developmental Origin of Health and Diseases” hypothesis

#### 4.4 Drugs and chemicals affecting epigenetic status

Drug addiction is an example of mental diseases acquired via epigenetic change. Cocaine and alcohol alter the epigenetic state (chromatin structure) on a subset of neuronal genes, inducing a drug addiction state (Kumar et al., 2005; Pacual et al., 2009). Chemical compounds related to plastics also potentially affect the epigenetic status of genes in the brain (Yaoi et al., 2007).

Imipramine, a major antidepressant, in turn has recently been found to restore a depressive state by altering the epigenetics (histone modification) of the *Bdnf* gene in the hippocampus (Tsankova et al., 2006). Valproic acid (VPA), a histone deacetylase (HDAC) inhibitor, is another drug that alters the epigenetic state. VPA normalizes histone acetylation of genes in the hippocampus, which leads to suppression of seizure-induced cognitive impairment by blocking seizure-induced aberrant neurogenesis (Jessberger et al., 2007). These observations indicate that chemicals that alter epigenetic gene expression, such as HDAC inhibitors, may become candidates for the treatment of neurodevelopmental diseases (Renthal et al., 2008).

The findings above are mainly obtained from animal experiments, and there is little evidence from human studies. However, epigenetic differences increase with age in monozygotic twins (Fraga et al., 2005), suggesting that epigenetic status may be altering during aging by environmental factors in humans.

## 5. Epigenetic therapies

Because epigenetic modifications are a reversible mechanism unlike mutations (substitutions of nucleotides), correction of the epigenetic defect is potentially easier than correction of the mutations (Fig. 5).

### 5.1 DNA methylation donor

Folic acid is the methyl-donor for transfer to cytosine. Therefore, in order to maintain DNA methylation, sufficient intake of folic acid is essential. Folic acid deficiency during pregnancy is now increasing in Japan. This increases the risk of having babies with neural tube defects (Watanabe et al., 2008). As mentioned above, inappropriate supply of nutrients from mother to the fetus also increases the susceptibility of fetus to develop diabetes mellitus due to epigenetic changes (Park et al., 2008). However, supplementation of folic acid during pregnancy protects the fetus by enriching DNA methylation of the promoter regions of *PPAR $\alpha$*  and glucocorticoid receptor genes in the liver, leading to suppress gene expression (Buedge et al., 2009). These findings indicate that proper nutrient intakes may alter the phenotype of the offspring through epigenetic changes.

Since 1980s, folic acid has empirically been used for the treatment of autistic children and adults with mental diseases, and several studies have shown that folic acid is effective in a subset of patients (Rimland, 1998; James et al., 2004; Moretti et al., 2005). Although the precise mechanism is not known, it is also possible that folic acid administration may correct the DNA methylation status in genes.

### 5.2 Nutrition

A honey bee secretion known as royal jelly can cause phenotypic change in genetically identical female honeybees to induce the development of a fertile queen. This effect may be mediated by epigenetic changes. A recent study showed that royal jelly removes global DNA methylation, silencing the expression of *Dnmt3* during larval development (Kucharski et al., 2008). The phenotypic change from a worker bee to a queen is reproduced by using siRNA that inhibit *Dnmt3* (Kucharski et al., 2008). More recently, many kinds of nutrition have been shown to have epigenetic effects and epigenetic therapeutics have been approved by the US Food and drug Administration for treating specific cancers and seizure disorders (Mack, 2006; Sharma et al., 2010).

### 5.3 Gene-specific therapy

Folic acid is relatively safe, since it is a nutrient. However, its effect is global, and it is not specific to a certain gene. It may be better if epigenetic correction is made only to a specific gene that is associated with a disease state. This kind of therapy can be achieved using pyrrole-imidazole (PI) polyamides, small synthetic molecules that recognize and attach to the minor groove of DNA, thereby inhibiting gene transcription by blocking transcription factor binding in a DNA sequence specific manner (Matsuda et al., 2011). Furthermore, PI polyamide conjugated with SAHA, a HDAC inhibitor, can alter the histone modification in a gene-specific manner, resulting in up-regulation of the target gene (Ohtsuki et al., 2010).

### 5.4 Exercise and environmental enrichment

It has recently been discovered that DNA sequence is different in each neuron (Coufal et al., 2009), and that epigenetic change underlies the somatic change (Muotri et al., 2005). This

phenomenon is based on retrotransposition, in which a repetitive L1 sequence is inserted into various genomic regions when it is hypomethylated, potentially altering expression of adjacent genes. Retrotransposon insertion is activated by deficiency of MeCP2 (Muotri, et al., 2010). Interestingly, the retrotransposition is also activated by voluntary exercise (running) in mice (Muotri et al., 2009), suggesting that exercise may alter the DNA methylation status in neurons.

Studies using a Rett syndrome mouse model that lacks MeCP2 show that environmental enrichment (e.g., availability of stimulating toys) during early postnatal development produces effects on neural development and ameliorates the neurological phenotypes associated with Rett syndrome (Lonetti et al., 2010; Kerr et al., 2010). This suggests that DNA methylation status may be corrected by an appropriate environment, compensating for the insufficient MeCP2 function.

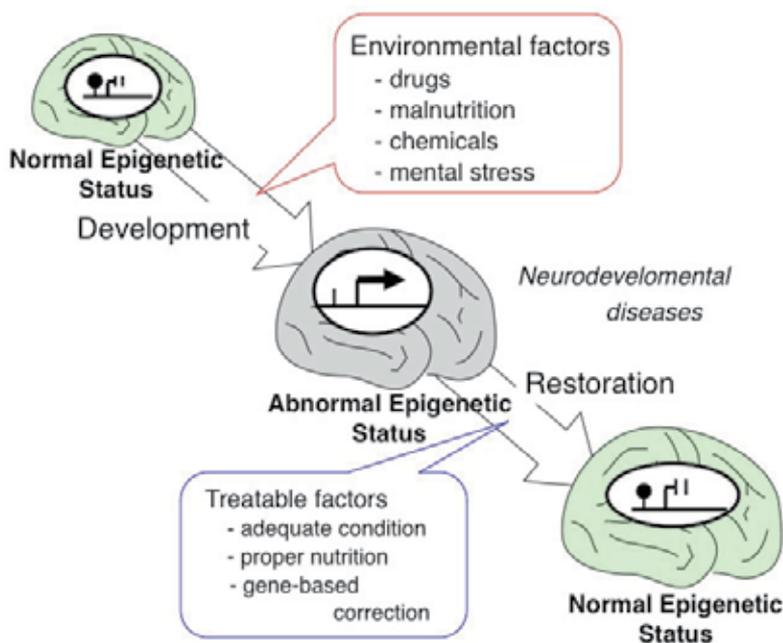


Fig. 5. Overview of epigenetic change and environmental factors in the brain

## 6. Conclusion

Epigenetics is a genetic code, not within the DNA but upon the DNA. Until recently, it has been believed that the epigenetic code is faithfully maintained at the step of DNA replication. However, various environmental factors potentially rewrite the epigenetic codes, which can lead to a disease condition. Moreover, a recent animal study has shown that mental stress not only rewrite the epigenetic code in the brain but also in the germline. Hence, the altered epigenetic code can be transmitted to the next generation, escaping the erasure of epigenetic marks that typically occurs during gametogenesis (Franklin et al., 2010). However, epigenetics is a reversible mechanism, and thus, epigenetic changes are treatable. Therefore, if the transgenerational inheritance of the epigenetic code is true in

humans and we transmit our own altered epigenetic code to our children, the code can be restored with the offer of appropriate environment and treatment. Although the number and kinds of environmental factors that can alter the epigenetic code are increasing, the precise mechanism is still largely unknown. Further understanding of how epigenetic modifications are changed during DNA replication is warranted in order to elucidate the genetic basis of environmental stress response.

## 7. References

- Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. (2001). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*, Vol.23, No.2, (October 2001), pp. 185-188, PMID: 10508514
- Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, Xie B, Daley GQ, Church GM. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature Biotechnology*, Vol.27, No.4 (April 2009), pp. 361-368, PMID: 19329998
- [Basic investigation report for handicapped children report by Ministry of Health, Welfare, and Labor, Japan.] 2005 (in Japanese).
- Baylin SB. (1997). Tying it all together: epigenetics, genetics, cell cycle, and cancer. *Science*, Vol.277, No.5334, (September 2010), pp. 1948-1949, PMID: 9333948
- Beck S. (2010). Taking the measure of the methylome. *Nature Biotechnology*, Vol.28, No.10 (October 2010), pp. 1026-1028, PMID: 20944589
- Bestor T, Laudano A, Mattaliano R, Ingram V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of Molecular Biology*, Vol.203, No.4, (October 1988), pp. 971-983, PMID: 3210246
- Bestor T, Laudano A, Mattaliano R, Ingram V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of Molecular Biology*, Vol.203, No.4, (October 1988), pp. 971-983, PMID: 3210246
- Bi W, Sapir T, Shchelochkov OA, Zhang F, Withers MA, Hunter JV, Levy T, Shinder V, Peiffer DA, Gunderson KL, Nezarati MM, Shotts VA, Amato SS, Savage SK, Harris DJ, Day-Salvatore DL, Horner M, Lu XY, Sahoo T, Yanagawa Y, Beaudet AL, Cheung SW, Martinez S, Lupski JR, Reiner O. (2009). Increased LIS1 expression affects human and mouse brain development. *Nature Genetics*, Vol.41, No.2, (February 2009), pp.168-177, PMID: 19136950
- Burdge GC, Lillycrop KA, Phillips ES, Slater-Jefferies JL, Jackson AA, Hanson MA. Folic acid supplementation during the juvenile-pubertal period in rats modifies the phenotype and epigenotype induced by prenatal nutrition. *Journal of Nutrition*, Vol.139, No.6 (June 2009), pp. 1054-1060, PMID: 19339705
- Carew JS, Giles FJ, Nawrocki ST. Histone deacetylase inhibitors: mechanisms of cell death and promise in combination cancer therapy. *Cancer Letters*, Vol.269, No.1 (September 2008), pp. 7-17, PMID: 18462867

- Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nature Review Genetics*, Vol.10, No.5 (May 2009), pp. 295-304, PMID: 19308066
- Chen WG, Chang Q, Lin Y, Meissner A, West AE, Griffith EC, Jaenisch R, Greenberg ME. (2003). Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science*, Vol.302, No.5646, (October 2003), pp. 885-889, PMID: 14593183
- Chunshu Y, Endoh K, Soutome M, Kawamura R, Kubota T. (2006). A patient with classic Rett syndrome with a novel mutation in MECP2 exon 1. *Clinical Genetics*, Vol.70, No.6, (December 2006), pp. 530-531, PMID: 17101000
- Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH. (2009). L1 retrotransposition in human neural progenitor cells. *Nature*, Vol.460, No.7259, (August 2009), pp. 1127-1131, PMID: 19657334
- Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, Antosiewicz-Bourget J, Egli D, Maherali N, Park IH, Yu J, Daley GQ, Eggan K, Hochedlinger K, Thomson J, Wang W, Gao Y, Zhang K. (2009). Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nature Biotechnology*, Vol.27, No.4 (April 2009), pp. 353-360, PMID: 19330000
- Dion V, Lin Y, Hubert L Jr, Waterland RA, Wilson JH. (2008). Dnmt1 deficiency promotes CAG repeat expansion in the mouse germline. *Human Molecular Genetics*, Vol.17, No.9, (May 2008), pp. 1306-1317, PMID: 18252747
- Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R, Herb B, Ladd-Acosta C, Rho J, Loewer S, Miller J, Schlaeger T, Daley GQ, Feinberg AP. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature Genetics*, Vol.41, No.12 (December 2009), pp. 1350-1353, PMID: 19881528
- Eden A, Gaudet F, Waghmare A, Jaenisch R. (2003). Chromosomal instability and tumors promoted by DNA hypomethylation. *Science*, Vol.300, No.5618, (April 2003), pp. 455, PMID: 12702868
- Edwards TM, Myers JP. (2007). Environmental exposures and gene regulation in disease etiology. *Environmental Health Perspectives*, Vol.115, No.9, (September 2007), pp. 1264-1270, PMID: 17805414
- Feng Q, Zhang Y. (2001). The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. *Genes and Development*, Vol.15, No.7 (April 2001), pp. 827-832, PMID: 11297506
- Fombonne E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, Vol.65, No.6, (January 2009), pp. 591-598, PMID: 19218885
- Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, Heine-Suñer D, Cigudosa JC, Urioste M, Benitez J, Boix-Chornet M, Sanchez-Aguilera A, Ling C, Carlsson E, Poulsen P, Vaag A, Stephan Z, Spector TD, Wu YZ, Plass C, Esteller M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy Science of the United States of America*, Vol.102, No.30, (July 2005), pp. 10604-10609, PMID: 1600993
- Franklin TB, Russig H, Weiss IC, Gräff J, Linder N, Michalon A, Vizi S, Mansuy IM. (2010). Epigenetic transmission of the impact of early stress across generations. *Biological Psychiatry*, Vol.68, No.5, (April 2010), pp. 408-415, PMID: 20673872

- Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *Journal of Biological Chemistry*, Vol.278, No.6, (February 2003), pp. 4035-4040, PMID: 12427740
- Gillman MW, Barker D, Bier D, Cagampang F, Challis J, Fall C, Godfrey K, Gluckman P, Hanson M, Kuh D, Nathanielsz P, Nestel P, Thornburg KL. (2007). Meeting report on the 3rd International Congress on Developmental Origins of Health and Disease (DOHaD). *Pediatric Research*, Vol.61, No.5 Pt 1, (May 2007), pp. 625-629, PMID: 17413866
- Gluckman PD, Seng CY, Fukuoka H, Beedle AS, Hanson MA. (2007). Low birthweight and subsequent obesity in Japan. *Lancet*, Vol.369, No.9567, (March 2007), pp. 1081-1082, PMID: 17398304
- Halkidou K, Gaughan L, Cook S, Leung HY, Neal DE, Robson CN. (2004). Upregulation and nuclear recruitment of HDAC1 in hormone refractory prostate cancer. *Prostate*, Vol.59, No.2 (May 2004), pp. 177-189, PMID: 15042618
- Hebbes TR, Thorne AW, Crane-Robinson C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO Journal*, Vol.7, No.5, (May 1988), pp. 1395-1402, PMID: 3409869
- Holoden C. (2009). Autism Now. *Science*, Vol.323, No.5914, (January 2009), pp. 565, DOI: 10.1126/science.323.5914.565c
- Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nature Genetics*, Vol.37, No.1, (January 2005), pp. 31-40, PMID: 15608638
- Howard G, Eiges R, Gaudet F, Jaenisch R, Eden A. (2008). Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene*, Vol.27, No.3 (January 2008), pp. 404-408, PMID: 17621273
- Inoue K, Kanai M, Tanabe Y, Kubota T, Kashork CD, Wakui K, Fukushima Y, Lupski JR, Shaffer LG. (2001). Prenatal interphase FISH diagnosis of PLP1 duplication associated with Pelizaeus-Merzbacher disease. *Prenatal Diagnosis*, Vol.21, No.13, (December 2001), pp. 1133-1136, PMID: 11787038
- Itoh M, Ide S, Takashima S, Kudo S, Nomura Y, Segawa M, Kubota T, Mori H, Tanaka S, Horie H, Tanabe Y, Goto Y. (2007). Methyl CpG-binding protein 2 (a mutation of which causes Rett syndrome) directly regulates insulin-like growth factor binding protein 3 in mouse and human brains. *Journal of Neuropathology and Experimental Neurology*, Vol.66, No.2, (February 2007), pp. 117-123, PMID: 17278996
- James SJ, Cutler P, Melnyk S, Jernigan S, Janak L, Gaylor DW, Neubrandner JA. (2004). Metabolic biomarkers of increased oxidative stress and impaired methylation capacity in children with autism. *American Journal of Clinical Nutrition*, Vol.80, No.6 (December 2004), pp. 1611-1617, PMID: 15585776
- Jenuwein T, Allis CD. (2001). Translating the histone code. *Science*, Vol.293, No.5532, (August 2001), pp. 1074-1080, PMID: 11498575
- Jessberger S, Nakashima K, Clemenson GD Jr, Mejia E, Mathews E, Ure K, Ogawa S, Sinton CM, Gage FH, Hsieh J. Epigenetic modulation of seizure-induced neurogenesis and cognitive decline. *Journal of Neuroscience*, Vol.27, No.22 (May 2007), pp. 5967-5975, PMID: 17537967
- Jones PA, Baylin SB. (2002). The fundamental role of epigenetic events in cancer. *Nature Review Genetics*, Vol.3, No.6, (June 2002), pp. 415-428, PMID: 12042769

- Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics*, Vol.19, No.2, (June 1998), pp. 187-191, PMID: 9620779
- Kerr B, Silva PA, Walz K, Young JI. Unconventional transcriptional response to environmental enrichment in a mouse model of Rett syndrome. *Public Library of Science One*, Vol.5, No.7, (July 2010), e11534, PMID: 20634955
- Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, Besenbacher S, Jonasdottir A, Sigurdsson A, Kristinsson KT, Jonasdottir A, Frigge ML, Gylfason A, Olason PI, Gudjonsson SA, Sverrisson S, Stacey SN, Sigurgeirsson B, Benediktsson KR, Sigurdsson H, Jonsson T, Benediktsson R, Olafsson JH, Johannsson OT, Hreidarsson AB, Sigurdsson G; DIAGRAM Consortium, Ferguson-Smith AC, Gudbjartsson DF, Thorsteinsdottir U, Stefansson K. (2009). Parental origin of sequence variants associated with complex diseases. *Nature*, Vol.462, No.7275 (December 2009), pp. 868-874, PMID: 20016592
- Kouzarides T. (2007). Chromatin modifications and their function. *Cell*, Vol.128, No.4, (February 2007), pp. 693-705, PMID: 17320507
- Kubota T, Das S, Christian SL, Baylin SB, Herman JG, Ledbetter DH. (1997). Methylation-specific PCR simplifies imprinting analysis. *Nature Genetics*, vol.16, No.1, (May 1997), pp. 16-17, PMID: 9140389
- Kubota T, Nonoyama S, Tonoki H, Masuno M, Imaizumi K, Kojima M, Wakui K, Shimadzu M, Fukushima Y. (1998). A new assay for the analysis of X-chromosome inactivation based on methylation-specific PCR. *Human Genetics*, Vol.104, No.1, (January 1999), pp. 49-55, PMID: 10071192
- Kubota T, Wakui K, Nakamura T, Ohashi H, Watanabe Y, Yoshino M, Kida T, Okamoto N, Matsumura M, Muroya K, Ogata T, Goto Y, Fukushima Y. (2002). The proportion of cells with functional X disomy is associated with the severity of mental retardation in mosaic ring X Turner syndrome females. *Cytogenetics and Genome Research*, Vol.99, No.1-4, (2002), pp. 276-284, PMID: 12900575
- Kubota T, Furuuchi H, Kamoda T, Iwasaki N, Tobita N, Fujiwara N, Goto Y, Matsui A, Sasaki H, Kajii T. (2004). ICF syndrome in a girl with DNA hypomethylation but without detectable DNMT3B mutation. *American Journal of Medical Genetics A*, Vol.129A, No.1, (September 2004), pp. 290-293, PMID: 15326630
- Kucharski R, Maleszka J, Foret S, Maleszka R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. *Science*, Vol.319, No.5871 (June 2009), pp. 1827-1830, PMID: 18339900
- Kumar A, Choi KH, Renthal W, Tsankova NM, Theobald DE, Truong HT, Russo SJ, Laplant Q, Sasaki TS, Whistler KN, Neve RL, Self DW, Nestler EJ. (2005). Chromatin remodeling is a key mechanism underlying cocaine-induced plasticity in striatum. *Neuron*, Vol.48, No.2 (October 2005), pp. 303-314, PMID: 16242410
- Liang G, Lin JC, Wei V, Yoo C, Cheng JC, Nguyen CT, Weisenberger DJ, Egger G, Takai D, Gonzales FA, Jones PA. (2004). Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proceedings of the National Academy Science of the United States of America*, Vol.101, No.19, (May 2004), pp. 7357-7362, PMID: 15123803
- Lillycrop KA, Phillips ES, Jackson AA, Hanson MA, Burdge GC. (2005). Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents

- epigenetic modification of hepatic gene expression in the offspring. *Journal of Nutrition*, Vol.135, No.6 (June 2005), pp. 1382-1386, PMID: 15930441
- Lillycrop KA, Phillips ES, Torrens C, Hanson MA, Jackson AA, Burdge GC. (2008). Feeding pregnant rats a protein-restricted diet persistently alters the methylation of specific cytosines in the hepatic PPAR alpha promoter of the offspring. *British Journal of Nutrition*, Vol.100, No.2, (August 2008), pp. 278-282, PMID: 18186951
- Lonetti G, Angelucci A, Morando L, Boggio EM, Giustetto M, Pizzorusso T. (2010). Early environmental enrichment moderates the behavioral and synaptic phenotype of MeCP2 null mice. *Biological Psychiatry*, Vol.67, No.7, (April 2010), pp. 657-665, PMID: 20172507
- López Castel A, Cleary JD, Pearson CE. (2010). Repeat instability as the basis for human diseases and as a potential target for therapy. *Nature Reviews Molecular Cell Biology*, Vol.11, No.3, (May 2010), pp. 165-170, PMID: 20177394
- Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, Vol.389, No.6648, (January 1997), pp. 251-260, PMID: 9305837
- Mack GS. Epigenetic cancer therapy makes headway. *Journal of the National Cancer Institute*, Vol.98, No.20 (October 2006), pp. 1443-1444, PMID: 17047192
- Martinowich K, Hattori D, Wu H, Fouse S, He F, Hu Y, Fan G, Sun YE. (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science*, Vol.302, No.5646, (October 2003), pp. 890-893, PMID: 14593184
- McGowan PO, Sasaki A, D'Alessio AC, Dymov S, Labonté B, Szyf M, Turecki G, Meaney MJ. (2009). Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature Neuroscience*, Vol.12, No.3, (December 2009), pp. 342-348, PMID: 19234457
- McNairn AJ, Gilbert DM. (2003). Epigenomic replication: linking epigenetics to DNA replication. *Bioessays*, Vol.25, No.7, (July 2003), pp. 647-656, PMID: 12815720
- Matsuda H, Fukuda N, Ueno T, Katakawa M, Wang X, Watanabe T, Matsui S, Aoyama T, Saito K, Bando T, Matsumoto Y, Nagase H, Matsumoto K, Sugiyama H. Transcriptional inhibition of progressive renal disease by gene silencing pyrrole-imidazole polyamide targeting of the transforming growth factor- $\beta$ 1 promoter. *Kidney International*, Vol.79, No.1 (January 2011), pp. 46-56, PMID: 20861821
- Moretti P, Sahoo T, Hyland K, Bottiglieri T, Peters S, del Gaudio D, Roa B, Curry S, Zhu H, Finnell RH, Neul JL, Ramaekers VT, Blau N, Bacino CA, Miller G, Scaglia F. Cerebral folate deficiency with developmental delay, autism, and response to folinic acid. *Neurology*, Vol.64, No.6 (March 2005), pp. 1088-1090 PMID: 15781839
- Morey L, Brenner C, Fazi F, Villa R, Gutierrez A, Buschbeck M, Nervi C, Minucci S, Fuks F, Di Croce L. MBD3, a component of the NuRD complex, facilitates chromatin alteration and deposition of epigenetic marks. *Molecular and Cellular Biology*, Vol.28, No.19 (October 2008), pp. 5912-5923, PMID: 18644863
- Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. (2009). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature*, Vol.435, No.7044, (June 2005), pp. 903-910, PMID: 15959507
- Muotri AR, Zhao C, Marchetto MC, Gage FH. (2009). Environmental influence on L1 retrotransposons in the adult hippocampus. *Hippocampus*, Vol.19, No.10, (November 2009), pp. 1002-1007, PMID: 19771587

