STRUCTURAL BASIS FOR OLIGOMERIZATION OF THE EUKARYOTIC

REPLICATION FACTOR MCM10

By

Wenyue Du

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biological Sciences

December, 2013

Nashville, Tennessee

Approved:

Brandt F. Eichman

Katherine Friedman

Melanie Ohi

Charles Singleton

Martin Egli

To my dearest Mom and Dad, and my beloved husband Yi, for their unconditional

love and support

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my mentor and Ph.D. advisor, Dr. Brandt Eichman, for his enormous support and guidance. You are not only a role model in how to do great science, but are also always there to encourage and lead me through the twists and turns of the scientific journey. Besides scientific mentorship, you are also very supportive for my overall professional development. I consider working with you to be a great privilege and I am grateful for the opportunity.

I would also like to thank the members of my thesis committee for their advice and guidance: Drs. Kathy Friedman, Melanie Ohi, Martin Egli, Charles Singleton, and past member Dr. Dan Kaplan. The project would not have been proceeding so well without your input. In particular, Dr. Kathy Friedman and Margaret Platts in her lab showed me everything about yeast-2-hybrid and helped me get off the ground. Dr. Melanie Ohi gave invaluable comments on the analytical ultracentrifugation. Drs. Martin Egli, Charles Singleton and Dan Kaplan always offered valuable input and thoughtful suggestions.

I am also grateful to all Eichman lab members, past and present, for their technical assistance, informal discussions, and friendship over the years. I also owe thanks to all the brilliant and kind people that I have worked with on my project, either formally or informally, many of whom have offered technical assistance on unfamiliar instrumentations, or guided me on data analysis software. I am eternally thankful to them and wish them the best of luck in their

future endeavors. This work would also not have been possible with the financial support from NIH Grant (R01 GM080570) and Vanderbilt Department of Biological Sciences.

These five years is a great time for me, which could not have been possible without the many good friends I have here at Vandy. We shared laughter and great times, and you guys were always there for me. For those friends who are also far away from home, each one of us is great to live our dreams and accomplish our goals.

Last but certainly not least, I could not have been here without the enduring love and support from my dear Mom and Dad, who give their daughter all she ever wanted and the freedom to live her own dream. Yi, my beloved husband, you are an eternal inspiration and I am delighted to share my life with you.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	L104D/L108D
4A	L104A/L108A/M115A/M118A
aa	amino acids
AAA+	ATPases associated with diverse cellular activities
AD	activation domain
ATP	adenosine triphosphate
AUC	analytical ultracentrifugation
BD	binding domain
BME	β-mercaptoethanol
bp	base pair
CC	coiled-coil
CCCC	cysteine- cysteine- cysteine- cysteine
CCCH	cysteine- cysteine- cysteine- histidine
Cdc	cell division cycle
CDK	cyclin-dependent kinase
CMG	Cdc45/Mcm2-7/GINS
CTD	C-terminal domain
DDK	Dbf4-dependent kinase
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA
DUE	DNA-unwinding element

EM	electron microscopy
EMSA	electrophoretic mobility shift assays
EV	empty vector
FA	Fanconi anemia
G	gram
G1 phase	Gap 1 phase
GINS	Go-Ichi-Nii-San
h	hour
HA	hemagglutinin
HRP	horseradish peroxidase
HU	hydroxyurea
ID	internal domain
kDa	kilo Dalton
L	liter
LB	Luria broth
LTag	large T-antigen
M	Molar
MAR	matrix attachment region
MBP	maltose binding protein
Mcm/MCM	Minichromosome maintenance protein
Mcm10ΔN	Mcm10$^{230-860}$
mg	milligram
min	minute

mL	milliliter
mM	millimolar
µg	microgram
µM	micromolar
Mw	Molecular weight
NMR	Nuclear magnetic resonance
nt	nucleotide
NTA	nitrilotriacetic
NTD	N-terminal domain
OB	oligonucleotide/ oligosaccharide binding
ORC	origin recognition complex
PCNA	proliferating cell nuclear antigen
PDB	protein databank
PEG	polyethylene glycol
PIP	PCNA interacting peptide
pol α	polymerase α
pre-IC	pre-initiation complex
pre-RC	pre-replicative complex
RCR	rolling-circle replication
RDR	recombination-dependent DNA replication
RFC	replication factor C
RPA	replication protein A
rpm	rotations per minute