- Muotri AR, Marchetto MC, Coufal NG, Oefner R, Yeo G, Nakashima K, Gage FH. (2010). L1 retrotransposition in neurons is modulated by MeCP2. *Nature*, Vol.468, No.7322, (November 2010), pp. 903-910, PMID: 21085180
- Murgatroyd C, Patchev AV, Wu Y, Micale V, Bockmühl Y, Fischer D, Holsboer F, Wotjak CT, Almeida OF, Spengler D. (2009). Dynamic DNA methylation programs persistent adverse effects of early-life stress. *Nature Neuroscience*, Vol.12, No.12, (December 2009), pp. 1559-1566, PMID: 19898468
- Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, Vol.393, No.6683, (May 1998), pp. 386-389, PMID: 9620804
- Nolen LD, Gao S, Han Z, Mann MR, Gie Chung Y, Otte AP, Bartolomei MS, Latham KE. (2005). X chromosome reactivation and regulation in cloned embryos. *Developmental Biology*, Vol.279, No.2, (March 2005), pp. 525-40, PMID: 15733677
- Obi T, Nishioka K, Ross OA, Terada T, Yamazaki K, Sugiura A, Takanashi M, Mizoguchi K, Mori H, Mizuno Y, Hattori N. (2008). Clinicopathologic study of a SNCA gene duplication patient with Parkinson disease and dementia. *Neurology*, Vol.70, No.3, (January 2008), pp. 238-241, PMID: 18195271
- Ohtsuki M, Kimura M, Minoshima M, Suzuki T, Ikeda M, Bando T, Nagase H, Shinohara K, Sugiyama H. (2009). Synthesis and properties of PI polyamide-SAHA conjugate. *Tetrahedron Letters*, Vol.50, No.52, (available online October 2009), pp. 7288-7292
- Okano M, Bell DW, Haber DA, Li E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, Vol.99, No.3, (October 1999), pp. 276-257, PMID: 10555141
- Painter RC, de Rooij SR, Bossuyt PM, Simmers TA, Osmond C, Barker DJ, Bleker OP, Roseboom TJ. Early onset of coronary artery disease after prenatal exposure to the Dutch famine. *American Journal of Clinical Nutrition*, Vol.84, No.2, (August 2006), pp. 322-327, PMID: 16895878
- Park JH, Stoffers DA, Nicholls RD, Simmons RA. (2008). Development of type 2 diabetes following intrauterine growth retardation in rats is associated with progressive epigenetic silencing of Pdx1. *Journal of Clinical Investigation*, Vol.118, No.6, (June 2008), pp. 2316-2324, PMID: 18464933
- Pascual M, Boix J, Felipe V, Guerri C. (2009). Repeated alcohol administration during adolescence causes changes in the mesolimbic dopaminergic and glutamatergic systems and promotes alcohol intake in the adult rat. *Journal of Neurochemistry*, Vol.108, No.4, (February 2009), pp. 920-931, PMID: 19077056
- Persico AM, Bourgeron T. (2006). Searching for ways out of the autism maze: genetic, epigenetic and environmental clues. *Trends in Neurosciences*, Vol.29, No.7, (July 2006), pp. 349-358, PMID: 16808981
- Qiu J. (2006). Epigenetics: unfinished symphony. *Nature*, Vol.441, No.7090, (May 2006), pp. 143-145, PMID: 16688142
- Reiner O, Carrozzo R, Shen Y, Wehnert M, Faustinella F, Dobyns WB, Caskey CT, Ledbetter DH. (1993). Isolation of a Miller-Dieker lissencephaly gene containing G protein beta-subunit-like repeats. *Nature*, Vol.364, No.6439, (August 1993), pp. 717-721, PMID: 835578

- Renthal W, Nestler EJ. Epigenetic mechanisms in drug addiction. *Trends in Molecular Medicine*, Vol.14, No.8, (August 2008), pp. 341-350, PMID: 18635399
- Rimland B. Controversies in the treatment of autistic children: vitamin and drug therapy. (1988). *Journal of Child Neurology*, Vol.3, Suppl (1988), pp. S68-72, PMID: 3058789
- Roa BB, Lupski JR. (1993). Molecular basis of Charcot-Marie-Tooth disease type 1A: gene dosage as a novel mechanism for a common autosomal dominant condition. *American Journal of the Medical Sciences*, Vol.306, No.3, (September 1993), pp. 177-184, PMID: 8128981
- Rodriguez J, Frigola J, Vendrell E, Risques RA, Fraga MF, Morales C, Moreno V, Esteller M, Capellà G, Ribas M, Peinado MA. (2006). Chromosomal instability correlates with genome-wide DNA demethylation in human primary colorectal cancers. *Cancer Research*, Vol.66, No.17, (September 2006), pp. 8462-9468, PMID: 16951157
- Sakashita K, Koike K, Kinoshita T, Shiohara M, Kamijo T, Taniguchi S, Kubota T. (2001). Dynamic DNA methylation change in the CpG island region of p15 during human myeloid development. *Journal of Clinical Investigation*, Vol.108, No.8, (October 2001), pp. 1195-1204, PMID: 1160262
- Satterlee JS, Schübeler D, Ng HH. (2010). Tackling the epigenome: challenges and opportunities for collaboration. *Nature Biotechnology*, Vol.28, No.10 (October 2010), pp. 1039-1044, PMID: 20944594
- Sharma S, Kelly TK, Jones PA. (2010). Epigenetics in cancer. *Carcinogenesis*, Vol.31, No.1, (January 2010), pp. 27-36, PMID: 19752007
- Shirohzu H, Kubota T, Kumazawa A, Sado T, Chijiwa T, Inagaki K, Suetake I, Tajima S, Wakui K, Miki Y, Hayashi M, Fukushima Y, Sasaki H. (2002). Three novel DNMT3B mutations in Japanese patients with ICF syndrome. *American Journal of Medical Genetics*, Vol.112, No.1, (September 2002), pp. 31-37, PMID: 12239717
- Silveira PP, Portella AK, Goldani MZ, Barbieri MA. (2007). Developmental origins of health and disease (DOHaD). *Jornal de Pediatria*, Vol.83, No.6 Pt 1, (November-December 2007), pp. 494-504, PMID: 18074050
- Song J, Noh JH, Lee JH, Eun JW, Ahn YM, Kim SY, Lee SH, Park WS, Yoo NJ, Lee JY, Nam SW. (2005). Increased expression of histone deacetylase 2 is found in human gastric cancer. *APMIS*, Vol.113, No.4(April 2005), pp. 264-268, PMID: 15865607
- St Clair D, Xu M, Wang P, Yu Y, Fang Y, Zhang F, Zheng X, Gu N, Feng G, Sham P, He L. (2005). Rates of adult schizophrenia following prenatal exposure to the Chinese famine of 1959-1961. *JAMA*, Vol.294, No.5, (August 2005), pp. 557-562, PMID: 16077049
- Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO Journal*, Vol.27, No.20, (October 2008), pp. 2681-2690, PMID: 18818694
- Tsankova NM, Berton O, Renthal W, Kumar A, Neve RL, Nestler EJ. (2006). Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. *Nature Neuroscience*, Vol.9, No.4, (April 2006), pp. 519-525, PMID: 16501568
- Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, Slagboom PE, Heijmans BT. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Human Molecular Genetics*, Vol.18, No.21, (November 2009), pp. 4046-4053, PMID: 19656776

- Waddington CH. (1942). Epigenotype. *Endeavour*, Vol.1, (1942), pp. 18-20.
- Wang J, Hevi S, Kurash JK, Lei H, Gay F, Bajko J, Su H, Sun W, Chang H, Xu G, Gaudet F, Li E, Chen T. The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nature Genetics*, Vol.41, No.1 (January 2009), pp. 125-129, PMID: 19098913
- Watanabe H, Fukuoka H, Sugiyama T, Nagai Y, Ogasawara K, Yoshiike N. (2008). Dietary folate intake during pregnancy and birth weight in Japan. *European Journal of Nutrition*, Vol.47, No.6, (September 2008), pp. 341-347, 2008, PMID: 18709472
- Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, Dymov S, Szyf M, Meaney MJ. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, Vol.7, No.8, (August 2004), pp. 847-854, PMID: 15220929
- Xue F, Tian XC, Du F, Kubota C, Taneja M, Dinnyes A, Dai Y, Levine H, Pereira LV, Yang X. (2002). Aberrant patterns of X chromosome inactivation in bovine clones. *Nature Genetics*, Vol.31, No.2, (Jun 2002), pp. 216-220, PMID: 12032569
- Yaoi T, Itoh K, Nakamura K, Ogi H, Fujiwara Y, Fushiki S. (2008). Genome-wide analysis of epigenomic alterations in fetal mouse forebrain after exposure to low doses of bisphenol A. *Biochemical and Biophysical Research Communications*, Vol.376, No.3, (November 2008), pp. 563-567, PMID: 18804091
- Yeargin-Allsopp M, Rice C, Karapurkar T, Doernberg N, Boyle C, Murphy C. (2003). Prevalence of autism in a US metropolitan area. *JAMA*, Vol.289, No.1, (January 2003), pp. 49-55, PMID: 12503976
- Zoghbi HY. (2003). Postnatal neurodevelopmental disorders: meeting at the synapse? *Science*, Vol.302, No.5646, (October 2003), pp. 862-830, PMID: 14593168

# Relations Between Replication and Transcription

Daniel Castro-Roa and Nikolay Zenkin  
*Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences  
Newcastle University, Newcastle upon Tyne  
United Kingdom*

## 1. Introduction

In the cell, RNA polymerase (RNAP) and the replisome share the same template DNA for their respective functions. The rate of replication is typically 20 times faster than transcription and six or more replication forks may be present at the same time on the chromosome (Gotta et al. 1991; Hirose et al. 1983). This implies that collisions between the two machineries are inevitable. In this chapter we will discuss the recent findings on the outcome of collisions between transcription and replication and their consequences, as well as cooperation between the two machineries.