S phase	synthesis phase
SAXS	small angle X-ray scattering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	second
SEC-MALS	size exclusion chromatography and multi-angle light scattering
SSB	single-strand DNA binding protein
ssDNA	single-stranded DNA
SV40	simian virus 40
T-antigen	tumor antigen
TCA	trichloroacetic acid
TCEP	tris(2-carboxyethyl)phosphine hydrochloride
TLS	translation/libration/screw-rotation
WT	wildtype
ZnF	zinc-finger

CHAPTER I

INTRODUCTION

DNA replication is the fundamental process for all living organisms to duplicate their genomes. From prokaryotes to eukaryotes, DNA replication is central to cell proliferation, and adopts a semiconservative mode to maintain high degree of conservation of the genetic material. Precise coordination and regulation mechanisms exist to guarantee the faithful transmission of the parental DNA information to daughter cells. This dissertation will focus on the structural and biochemical characterization of the oligomeric state of one essential protein in DNA replication, the minichromosome maintenance protein 10 (Mcm10). This work may provide a better understanding of the mechanism of DNA replication and replication-related diseases, including cancer.

Cell cycle

DNA replication occurs at a precise time in the cell cycle. The cell cycle in prokaryotic cells is achieved by the binary fission process. In eukaryotes, the cell cycle consists of interphase (G1, S, and G2 phases), and M phase (mitosis, cell division). G0 phase refers to the state of quiescence when cells have temporarily or reversibly stopped dividing. Checkpoints exist to ensure temporal accuracy, verify whether each phase of the cell cycle has been accurately completed before progression into the next phase and ensure that the cell does not continue

in the cell cycle in the presence of DNA damage or inappropriately attached kinetochores (Figure 1). G1 (Gap 1) phase is the growth phase when all biosynthetic activities of the cell reach a high rate, although there are differences in cell cycle durations across species and cell types. For example, some organisms such as *S. pombe*, have essentially no G1 phase and do all of their growing in G2 phase (Oliva 2005). Also, early embryonic cell cycles often omit G1 and G2 phases to quickly proceed through rounds of S phase and mitosis. Proteins and enzymes are formed in G1 phase to prepare for DNA replication (Garrett 2013). The G1 checkpoint control mechanism ensures accuracy for entry to the next phase, S phase. DNA replication happens during S (synthesis) phase. A single chromosome is replicated into two identical copies called sister chromatids. This process requires coordination of multiple protein players to form the replisome. G2 phase is the gap between DNA synthesis and mitosis. The cells continue to grow and G2 checkpoint ensures the cells are ready to enter M phase. During mitosis, eukaryotic cells separate the duplicated chromosomes in two identical sets into two daughter cells: the sister chromatids separate and each chromatid is linked to only one pole via kinetochore microtubules; the individual chromatids are then drawn to the poles as the kinetochore microtubules shorten. The metaphase checkpoint in the middle of mitosis ensures that the cell has achieved correct bipolar attachment of all chromosomes.
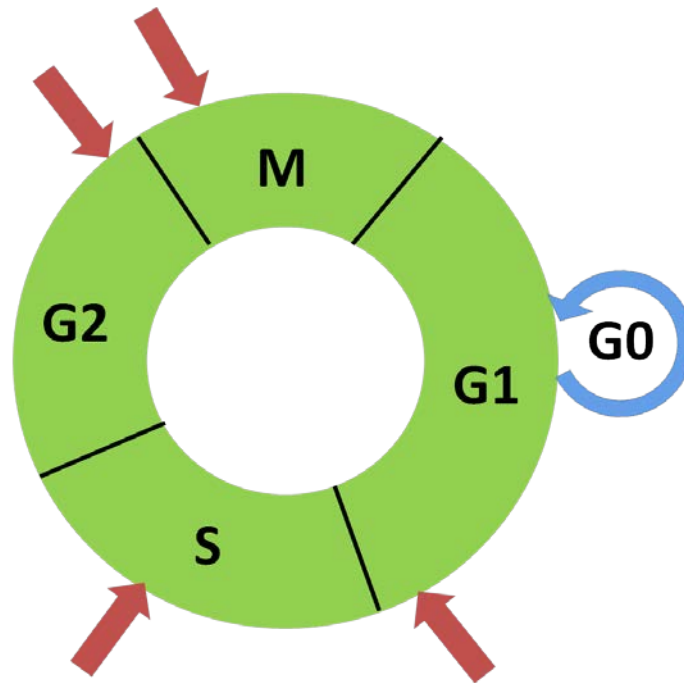
Figure 1. Cell cycle schematic. The red arrows indicate checkpoints. (Figure adapted from personal website of Eishi Noguchi, http://eishinoguchi.com/checkpoint.htm)

DNA replication

Overview

The basic mechanism of DNA replication is conserved from prokaryotes to eukaryotes. The double-stranded DNA is separated into single-strands, and each strand serves as a template for the complementary strand to be synthesized based on base pair matching through hydrogen bonding (A-T and C-G). DNA replication starts at specific locations on the chromosome, known as origins of replication. When dsDNA unwinds, two replication forks are formed, with multiple proteins associated to help with replication (Figure 2). The dsDNA is unwound by the helicase into two complementary single strands. Exposed ssDNA is stabilized by single strand binding proteins. DNA synthesis cannot start *de novo*. Primase

synthesizes short 8-12 nt RNA primers, which are then extended by polymerase α to about 30 nt to serve as the initiation primer for new DNA synthesis. A single priming event is sufficient for the leading strand, while multiple priming events are needed on the lagging strand. DNA synthesis always occurs in the 5' to 3' direction, which is continuous on the leading strand, and by short segments (1,000-2,000 nucleotides long in *E. coli*, 100-200 nucleotides long in eukaryotes) on the lagging strand, known as Okazaki fragments. Positive supercoiling tension is built up ahead of the helicase and needs to be released by topoisomerase. The nicks between the Okazaki fragments are then sealed by DNA ligase. The initial RNA primer is removed by nucleases such as FEN1, RNase I and Dna2.
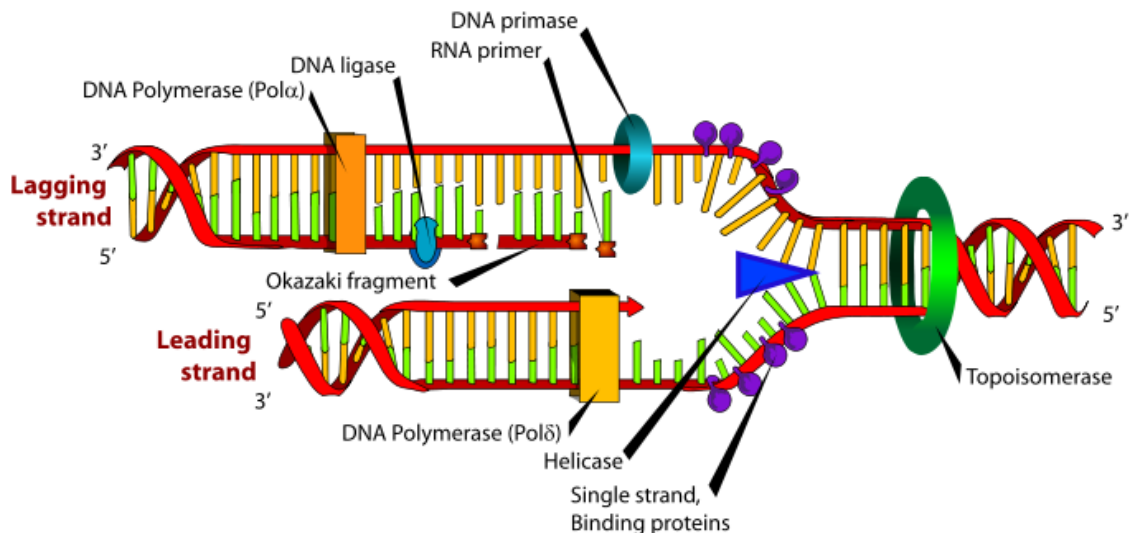
Figure 2. DNA replication fork (Figure adapted from Garrett and Grisham 2013). The primer and Okazaki fragments are not to scale.