## 2. Stability of RNA polymerase complexes with nucleic acids

In all living organisms transcription, the first step in gene expression, is accomplished by multisubunit RNAP. RNAP is highly evolutionary conserved, both structurally and functionally, in all three domains of life. Although there are differences in the mechanisms of initiation and regulation of transcription, mechanisms of catalysis are remarkably similar in all living organisms. Fully functional eukaryotic and archeal models involve 12-14 subunits (depending on the polymerase type and organism) with a total molecular weight greater than 500 kDa (Cramer et al. 2001). The simplified versions found in bacteria are composed of five subunits (subunits composition  $2\alpha$ ,  $\beta'$ ,  $\beta$  and  $\omega$ ) with a molecular mass of approximately 400 kDa. Transcription is a cyclic pathway roughly composed of three steps: initiation, elongation and termination. RNAP is a flexible machine able to adopt different states required for various transcription stages and the mode and stability of binding of the enzyme to DNA at these stages are different. Various modes of RNAP binding to nucleic acids throughout the transcription cycle is one of the key factors which determine the fate of the replication fork progression along the DNA.

First, we briefly describe steps of transcription cycle and properties of complexes formed by RNAP with the nucleic acids at these steps.

The first stage of transcription is initiation. This phase begins with the search of the promoter by the core enzyme equipped with the transcription initiation  $\sigma$  factor through a scanning mechanism (Park et al. 1982; von Hippel 2007). The means by which RNAP is thought to find the promoter involves the tracking of a groove of the double helix throughout electrostatic interactions reinforced by the entropy that results from the

displacement of the counterion cloud that surrounds the DNA (Sakata-Sogawa et al. 2004). These initial interactions of the enzyme with DNA are weak and therefore unstable.

$\sigma$  subunit recognizes a promoter, which usually is comprised of two hexameric sequences around -10 and -35 positions relative to transcription start site, to which it has high affinity. It has been demonstrated that promoter sequences are highly flexible making them more prone to RNAP binding (Ozoline et al. 1999; Travers 2004). The interactions of RNAP with promoter DNA are expanded by the wrapping of the DNA (demonstrated by DNA footprinting experiments) on the surface of the enzyme making the complex more stable (Ozoline et al. 1995). Wrapping also facilitates further rearrangements in both, RNAP and DNA. During this state, known as promoter closed complex (Li et al. 1998), the DNA helix remains double stranded and the complex, though being relatively stable, remains sensitive to high ionic strength and competitors such as heparin (Coulombe et al. 1999). The stability of this complex depends on the sequence of the promoter. Generally, the farther the sequence of a promoter (-10 and -35 elements) from the consensus, the less stable promoter complex will be formed on it (Fenton et al. 2001). Promoter sequence also determines the capacity of RNAP to compete with repressors and nucleoid proteins (Grainger et al. 2006).

After formation of the closed complex the double helix of DNA is destabilised by action of specific residues in the  $\sigma$  factor on a precise region of the promoter (Aiyar et al. 1994; deHaseth et al. 1995; Murakami et al. 2002). Then, the enzyme melts the double helix of the DNA and form a stretch ~17 nucleotides (nt) of unwound DNA known as the transcription bubble (for some particular  $\sigma$  factors, the energy required is obtained by ATP hydrolysis (Merrick 1993)). The melting generates further rearrangements of the DNA inside RNAP, placing the downstream DNA into the enzyme's DNA-binding clamp and positioning the template DNA in register with the active centre making the complex catalytically competent (Murakami et al. 2002; Vassylyev et al. 2002). This new configuration of RNAP is referred to as open promoter complex (Li et al. 1998; Mekler et al. 2002). At this point RNAP occupies a total of ~35 bp of the DNA and has undergone several structural rearrangements that provide higher stability compared to afore mentioned closed promoter complexes. Open promoter complex is capable to withstand higher ionic strength (200 mM KCl) and becomes resistant to competitors (Reppas et al. 2006; von Hippel et al. 1984).

Initiation of transcription starts with synthesis of short RNA transcripts (2-9 nucleotides long). Given that RNAP remains anchored to the promoter by  $\sigma$  factor, synthesis involves pulling of a stretch of downstream DNA of the same size inside the main channel of RNAP. This phenomenon is referred to as "scrunching". Scrunching results in increase of the size of the transcription bubble (given that its upstream edge is kept at the same position by  $\sigma$  subunit). Initiation ends when the energy accumulated during "scrunching" surpasses the energy that is anchoring RNAP to the promoter releasing the enzyme from it followed by the subsequent ejection of the  $\sigma$  factor leaving RNAP ready to elongate the RNA chain.

The stability of RNAP during elongation is greatly enhanced by multiple protein-nucleic acids contacts that are formed after synthesis of ~8 nucleotide long RNA and  $\sigma$  subunit release. In simple terms, if using a comparison of RNAP to a crab claw, the claw in initiation is more open, while in elongation it closes on the RNA-DNA hybrid almost fully surrounding it. These interactions of RNAP with nucleic acids make elongation complex highly stable and resistant to very high ionic strength (1 M KCl) and to competitors (Kuznedelov et al. 2002). The structure of the elongation complex is schematically shown in Figure 1. The characteristic feature of the elongation complex, that plays major role in its

stability, is the 9-10 base pair RNA-DNA hybrid. The length of the hybrid as well as the size of the transcription bubble remain the same throughout elongation.

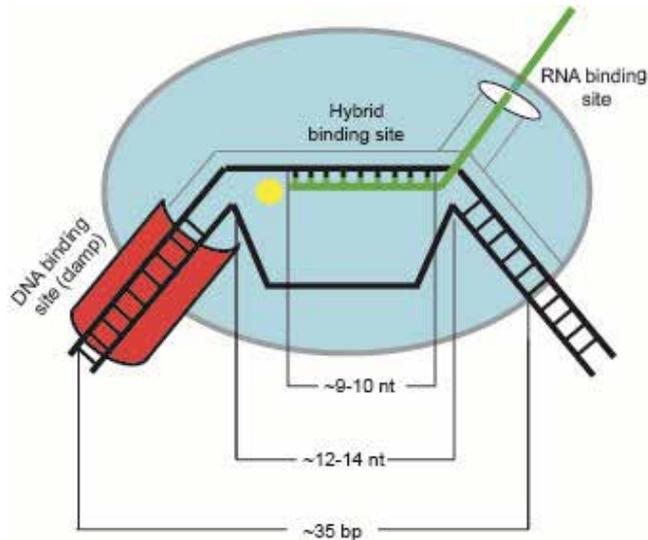


Fig. 1. Schematic representation of RNAP core enzyme (blue) during elongation. Catalytic  $Mg^{2+}$  ions of active centre are shown as yellow circle. DNA is black, RNA is green. The three major interaction sites with nucleic acids that provide stability to the elongation complex and their length are shown.

Besides unusual stability of the elongation complex, transcribing RNAP generates a considerable force. The calculated force of the actively transcribing RNAP is 20 pN per molecule of RNAP and is additive when more RNAP molecules collide with each other (Wang et al. 1998). This allows RNAP to overcome unwanted events and also dislodge proteins bound to DNA template (Epshtein et al. 2003). This powerful and stably bound to DNA machine is capable to proceed through millions (in eukaryotes) of base pairs without losing contact with the DNA template and nascent RNA.

Though RNAP is supposed to processively transcribe through long distances without interruption it is a subject to stringent regulation. RNAP recognizes different signals in the DNA that slow down its movement via various mechanisms. Some pause signals lead to structural modification of RNAP active centre, which slows down catalysis of the phosphodiester bond (Landick 2009). An unusual property of RNAP that also plays role in pausing is its ability to move backwards along the DNA template, a phenomenon called backtracking. During backtracking, RNAP shifts backwards in a manner of a zipper: the 3' end of RNA disengages from the template DNA strand and the active centre, while the rear end of RNA-DNA hybrid RNA anneals back to the template. This keeps the length of the RNA-DNA hybrid the same as in active elongation complex, which means that this arrested complex is as stable as the active one. These arrested complexes require separate factors (e.g. cleavage Gre factors, transcription-DNA repair coupling factor Mfd (Borukhov et al. 1993; Park et al. 2002)) for their resolution. Pausing and backtracking increase the probability of RNA polymerase encountering the replisome, and may be detrimental for genome integrity and cell viability if not resolved.

When RNAP encounters a termination sequence (typically a GC rich palindromic sequence followed by a stretch of uridines ~8nt long) it forms a weak RNA:DNA hybrid containing the poly U track of the terminator (Gusarov et al. 1999). This pauses (by backtracking mechanism) RNAP and gives enough time for a hairpin loop to form in the RNA exit channel followed by destabilization of the RNA:DNA hybrid leading to RNA release and aperture of the DNA binding clamp resulting in RNAP dissociation (Epshtein et al. 2007; Santangelo et al. 2004)

### 3. Interactions of transcription with replication

#### 3.1 Collisions of RNAP and the replisome

The effect on collisions between RNAP and the replisome depends greatly in their directionality. Co-directional interactions occur in the leading strand whereas head-on collisions take place on the lagging strand (Figure. 2). Interestingly, analyses of genome organisation have shown that most of the essential genes, highly transcribed genes, and longer genes (Huvet et al. 2007; Omont et al. 2004; Price et al. 2005; Rocha et al. 2003) are oriented in the same direction as replication on the leading strand. This arrangement was also observed in *B. subtilis*, *Borrelia burgdorferi*, *Treponema pallidum*, *Haemophilus influenzae*, *Helicobacter pylori*, *Mycoplasma genitalium*, *mycoplasma pneumoniae*, and in some bacteriophages (McLean et al. 1998). The difference in gene arrangement is considered to be the outcome of natural selection on genome organisation (Mirkin et al. 2005). Although the evolution pressure that resulted in this organisation is still unclear, it is thought to be, at least in part, determined by differences of the interactions of RNAP with the replication fork during co-directional versus head-on collisions (Brewer 1988). This could be detrimental for RNAP completion of transcripts affecting the production of correct, full length proteins or most likely, because it could stall the replisome inhibiting the replication fork progression (Deng et al. 2005; Srivatsan et al. 2010). These mechanisms will be discussed later.

##### 3.1.1 The effects of RNAP on replication and genome integrity

The manner in which replication forks compete with RNAP has been evaluated *in vivo* and *in vitro*. Unsurprisingly, most of the research has been performed in *B. subtilis* and *E. coli*. Early studies in *E. coli* analysed the outcome of the collision between the two machineries by inserting in the chromosome an IPTG inducible and unidirectional ColE1 replication origin upstream (for analysis of co-directional collisions with RNAP) or downstream (for analysis of head on collisions with RNAP) of the ribosomal operon *rrnB*.

This particular operon (5.4 kb) was chosen because of its length and high rate of transcription would make it easier to be visualised using electron microscopy (French 1992). In the case of co-directional orientation, replication fork progression was the same in both non-transcribed and transcribed regions where no accumulation of the replisome was observed. Interestingly, during replication of the transcribed regions RNAP molecules were absent behind the replication fork but were still present in front of it. Additionally, after the replisome had completely passed through the transcribed region repopulation by RNAP took place. These observations suggested that the replisome was not neither slowed down nor displaced from the DNA by transcribing RNAP (Figure 3).

This observation is consistent with recent *in vitro* work which showed that after co-directional collision, RNAP is dislodged from the DNA and the replication fork resumes elongation by using the displaced RNA as a primer (Pomerantz et al. 2008). The mechanism in which RNAP acts as primase will be discussed later.

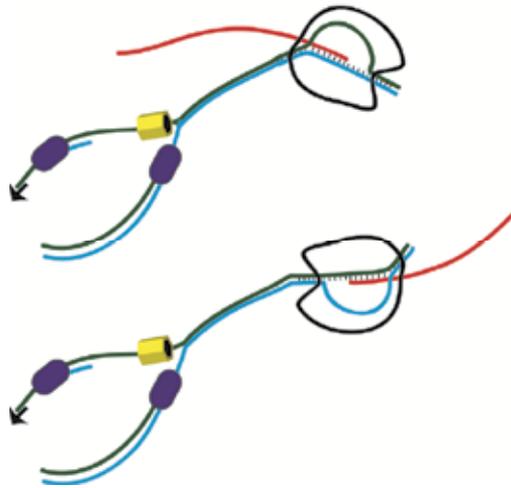


Fig. 2. Schematic representation of Co-directional and head-on collision between the replication fork and RNAP. RNAP with its characteristic claw-like shape is shown in black (right part) with nascent RNA (red line) and template (green) and non-template (blue) strands. The replisome (on the left) is represented by yellow 3D hexagon (DnaB) and DNA polymerases as purple ovals in lagging and leading strands respectively.

In the case of head-on collisions, RNAP was also dislodged but the replication fork progressed through the operon much more slowly than in the co-directional counterpart. In other studies using DNA microarrays of wild type *B. subtilis* and of mutants carrying a long stretch of inverted DNA in the chromosome (to induce head-on collisions) the reduction of the speed of the replisome was also observed. In the wild type scenario, no interference of the fork progression by transcriptional machinery was detected. Surprisingly, in the case of the inverted mutants, replication was generally slowed down in both highly and low transcribed regions reinforcing the theory that the genome is organised to favour the co-directionality of both machineries (Wang et al. 2007).

The slowing down of the replication fork caused by head-on collisions could result in genomic instability due to inaccurate restart of the replisome by new collisions with transcribing RNAP molecules. Another possibility that could lead to instability is that, after collision, stalled replication fork could unwind generating a four stranded DNA structure that resembles a Holliday junction (a process known as replication fork regression) (Atkinson et al. 2009). This could lead to double strand breaks by cleavage of the DNA performed by Holiday junction resolvase affecting cell fitness and viability (McGlynn et al. 2000). In this study, performed in *E. coli* it was proposed that halted RNAP represents an impediment for replication fork progression and upon collisions it may require to restart. To reach this conclusion, cells were irradiated with UV light damaging the DNA. Transcribing RNAP cannot surpass the lesions on the DNA and therefore it stalls. Replication was observed to restart upon encountering the stalled RNAP which was assisted by endonucleolytic cleavage of the DNA by the Holiday junction resolvase RuvABC. Mutants lacking the RuvABC system, in the presence of high concentrations of the alarmone ppGpp were capable to survive UV irradiation demonstrating that few events of replication restart were occurring. This finding suggests that transcription factors are also involved in the resolution of conflicts between the two machineries.

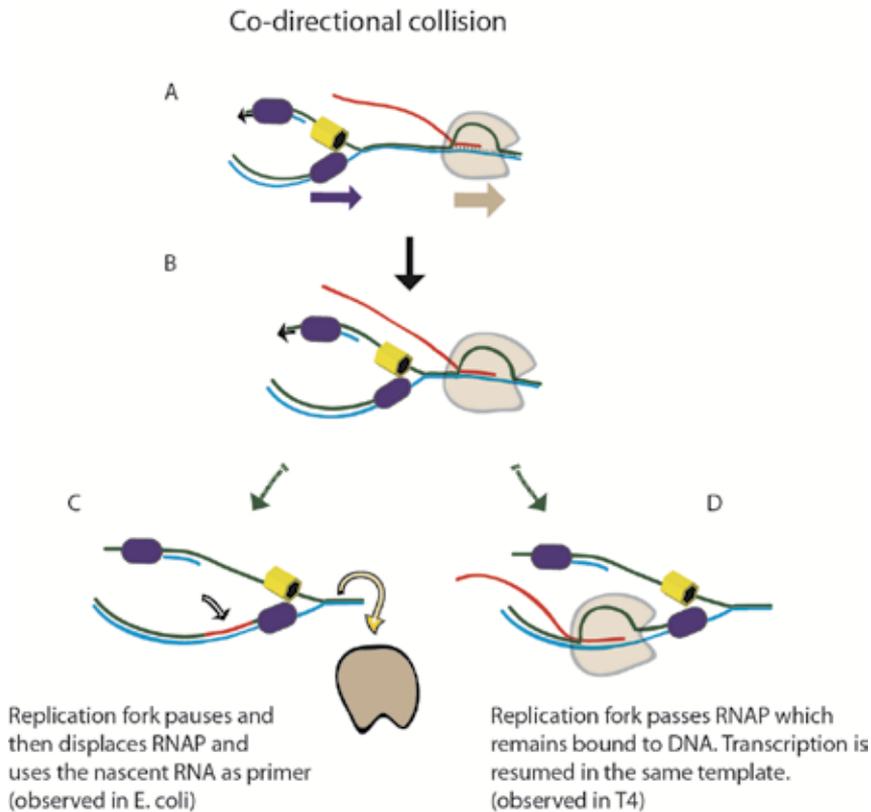


Fig. 3. Schematic representation of the observed outcome of the co-directional collision between the replication fork and RNAP. A) Machineries with arrows indicating their directionality. B) Replication fork colliding with RNAP. C) Replication fork displaces RNAP and utilises the newly synthesised RNA as primer. D) RNAP is temporarily displaced from the DNA allowing the replication fork to pass by. Then, transcription is resumed in the same template.

The mechanism in which the replisome is slowed down could be explained by an observation made *in vitro* with RNAP and replication system purified from *E. coli*. It was demonstrated that during head-on collisions RNAP was also displaced from the DNA, as in the case of co-directional collisions, but the replisome was stalled, though it did not fall apart and eventually was capable to resume elongation (Pomerantz et al. 2010). Earlier studies with plasmids suggested (Mirkin et al. 2005) that the outcome of head-on collision is due to the actual physical interactions between the two machineries (Figure 4). Another possible explanation for this phenomenon was that the topology of the DNA (positive supercoiling generated in front of the transcribing RNAP) could be the cause of the decrease in rate of replication (Brewer 1988; Deshpande et al. 1996).

Besides studying the interactions of actively elongating RNAP, different strategies were used to investigate the effects of initiating and terminating RNAP on the replication fork. In *E. coli*, an *in vivo* system was set up using two plasmids. One contained the strong T7A1 promoter where the equilibrium from promoter clearance towards abortive initiation was shifted by modifying the initial transcribed sequence. This would allow RNAP to be stably

bound to the promoter without entering into elongation. When challenged with replication in co-directional orientation, it was not detrimental for the replisome but during head-on collisions it turned out to be inhibitory. The second plasmid contained the same promoter but in this case RNAP was allowed to enter into elongation normally. Further downstream a transcription termination signal was placed. Interestingly, the replisome was stalled at co-oriented collisions with the terminating RNAP which, as mentioned before, could be in the backtracked state that occurs prior to termination. It is possible that the resulting effect upon encountering promoters and terminators might serve as polar “replication punctuation marks” that could facilitate clearing of mutations acquired in transcribed parts of the genome, by the mismatch repair or gene conversion machineries (Mirkin et al. 2006).

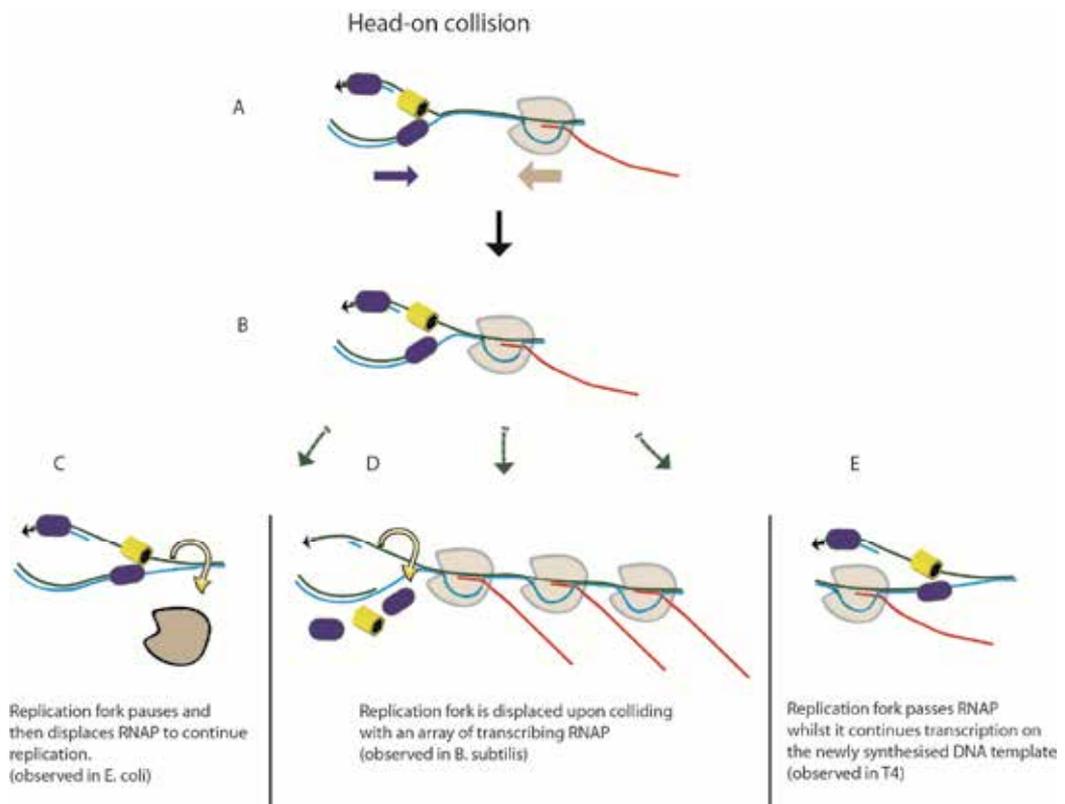


Fig. 4. Schematic representation of the observed outcomes of the head-on collision between the replication fork and RNAP. A) Machineries with arrows indicating their directionality. B) Replication fork colliding with RNAP. C) Replication fork briefly pauses and then displaces RNAP from the DNA D) Replication fork collapses after colliding with an array of transcribing RNAP. E) RNAP is temporarily dislodged from the DNA by the replisome but it is still bound to the DNA. After the replication fork has passed, transcription is resumed in the new DNA template

In contrast to the observations made with *E. coli*, studies conducted *in vitro* with reconstituted components using the bacteriophage T4 as a model showed that progression of the replication fork was paused regardless of the direction of the collision with longer stalling during head-on encounters (Liu et al. 1995). However, it was suggested that after co-

directional collisions RNAP was still bound to the DNA as the replication passed it by and was capable to resume transcription in the same strand. It was also postulated that during head-on collisions, RNAP would continue transcribing but both the enzyme and the transcript would switch strand, from the already existing (where the collision takes place) to the newly synthesised one (Liu et al. 1993).