DNA replication is divided into three stages: initiation, elongation, and termination. Initiation refers to the process in which a cell recruits and assembles

multiple proteins into a macromolecular machine, the replisome. Elongation is the stage during which bulk DNA synthesis occurs. The proteins mentioned above are actively involved during elongation. When replication forks travelling in opposite directions meet one another, indicating that all DNA in between the origins of these forks has been successfully copied, the elongation process is terminated. These three fundamental steps apply to all forms of life, with variations and complexity in different forms. In this chapter, the mechanisms of DNA replication in all three kingdoms of life will be discussed individually. One advantage of studying bacterial, phage/viral, and archaeal replication is the fact that the enzymes can be purified for *in vitro* studies. Also, lower organisms usually require simpler protein machineries. For example, some of the protein complexes that are heteromeric in eukaryotes are homomeric in prokaryotes.

Prokaryotic Replication

In the bacterial chromosome, there is typically one single origin of replication, *oriC*. In *Escherichia coli*, the most thoroughly studied model organism for bacterial DNA replication, *oriC* is approximately 250 base pairs in length and contains multiple sequence-specific binding sites for the initiator protein DnaA, named DnaA boxes (reviewed in Mott 2007). DnaA binds to the DnaA box consensus sequence 5'-TTATCCAC-3' with high affinity (Schaper 1995). In the presence of ATP, DnaA also binds to "I sites", which are found interspersed among the DnaA boxes. The "I sites" differ subtly from the DnaA box consensus element. DnaA also binds to the "ATP-DnaA boxes" in the presence of ATP,

which lie within an AT-rich DNA-unwinding element (DUE) composed of three 13-mer repeats. At the onset of DNA replication initiation, DnaA binds with *oriC*. As DnaA localizes to *oriC*, it homo-oligomerizes to form a large nucleoprotein complex, which is visible by electron microscopy (Funnell 1987). DnaA assembly, along with negatively supercoiled DNA and nucleoid architectural factors, facilitate DNA melting which generates single-stranded DNA, followed by helicase loading and further assembly of the replisomal machinery. Inactivation of DnaA is achieved by stimulation of ATP hydrolysis upon replisome assembly. The lack of ATP induces rearrangements in the AAA$^+$-specific motifs in DnaA, resulting in the dissociation of the oligomeric complex (reviewed in Mott and Berger 2007).

The helicase DnaB is recruited to the melted DNA in the form of a hexamer, with the help of six DnaC proteins (Caruthers 2002). Single-strand DNA binding protein (SSB) binds to the exposed ssDNA. SSB in bacteria is a homotetramer. Each monomer contacts DNA through an oligonucleotide/oligosaccharide (OB) fold and achieves protein-protein interaction through an acidic C-terminal domain (Raghunathan 1997; Raghunathan 2000). Primase (DnaG) then primes the unwound DNA as template for DNA synthesis.