Although the mechanism in which this strand swapping might occur is still unclear it was proposed that RNAP and transcript transiently disengage from the template but remained bound to the DNA and an active elongation complex reforms after the replisome has passed through.

Most recent studies monitoring the association of the replicative helicase and replication start proteins by ChIP, ChIP-qPCR and ChIP-chip also observed that collision leads to replication restart *in vivo* in *B. subtilis* in highly transcribed genes both co-directional and head-on collisions (Merrikh et al. 2011).

Analysis of the rolling circle M13 replication of plasmids in head-on orientation towards RNAP revealed a significant loss of plasmid stability along with deletions in and downstream the transcribed region. Notably, it was also observed that replication of the parental plasmid decreased (due to interference of the replisome by RNAP) leading to an enrichment of the plasmids that had undergone deletions (Vilette et al. 1996).

Mutations as result of head-on collisions between the replisome and transcription have also been reported for bacterial chromosomal DNA. In *B. subtilis*, analysis of the *rpoB* gene (which encodes for the RNAP  $\beta$  subunit) positioned in the lagging strand, head-on towards replication fork progression, showed an increased mutation rate ( $\sim 3$  fold) compared to the isogenic control with no inversion (Wang et al. 2007). Since the observation was done in one of the longest transcriptional units, the study also suggested that replication fork progression is affected more dramatically when not just a single RNAP but an array of RNAP molecules are transcribing the gene where the collision takes place. This supports the hypothesis that states that long genes are arranged in a co-directional manner to avoid this sort of events (Omont et al. 2004). In the same study major disruption of the replisome was detected on ribosomal genes orientated head-on to the replication fork by following RecA filaments which indicate generation of single stranded DNA or double stranded DNA ends. In eukaryotic cells, even though the rate of replication is similar to the rate of transcription and there are specific barriers that block the replisome from entering ribosomal genes (Brewer et al. 1988), interactions of the two machineries have also been characterised. In *Saccharomyces cerevisiae* by genome wide analysis it was shown that replication fork is paused by RNAP II in highly transcribed genes (Azvolinsky et al. 2009).

Other studies on yeast showed that genome integrity is compromised by impairment of the replisome during head-on collisions with RNAP II which might promote transcription associated recombination (Kim et al. 2007; Prado et al. 2005).

All these findings relate directly to the mechanisms of overcoming obstacles by the two machineries with major emphasis on the outcome of the replisome in terms of its integrity and stability that clearly is affected depending on the directionality of the collisions. However there are other direct or indirect roles played by RNAP in replication.

#### **4. Effects of transcription on replication initiation and completion**

DNA replication is normally initiated from a defined position on the chromosome, the origin, when assembly of the replication complex takes place. There is evidence that

transcription might be involved in the regulation of the initiation of replication. After unwinding of the DNA duplex, DNA polymerases (DNAP) cannot initiate the synthesis of a new DNA chain, without pre-existing primers. Due to the 5'→3' directionality of the DNAP one of the strands is polymerised continuously (leading strand) and the other one in a discontinuous manner (lagging strand). The leading strand needs to be primed at the beginning, whereas the lagging strand requires being primed frequently to generate short chains (Okazaki fragments) until its completion (Masai et al. 1996). This process is achieved by a special primase, which is part of the replication machinery, but action of RNAP has also been described in the priming of replication (Hassan et al. 1994).

The mechanisms in which the transcriptional machinery is involved in replication initiation and priming have been studied over the past 30 years and will be discussed below.

#### 4.1 Transcription of *mioC* and its effect on replication initiation

The replication origin *oriC* is a highly conserved region in enteric bacteria. In its ~245 bp it contains five *dnaA* binding sites (Kornberg 1991; Messer et al. 2001) and the *mioC* gene which contains no apparent transcriptional terminator which allows most transcripts to read through the *oriC* region (Nozaki et al. 1988; Schauzu et al. 1987). First indication of RNAP involvement in initiation was obtained in early experiments in *E. coli* which demonstrated that replication was sensitive to rifampicin (which inhibits RNAP) independently from the effects on protein synthesis (Lark 1972; Messer 1972). Temperature sensitive mutants of DnaA are lethal at the non-permissive temperature. Interestingly, this effect is suppressed by specific mutations on *rpoB* (Atlung 1984; Bagdasarian et al. 1977). Also, an increased copy number of *oriC* containing chromosomes was observed in *rpoB* or *rpoC* mutants (Rasmussen et al. 1983). This suggested that modified RNAP could assist replication initiation but the actual possible mechanism remained obscure.

*In vitro* experiments of the replication from *oriC* showed that RNAP is required under specific conditions that affect DNA unwinding such as reduced negative superhelicity or reduced temperature (Baker et al. 1988). Similar to the case of phage  $\lambda$  replication where transcription from the  $P_R$  promoter induces opening of the DNA under the same conditions (Keppel 1988). It was also suggested that RNAP involvement was not due to priming activity because transcription terminated by incorporation of 3'-dATP had no effect on replication initiation, meaning that there is no need for the transcript to cross the *oriC* sequence. (Baker et al. 1988). Contrary to these observations, when primase was omitted from the reactions in the presence of RNAP, the replication efficiency was reduced tenfold but not inhibited contemplating the possibility that the transcripts could have been used as primers (Ogawa et al. 1985).

Further analysis showed that transcription originated from the *mioC* promoter enters the *oriC* region and seems to be involved in control of initiation of replication (deWind 1987). A study revealed that the copy number of minichromosomes in the presence of *mioC* transcription, was 2 fold higher compared to the *mioC* deletion control. An apparently contradicting observation was made when the wild type *mioC* promoter was replaced by stronger *lac* or *Cm<sup>r</sup>* promoters. Stronger transcription that entered *oriC* resulted in inhibition of minichromosome replication (Tanaka et al. 1983). Other experiments also performed in minichromosomes demonstrated that DnaA inhibits transcription from *mioC* promoter by binding to the *dnaA* box which is located right upstream of it (Lothar et al. 1985; Stuitje et al. 1986).

Notably, in a recent study on *E. coli* chromosome, a *mioC* gene mutant in which TN5 transposon was inserted in the promoter region was isolated. This mutant allowed *mioC* transcription constitutively on the chromosome throughout the cell cycle. It was shown that the read-through of mutant *mioC* through *oriC* impaired the initiation of chromosomal replication and also suggested that the transcriptional fluctuation during the cell cycle might be essential for the control of initiation of replication (Su'etsugu et al. 2003). This was in agreement with observations made in minichromosomes where transcription of *mioC* gene appeared to be repressed prior to the initiation of replication and derepressed right after initiation of replication (Ogawa et al. 1994; Theisen et al. 1993). The proposed mechanisms by which transcription from *mioC* could inhibit replication were that transcription could impair formation of initiation complex at the origin via physical interactions which would displace DnaA from DNA or through the generation of local changes in the DNA topology destabilizing the initiation complex. However, while the replication time in cells lacking *mioC* promoter was unaffected (Bates et al. 1997; Lobner-Olesen et al. 1992), transcription from this promoter became essential for cell viability when the *dnaA* box R4 was deleted suggesting a positive role for *mioC* transcription in replication initiation (Bates et al. 1997). Despite of all the data obtained over 3 decades, no clear consensus on the role of transcription from the *mioC* gene on replication has been achieved. The lack of consensus in the observations among these studies could be due to the differences in the plasmid structures that could result in different transcription modes which affect the effect on replication (Masai et al. 1988). Further investigation is required to elucidate the mechanisms of regulation of the two machineries.

#### 4.2 The role of RNAP as primase

As mentioned in the previous section, replication initiation and elongation of the lagging strand needs to be primed. DNAP is incapable of initiating DNA synthesis *de novo* and require an accessible and properly positioned hydroxyl group to attach a deoxyribonucleotide. Such insufficiency might have emerged through evolution to increase processivity of these enzymes by impairing the sequence specific interactions with template, which are needed for specific initiation of *de novo* synthesis (Zenkin et al. 2008). In the majority of cases DNA replication is primed by enzymes called primases. Primases are DNA-dependent RNA polymerases that synthesise oligoribonucleotides that remain annealed to template DNA and are used by DNAP as primers. Some other strategies have been acquired by different replicons to fulfil these requirements. For example, preformed tRNA (by retroviral reverse transcriptases), DNA primers generated by endonucleolytic cleavage (by gpA of  $\phi$ X174), and serine OH group (adenoviral 55 kDa terminal protein). Strikingly, some replicons rely on RNAP and not on primases for initiation of replication of their genomes. This event was first observed 30 years ago in reconstituted replication system by A. Kornberg (Brutlag et al. 1971; Wickner et al. 1972). Before the discovery of DnaG (Bouche et al. 1975; Rowen et al. 1978; Schekman et al. 1972) RNAP was thought to be the only primase in the cell.

RNAP priming activity is directly utilised in the replication of the leading strand of filamentous and T-odd phages. The filamentous phages, such as M13, fd and f1, contain a simple, relatively small (6407 nt long), single stranded circular genome which contains 10 genes (Herrmann et al. 1978; Ohsumi et al. 1978). An important feature of the genome is an intergenic region that takes up to 8% of the genome. This region does not encode proteins but serves as the origin for the + and - strand synthesis (Chen et al. 1978). This region contains five

palindromes (Zinder et al. 1985). After infection, adsorption and penetration, the replication of the single stranded genome starts. RNAP equipped with  $\sigma^{70}$  binds to one of the hairpins (which has no homology to any *E. coli* promoter) in the (+) strand of the intergenic region. Though RNAP bound to this region protects ~125 nt, only ~20 bp long hairpin is enough for specific recognition. RNAP synthesises a short (18 nt long) RNA chain (Zenkin et al. 2006). This RNA serves as a primer and is further extended by DNAP (Kaguni et al. 1982).

The mechanism in which RNAP synthesises the M13 replication primer has been recently elucidated (Figure 5). It was observed that the minimal M13 origin occupies the downstream DNA binding channel and positions the single-stranded template in the active centre of RNAP leaving the enzyme ready to initiate synthesis of RNA (this is very unusual mode of initiation as compared with "normal" transcription where, upon RNAP binding, the double stranded DNA is melted and abortive synthesis of RNA occurs until the enzyme has stabilised both the transcript and the DNA template in the active centre. This series of events allows further synthesis of RNA). RNAP lacking the  $\sigma$  subunit (core enzyme), was observed to be capable to specifically initiate RNA synthesis on the M13 *ori* hairpin. The synthesis of a dinucleotide suggested that the stabilisation of the initiation complex was achieved through "structure-specific" interactions of the downstream DNA binding channel of RNAP with the *ori* hairpin. However, no synthesis beyond the first phosphodiester bond was observed. When RNAP containing  $\sigma$  subunit was used, stabilisation of short transcripts was achieved, allowing further extension of the RNA.  $\sigma$  factor was shown to be needed at least until a trinucleotide was formed. When the growing RNA:DNA hybrid reached the size of 10-12 bases, steric constraints lead to the disengagement of the rear part of the hybrid from the active cleft and leads to formation of overextended hybrid. Upon synthesis of an 18-20 bases long RNA irreversible structural rearrangement is induced that leads to accommodation of the hybrid in the downstream DNA binding channel. After the rearrangement, the exposed 3' end of the RNA becomes available to be used by DNAP (Zenkin et al. 2008). It is still unclear if DNAP recognises the 3' end of the RNA in complex with RNAP or if RNAP is displaced before initiation of DNA synthesis.