During elongation, DNA polymerase III is loaded onto primed DNA (Corn 2006). The catalytic mechanism of DNA polymerase III involves the use of two metal ions in the active site. While new DNA synthesis is continuous on the leading strand, on the lagging strand, many RNA primers are needed, and many short fragments known as Okazaki fragments are synthesized. These RNA

6

primers on Okazaki fragments are then degraded by RNase H and DNA polymerase I. RNase H is a non-specific endonuclease and catalyzes the cleavage of RNA via a hydrolytic mechanism. It cleaves the 3'-O-P bond of RNA in a DNA/RNA duplex to produce 3'-hydroxyl and 5'-phosphate terminated products. The nicks between Okazaki fragments are fixed by ligase. As the helicase unwinds the DNA, topoisomerase/gyrase is needed to break the DNA to relieve the positive supercoiling tension and then re-anneals the lesion (Cozzarelli 1980).

In *E. coli*, DNA replication is terminated when the two opposing replication forks physically collide 180 degrees opposite from the origin on the circular chromosome (Dalgaard 2009). When one fork is slowed or stopped and the replication fork from the right hemisphere starts to replicate the left hemisphere in the wrong direction, the fork collides with the Tus protein, which binds to the *cis*-acting Ter elements on the chromosome. However, the *tus* gene is not conserved among bacteria, and *tus* mutants lack measurable phenotypes. The rationale for the large number of Ter copies and their distribution over such a large portion of the chromosome is poorly understood. Hendrickson and Lawrence used bioinformatics techniques and found that in both *E. coli* and *Bacillus subtilis*, changes in mutational bias patterns indicated that replication termination most likely occurs at or near the *dif* site, located 180 degrees from the origin. They suggested that the Ter site might participate in halting replication forks originating from DNA repair events, and not those originating at the chromosomal origin of replication (Hendrickson 2007).

Bacteriophage Replication

Bacteriophages (prokaryotic viruses) provide a unique and useful model system for the study of DNA replication (reviewed in Weigel 2006). Genes encoding replication functions tend to be located close to each other in many phage genomes, known as "replication modules". There are virulent phages that are only lytic (e.g., T4 phage), and temperate phages that can be both lytic and lysogenic (e.g., phage lambda of *E.coli*). In the lytic cycle, the host (bacterial cells) are lysed and destroyed immediately after phage replication, whereas the lysogenic cycle allows the host and phage to co-exist initially through integration or circularization of the viral genome until environmental factors trigger the lytic cycle, such as the depletion of nutrients.

Unlike prokaryotic and eukaryotic replication, phage replication can utilize multiple mechanisms to prime replication depending on the virus and the situation. There are four mechanisms known: 1) the 3'-OH end of nicked dsDNA in the rolling-circle replication (RCR) mechanism; 2) the 3'-OH end of a short RNA primer synthesized by primase in the theta(θ)-type DNA replication; 3) the transcripts that remain bound to the templates; and 4) the 3'-end of ssDNA invading a duplex in recombination-dependent DNA replication (RDR) (reviewed in Weigel and Seitz 2006).

Bacteriophage T7, an obligate lytic phage that infects *E. coli* (Dunn 1983; Studier 1983), is a model system to study bacteriophage replication. In fact, since there are only a few proteins involved in bacteriophage T7 replication, it is a convenient model to study the dynamics of a minimal replication fork (Doublie

1998). The important proteins involved in T7 viral DNA synthesis include a hexameric T7 primase-helicase that unwinds and primes the DNA, a T7 single-stranded DNA binding protein that binds unwound ssDNA, an 80 kDa T7 DNA polymerase, and *E. coli* thioredoxin as a processivity factor that prevents polymerase from dissociating from the DNA template. The polymerase structure resembles an open right hand, with mostly alpha helices important in nucleotide recognition. The thumb domain binds to thioredoxin, with catalytic activity contributed by the fingers domain and DNA cradled by the palm domain. The 3'-5' exonuclease domain is important for the proofreading function of the polymerase. The crystal structure of T7 DNA polymerase bound to thioredoxin, a primer template, and a nucleotide triphosphate has been determined. The complex showed an unprecedented closed conformation of the polymerase active site that allows catalytically important, conserved residues to engage the incoming nucleotide. Two aspartate residues that are strictly conserved in the Polymerase I family, and that have homologues in other DNA polymerases and in RNA polymerases, ligate two metal ions in the polymerase active site in positions that are consistent with a two-metal mechanism of phosphoryl transfer (Doublie 1998).

Eukaryotic Replication

Eukaryotes have a much more complex DNA replication process and machinery compared to prokaryotes. Complex arrays of proteins are assembled and reorganized. Sophisticated regulatory mechanisms exist to coordinate the

initiation, elongation, and termination processes of DNA replication. Eukaryotic replication also requires initiation at multiple origins to duplicate the large genomes. DNA replication in higher eukaryotes has high plasticity. Fertilized eggs of amphibians replicate the entire genome within a matter of 20-30 min, whereas the genome replication in somatic cells takes 8-10 hours. The former requires very frequent initiation through utilization of virtually all the potential origins, while the latter only has spatially and temporally regulated origin firing (reviewed in Masai 2010).

SV40 Replication

   Simian virus 40 (SV40), a polyomavirus that is found in both monkeys and humans, is a well-studied model system for eukaryotic DNA replication. The SV40 genome is a circular duplex DNA with a single replicon. SV40 is a simple but powerful model to study eukaryotic DNA replication due to its association with host histones in nucleosomes and its dependence on the host cell milieu for replication factors and precursors. SV40 DNA replication requires only one viral protein, the large tumor antigen (T-antigen), and all other components originate from the cell. The T-antigen functions as an initiator protein that binds to the origin of replication, and also as a helicase that unwinds duplex DNA (reviewed in Waga 1998). Replication protein A (RPA) binds to unwound single-stranded DNA. The T-antigen-RPA complex binds DNA polymerase α-primase, which synthesizes a short primer at the replication origin (Bullock 1989; Matsumoto 1990; Bullock 1991; Murakami 1992). Further processive replication is achieved

by proliferating cell nuclear antigen (PCNA), its loader replication factor C (RFC), and polymerase δ. Despite the conservation, there are several major differences between the SV40 and host DNA replication mechanisms (reviewed in Fanning 2009). SV40 encodes its own DNA helicase, the large tumor antigen (Stahl 1986); while chromosomal replication employs the Cdc45/Mcm2-7/GINS helicase (Moyer 2006). SV40 replication does not use DNA polymerase epsilon (Zlotkin 1996; Pospiech 1999), which is important for chromosomal replication (Seki 2006; Shikata 2006). Also, SV40 DNA replicates during the S/G2 phase, but CDK-phosphorylated DNA polymerase alpha-primase cannot support viral DNA replication in a cell-free reaction (Ott 2002), suggesting the SV40 DNA replication requires the cellular environment and protein factors from the host cell. Importantly, SV40 and polyomavirus infection induce DNA damage signaling that promotes viral DNA replication (Wu 2004; Dahl 2005; Shi 2005; Zhao 2008), but ordinarily inhibits chromosomal replication and cell cycle transitions (Cimprich 2008; Lavin 2008; Fanning and Zhao 2009).

Initiation

Eukaryotes define the site of initiation on the genome through binding of the origin recognition complex (ORC). There is extensive variation in terms of sequence preference for origins in different organisms (reviewed in Masai 2010). Budding yeast ORC specially recognizes the 11-bp or 17-bp conserved sequence, the ARS consensus sequence (Palzkill 1988; Marahrens 1992; Theis 1997). Fission yeast ORC and *Drosophila* ORC preferably bind to AT-rich

sequence (Chuang 1999). In mammalian cells, the ORC binding sequences usually are AT-rich (Altman 2004; Paixao 2004; Wang 2004), with dinucleotide repeats (Altman and Fanning 2004), asymmetrical purine-pyrimidine sequences (Wang 2004), and matrix attachment region (MAR) sequences (Debatisse 2004; Schaarschmidt 2004), but it has been impossible to define specific required sequence elements.