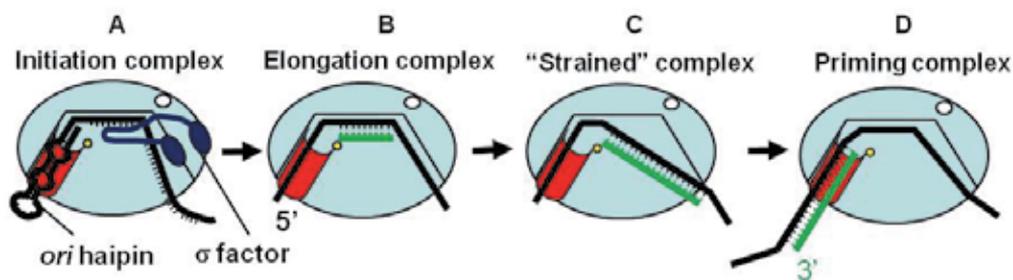


Fig. 5. Schematic representation of the mechanism of primer synthesis by RNAP. A) RNAP (blue) accommodates the *ori* hairpin (black) in the downstream DNA binding channel (red). Synthesis of RNA starts and the transcript is stabilised by  $\sigma$  factor (purple). B) Synthesis of an RNA (green) 8-9 nt long, proceeds in a normal way. C) Upon reaching the RNA:DNA hybrid of the size of 10-12 nt long, the rear end of the hybrid disengages from the active cleft. D) Upon reaching critical length of 18-20 nt, RNAP backslides to reposition the hybrid in the downstream DNA binding channel making the 3' end of RNA available to serve as primer for replication.

RNAP synthesised RNA that promote replication have also been characterised in the *E. coli* double stranded plasmid Cole1. As observed *in vitro*, initiation of replication of the continuous strand depends on three host enzymes: RNAP, RNase H and DNAP I. It was shown that if RNAP was inhibited by rifampicin replication of the plasmid would not occur (Itoh et al. 1980; Itoh et al. 1982). RNAP was shown to synthesise a 555 nt transcript which enters a region where the transition between RNA and DNA occurs (Selzer et al. 1982). The transcript, called RNA II contains a series of stems and loops which are necessary for the formation of an extended RNA:DNA hybrid at the end of the transcript (Masai et al. 1996). RNA II is tightly regulated by the antisense RNA I which binds to the 5' of RNA II affecting its secondary structure inhibiting the formation of the RNA:DNA hybrid (Tamm et al. 1983). Cleavage of RNA:DNA hybrid mediated by RNase H exposes the 3'-OH ends needed for DNA synthesis by DNAP I (Naito et al. 1984) which after polymerisation of 400 nt is replaced by the highly processive DNAP III which will continue the replication of the strand. 5' end secondary structure is required to abolish reformation of upstream DNA duplex behind RNAP, thus forcing RNAP to transcribe single stranded template. Therefore, the mechanism of primer formation may be the same as for single stranded M13 phage.

In a recent study conducted *in vitro* with a reconstituted *E. coli* replisome lacking primase, the transcriptional machineries from both *E. coli* and T7 RNA polymerases were challenged with the replication fork progressing in a co-directional manner (Pomerantz et al. 2008). Replication was initiated on a forked DNA template. The progression of replication was followed by the analysis of the radioactively labeled nascent DNA strand. RNAP, (either *E. coli* or T7) elongation complex was stalled in front of replisome by the omission of one of the 4 NTPs. Strikingly, after displacement of RNAP from the template, the nascent transcript apparently remained annealed to DNA template and was used by replisome as a primer. This can explain the outcomes of co-directional collisions observed *in vivo* (see above). The mechanism of primer formation in this case may also be similar to M13, given that the replisome approaching RNAP from behind restricts the upstream DNA duplex formation resulting in RNAP synthesising on single stranded DNA.

In eukaryotic cells evidence of RNAP acting as a specific primase which initiates the synthesis of the heavy-strand and the light-strand of the double stranded 16.6 kb long mitochondrial DNA has also been observed (Fuste et al. 2010; Wanrooij et al. 2008).

## 5. Conclusions

Altogether, through the data generated in four decades, it has become evident that transcription has played a major role in the evolution of genome organisation, architecture and integrity. Even though the mechanisms and the outcomes of the collisions of the replisome with RNAP, the regulatory function of transcription on replication initiation and its contribution in elongation by serving as primase have been demonstrated, many other questions remain unanswered. Further studies on the mechanisms in which *mioc* acts on *oriC* replication are needed. Also, the mechanism(s) of action of transcription factors such as *DksA* and the alarmone *ppGpp* modulate not only the activity of RNAP but also its interactions with replication and other machineries require to be properly elucidated. The phenomenon of template swapping by RNAP upon head-on collisions and the by-passing of the replication fork in co-directional collisions with the T4 replication machinery remains obscure and needs to be clarified.

## 6. References

- Aiyar, S. E., et al. (1994). "Mutations in sigma factor that affect the temperature dependence of transcription from a promoter, but not from a mismatch bubble in double-stranded DNA." *Biochemistry* 33(38): 11501-11506.
- Atkinson, J., et al. (2009). "Replication fork reversal and the maintenance of genome stability." *Nucleic acids research* 37(11): 3475-3492.
- Atlung, T. (1984). "Allele-specific suppression of dnaA(Ts) mutations by rpoB mutations in Escherichia coli." *Molecular & general genetics : MGG* 197(1): 125-128.
- Azvolinsky, A., et al. (2009). "Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae." *Molecular cell* 34(6): 722-734.
- Bagdasarian, M. M., et al. (1977). "Suppression of the DnaA phenotype by mutations in the rpoB cistron of ribonucleic acid polymerase in Salmonella typhimurium and Escherichia coli." *Journal of bacteriology* 130(2): 577-582.
- Baker, T. A., et al. (1988). "Transcriptional activation of initiation of replication from the E. coli chromosomal origin: an RNA-DNA hybrid near oriC." *Cell* 55(1): 113-123.
- Bates, D. B., et al. (1997). "The absence of effect of gid or mioC transcription on the initiation of chromosomal replication in Escherichia coli." *Proceedings of the National Academy of Sciences of the United States of America* 94(23): 12497-12502.
- Borukhov, S., et al. (1993). "Transcript cleavage factors from E. coli." *Cell* 72(3): 459-466.
- Bouche, J. P., et al. (1975). "dnaG gene product, a rifampicin-resistant RNA polymerase, initiates the conversion of a single-stranded coliphage DNA to its duplex replicative form." *The Journal of biological chemistry* 250(15): 5995-6001.
- Brewer, B. J. (1988). "When polymerases collide: replication and the transcriptional organization of the E. coli chromosome." *Cell* 53(5): 679-686.
- Brewer, B. J., et al. (1988). "A replication fork barrier at the 3' end of yeast ribosomal RNA genes." *Cell* 55(4): 637-643.
- Brutlag, D., et al. (1971). "A possible role for RNA polymerase in the initiation of M13 DNA synthesis." *Proceedings of the National Academy of Sciences of the United States of America* 68(11): 2826-2829.
- Chen, T. C., et al. (1978). "Replication of bacteriophage M13. XIII. Structure and replication of cloned M13 miniphage." *Journal of molecular biology* 125(2): 107-121.
- Coulombe, B., et al. (1999). "DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms." *Microbiol Mol Biol Rev*: 457-478.
- Cramer, P., et al. (2001). "Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution." *Science* 292(5523): 1863-1876.
- deHaseth, P. L., et al. (1995). "Open complex formation by Escherichia coli RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA." *Mol Microbiol* 16(5): 817-824.
- Deng, S., et al. (2005). "Organization of supercoil domains and their reorganization by transcription." *Molecular microbiology* 57(6): 1511-1521.
- Deshpande, A. M., et al. (1996). "DNA replication fork pause sites dependent on transcription." *Science* 272(5264): 1030-1033.
- deWind, N., Parren, P., Stuitje, A. R. and Meijer, < (1987). "Evidence for the involvement of the 16kD gene promoter in initiation of chromosomal replication of Escherichia coli

- strains carrying a B/r-derived replication origin." *Nucleic acids research* 15: 4907-4914.
- Epshtein, V., et al. (2007). "An allosteric path to transcription termination." *Molecular cell* 28(6): 991-1001.
- Epshtein, V., et al. (2003). "Transcription through the roadblocks: the role of RNA polymerase cooperation." *EMBO J* 22(18): 4719-4727.
- Fenton, M. S., et al. (2001). "Function of the bacterial TATAAT -10 element as single-stranded DNA during RNA polymerase isomerization." *Proceedings of the National Academy of Sciences of the United States of America* 98(16): 9020-9025.
- French, S. (1992). "Consequences of replication fork movement through transcription units in vivo." *Science* 258(5086): 1362-1365.
- Fuste, J. M., et al. (2010). "Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication." *Molecular cell* 37(1): 67-78.
- Gotta, S. L., et al. (1991). "rRNA transcription rate in Escherichia coli." *Journal of bacteriology* 173(20): 6647-6649.
- Grainger, D. C., et al. (2006). "Association of nucleoid proteins with coding and non-coding segments of the Escherichia coli genome." *Nucleic acids research* 34(16): 4642-4652.
- Gusarov, I., et al. (1999). "The mechanism of intrinsic transcription termination." *Molecular cell* 3: 495-504.
- Hassan, A. B., et al. (1994). "Does transcription by RNA polymerase play a direct role in the initiation of replication?" *Journal of cell science* 107 ( Pt 6): 1381-1387.
- Herrmann, R., et al. (1978). "Transposition of a DNA sequence determining kanamycin resistance into the single-stranded genome of bacteriophage fd." *Molecular & general genetics : MGG* 159(2): 171-178.
- Hirose, S., et al. (1983). "Initiation site of deoxyribonucleotide polymerization at the replication origin of the Escherichia coli chromosome." *Molecular & general genetics : MGG* 189(3): 422-431.
- Huvet, M., et al. (2007). "Human gene organization driven by the coordination of replication and transcription." *Genome research* 17(9): 1278-1285.
- Itoh, T., et al. (1980). "Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H." *Proceedings of the National Academy of Sciences of the United States of America* 77(5): 2450-2454.
- Itoh, T., et al. (1982). "Purification of ribonuclease H as a factor required for initiation of in vitro ColE1 DNA replication." *Nucleic acids research* 10(19): 5949-5965.
- Kaguni, J. M., et al. (1982). "Enzymatic replication of E. coli chromosomal origin is bidirectional." *Nature* 296(5858): 623-627.
- Keppel, F., Fayet, O., Georgopoulos, C., Ed. (1988). *The bacteriophages*, Plenum Press, New York.
- Kim, N., et al. (2007). "Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication." *DNA repair* 6(9): 1285-1296.
- Kornberg, A. (1991). "Control of initiation of the Escherichia coli chromosome." *Cold Spring Harbor symposia on quantitative biology* 56: 275-278.
- Kuznedelov, K., et al. (2002). "Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is

- not involved in the maintenance of RNA-DNA hybrid length." *EMBO J* 21(6): 1369-1378.
- Landick, R. (2009). "Transcriptional pausing without backtracking." *Proceedings of the National Academy of Sciences of the United States of America* 106(22): 8797-8798.
- Lark, K. G. (1972). "Evidence for the direct involvement of RNA in the initiation of DNA replication in *Escherichia coli* 15T." *Journal of molecular biology* 64(1): 47-60.
- Li, X. Y., et al. (1998). "Characterization of the closed complex intermediate formed during transcription initiation by *Escherichia coli* RNA polymerase." *The Journal of biological chemistry* 273(36): 23549-23557.
- Liu, B., et al. (1995). "Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex." *Science* 267(5201): 1131-1137.
- Liu, B., et al. (1993). "The DNA replication fork can pass RNA polymerase without displacing the nascent transcript." *Nature* 366(6450): 33-39.
- Lobner-Olesen, A., et al. (1992). "Different effects of *mioC* transcription on initiation of chromosomal and minichromosomal replication in *Escherichia coli*." *Nucleic acids research* 20(12): 3029-3036.
- Lothar, H., et al. (1985). "dnaA protein-regulated transcription: effects on the in vitro replication of *Escherichia coli* minichromosomes." *The EMBO journal* 4(2): 555-560.
- Masai, H., et al. (1988). "Operon structure of *dnaT* and *dnaC* genes essential for normal and stable DNA replication of *Escherichia coli* chromosome." *The Journal of biological chemistry* 263(29): 15083-15093.
- Masai, H., et al. (1996). "Mechanisms of primer RNA synthesis and D-loop/R-loop-dependent DNA replication in *Escherichia coli*." *Biochimie* 78(11-12): 1109-1117.
- McGlynn, P., et al. (2000). "Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression." *Cell* 101(1): 35-45.
- McLean, M. J., et al. (1998). "Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes." *Journal of molecular evolution* 47(6): 691-696.
- Mekler, V., et al. (2002). "Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex." *Cell* 108: 599-614.
- Merrick, M. J. (1993). "In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N)." *Mol Microbiol* 10(5): 903-909.
- Merrikh, H., et al. (2011). "Co-directional replication-transcription conflicts lead to replication restart." *Nature* 470(7335): 554-557.
- Messer, W. (1972). "Initiation of deoxyribonucleic acid replication in *Escherichia coli* B-r: chronology of events and transcriptional control of initiation." *Journal of bacteriology* 112(1): 7-12.
- Messer, W., et al. (2001). "Bacterial replication initiator DnaA. Rules for DnaA binding and roles of DnaA in origin unwinding and helicase loading." *Biochimie* 83(1): 5-12.
- Mirkin, E. V., et al. (2006). "Transcription regulatory elements are punctuation marks for DNA replication." *Proceedings of the National Academy of Sciences of the United States of America* 103(19): 7276-7281.
- Mirkin, E. V., et al. (2005). "Mechanisms of transcription-replication collisions in bacteria." *Molecular and cellular biology* 25(3): 888-895.
- Murakami, K. S., et al. (2002). "Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex." *Science* 296(5571): 1285-1290.

- Murakami, K. S., et al. (2002). "Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution." *Science* 296(5571): 1280-1284.
- Naito, S., et al. (1984). "Escherichia coli mutants suppressing replication-defective mutations of the ColE1 plasmid." *Proceedings of the National Academy of Sciences of the United States of America* 81(2): 550-554.
- Nozaki, N., et al. (1988). "In vitro transcription of the origin region of replication of the Escherichia coli chromosome." *The Journal of biological chemistry* 263(28): 14176-14183.
- Ogawa, T., et al. (1985). "Initiation of enzymatic replication at the origin of the Escherichia coli chromosome: contributions of RNA polymerase and primase." *Proceedings of the National Academy of Sciences of the United States of America* 82(11): 3562-3566.
- Ogawa, T., et al. (1994). "Cell cycle-dependent transcription from the gid and mioC promoters of Escherichia coli." *Journal of bacteriology* 176(6): 1609-1615.
- Ohsumi, M., et al. (1978). "The isolation and characterization of an in vivo recombinant between the filamentous bacteriophage f1 and the plasmid pSC101." *Virology* 89(2): 438-449.
- Omont, N., et al. (2004). "Transcription/replication collisions cause bacterial transcription units to be longer on the leading strand of replication." *Bioinformatics* 20(16): 2719-2725.
- Ozoline, O. N., et al. (1999). "DNA bendability--a novel feature in E. coli promoter recognition." *Journal of biomolecular structure & dynamics* 16(4): 825-831.
- Ozoline, O. N., et al. (1995). "Structure of open promoter complexes with Escherichia coli RNA polymerase as revealed by the DNase I footprinting technique: compilation analysis." *Nucleic Acids Res.* 23: 4533-4541.
- Park, C. S., et al. (1982). "Molecular mechanism of promoter selection in gene transcription. II. Kinetic evidence for promoter search by a one-dimensional diffusion of RNA polymerase molecule along the DNA template." *The Journal of biological chemistry* 257(12): 6950-6956.
- Park, J. S., et al. (2002). "E. coli Transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation." *Cell* 109(6): 757-767.
- Pomerantz, R. T., et al. (2008). "The replisome uses mRNA as a primer after colliding with RNA polymerase." *Nature* 456(7223): 762-766.
- Pomerantz, R. T., et al. (2010). "Direct restart of a replication fork stalled by a head-on RNA polymerase." *Science* 327(5965): 590-592.
- Prado, F., et al. (2005). "Impairment of replication fork progression mediates RNA polII transcription-associated recombination." *The EMBO journal* 24(6): 1267-1276.
- Price, M. N., et al. (2005). "Interruptions in gene expression drive highly expressed operons to the leading strand of DNA replication." *Nucleic acids research* 33(10): 3224-3234.
- Rasmussen, K. V., et al. (1983). "Conditional change of DNA replication control in an RNA polymerase mutant of Escherichia coli." *Journal of bacteriology* 154(1): 443-451.
- Reppas, N. B., et al. (2006). "The transition between transcriptional initiation and elongation in E. coli is highly variable and often rate limiting." *Molecular cell* 24(5): 747-757.
- Rocha, E. P., et al. (2003). "Essentiality, not expressiveness, drives gene-strand bias in bacteria." *Nature genetics* 34(4): 377-378.
- Rowen, L., et al. (1978). "Primase, the dnaG protein of Escherichia coli. An enzyme which starts DNA chains." *The Journal of biological chemistry* 253(3): 758-764.

- Sakata-Sogawa, K., et al. (2004). "RNA polymerase can track a DNA groove during promoter search." *Proceedings of the National Academy of Sciences of the United States of America* 101(41): 14731-14735.
- Santangelo, T. J., et al. (2004). "Forward translocation is the natural pathway of RNA release at an intrinsic terminator." *Molecular cell* 14(1): 117-126.
- Schauzu, M. A., et al. (1987). "Transcripts within the replication origin, oriC, of *Escherichia coli*." *Nucleic acids research* 15(6): 2479-2497.
- Schekman, R., et al. (1972). "Initiation of DNA synthesis: synthesis of phiX174 replicative form requires RNA synthesis resistant to rifampicin." *Proceedings of the National Academy of Sciences of the United States of America* 69(9): 2691-2695.
- Selzer, G., et al. (1982). "Specific cleavage of the p15A primer precursor by ribonuclease H at the origin of DNA replication." *Proceedings of the National Academy of Sciences of the United States of America* 79(23): 7082-7086.
- Srivatsan, A., et al. (2010). "Co-orientation of replication and transcription preserves genome integrity." *PLoS genetics* 6(1): e1000810.
- Stuitje, A. R., et al. (1986). "Dissection of promoter sequences involved in transcriptional activation of the *Escherichia coli* replication origin." *Nucleic acids research* 14(5): 2333-2344.
- Su'etsugu, M., et al. (2003). "Transcriptional control for initiation of chromosomal replication in *Escherichia coli*: fluctuation of the level of origin transcription ensures timely initiation." *Genes to cells : devoted to molecular & cellular mechanisms* 8(9): 731-745.
- Tamm, J., et al. (1983). "Structural analysis of RNA molecules involved in plasmid copy number control." *Nucleic acids research* 11(18): 6381-6397.
- Tanaka, M., et al. (1983). "A novel type of *E. coli* mutants with increased chromosomal copy number." *Molecular & general genetics : MGG* 192(1-2): 51-60.
- Theisen, P. W., et al. (1993). "Correlation of gene transcription with the time of initiation of chromosome replication in *Escherichia coli*." *Molecular microbiology* 10(3): 575-584.
- Travers, A. A. (2004). "The structural basis of DNA flexibility." *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 362(1820): 1423-1438.
- Vassilyev, D. G., et al. (2002). "Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution." *Nature* 417(6890): 712-719.
- Vilette, D., et al. (1996). "Transcription-induced deletions in plasmid vectors: M13 DNA replication as a source of instability." *Molecular & general genetics : MGG* 252(4): 398-403.
- von Hippel, P. H. (2007). "From "simple" DNA-protein interactions to the macromolecular machines of gene expression." *Annual review of biophysics and biomolecular structure* 36: 79-105.
- von Hippel, P. H., et al. (1984). "Protein-nucleic acid interactions in transcription: A molecular analysis." *Ann. Rev. Biochem.* 53: 389-446.
- Wang, J. D., et al. (2007). "Genome-wide coorientation of replication and transcription reduces adverse effects on replication in *Bacillus subtilis*." *Proceedings of the National Academy of Sciences of the United States of America* 104(13): 5608-5613.
- Wang, M. D., et al. (1998). "Force and velocity measured for single molecules of RNA polymerase." *Science* 282: 902-906.

- Wanrooij, S., et al. (2008). "Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro." *Proceedings of the National Academy of Sciences of the United States of America* 105(32): 11122-11127.
- Wickner, W., et al. (1972). "RNA synthesis initiates in vitro conversion of M13 DNA to its replicative form." *Proceedings of the National Academy of Sciences of the United States of America* 69(4): 965-969.
- Zenkin, N., et al. (2006). "The mechanism of DNA replication primer synthesis by RNA polymerase." *Nature* 439(7076): 617-620.
- Zenkin, N., et al. (2008). "RNA polymerase -the third class of primases." *Cell Mol Life Sci* 65(15): 2280-2288.
- Zinder, N. D., et al. (1985). "Multiregulatory element of filamentous bacteriophages." *Microbiological reviews* 49(2): 101-106.





*Edited by Jelena Kušić-Tišma*

DNA replication, the process of copying one double stranded DNA molecule to produce two identical copies, is at the heart of cell proliferation. This book highlights new insights into the replication process in eukaryotes, from the assembly of pre-replication complex and features of DNA replication origins, through polymerization mechanisms, to propagation of epigenetic states. It also covers cell cycle control of replication initiation and includes the latest on mechanisms of replication in prokaryotes. The association between genome replication and transcription is also addressed. We hope that readers will find this book interesting, helpful and inspiring.

Photo by agsandrew / iStock

**IntechOpen**