Potential sites of replication initiation are defined through ORC binding. Once origins are identified, a ring-shaped complex of six minichromosome maintenance proteins, Mcm2-7, are loaded onto the chromosome with the help of Cdc6 and Cdt1 (reviewed in Blow 2005): events that define the formation of pre-RC (pre-replication complex). Current models posit that there are excess numbers of pre-RCs on the genome as potential origins, but only some of them are actually used for initiation (Edwards 2002; Hyrien 2003). The mechanism through which certain origins are activated is still unknown, but it could be caused by the effects of transcriptional units (Snyder 1988; Haase 1994; Saha 2004), topology (Dayn 1992), the availability of single-stranded thymine-rich DNA (You 2003), and origin interference effects from adjacent loci (Brewer 1993; Tadokoro 2002; Lebofsky 2006). Studies using DNA fiber autoradiography and fluorescence microscopy showed that fast fork rates are associated with less frequent origin initiation, and slow fork rates are associated with more frequent initiation (Taylor 1977; Anglana 2003; Gilbert 2007).

Mcm2-7 forms a heterohexameric complex, which is the core of the DNA replication helicase. The loading of Mcm2-7 is ATP-dependent and requires

ORC, Cdc6 and Cdt1. Remus *et al* used electron microscopy to show Cdt1•Mcm2-7 form a single heptamer. Mcm2-7 double hexamers are loaded head-to-head and connected via their N-terminal rings. Mcm2-7 can slide passively along double-stranded DNA, which runs through the channel in the center (Remus 2009). The eukaryotic Mcm2-7 double hexamer structure is similar to that of the archaeal helicase MCM, which is thought to be the evolutionary precursor (Fletcher 2003; Liu 2009). However, the archaeal complex consists of six copies of a single protein, while the Mcm2-7 complex is a heterohexamer with six distinct, but related subunits.

Eukaryotic DNA replication initiation is a precisely regulated event that requires the ordered assembly of multiple protein complexes at replication origins. DNA replication needs to occur only once per cell cycle. The origin is licensed once Mcm2-7 is loaded onto DNA. Cells regulate and prevent origin re-licensing (licensed twice in one cell cycle) in yeast through control of cyclin levels (reviewed in Blow and Dutta 2005), including rapid CDK-dependent destruction of the Cdc6/Cdc18 protein during the S phase (reviewed in Masai 2010). In metazoa, re-licensing is mainly prevented through regulation of Cdt1 activity by degradation of its specific inhibitor, geminin (Blow and Dutta 2005; Arias 2007). Another approach may be the degradation of Orc1 by ubiquitin-dependent proteolysis (Mendez 2002; Ohta 2003) and inhibition of re-association of the MCM complex to chromatin (reviewed in Masai 2010).

Cdt1 and geminin help coordinate replisome assembly with the cell cycle. Cdt1 is loaded onto chromatin with the help of ORC and Cdc6 (Bell 2002; Fujita

13

2006). Geminin prevents the loading of Mcm2-7 onto DNA (re-licensing) by binding to Cdt1 (McGarry 1998; Wohlschlegel 2000; Tada 2001; Yanagi 2002; Lee 2004; Saxena 2005). The protein levels of Cdt1 and geminin oscillate throughout the cell cycle. Cdt1 level is high in G1 phase and M-G1 phase, but low in S phase (Nishitani 2001); whereas geminin is low in G1 phase, but high in S, G2, and M phases. Cdt1 is strictly eliminated during the S phase by two regulation mechanisms to prevent re-licensing. During S phase, Cdt1 is degraded by ubiquitination-dependent proteolysis in a CDK-dependent manner or through interaction with proliferating cell nuclear antigen (PCNA), with the remaining Cdt1 inhibited by geminin (Takeda 2005; Arias 2006; Fujita 2006; Nishitani 2006; Senga 2006; Xouri 2007; Xouri 2007).

At the onset of S phase, Mcm2-7 is activated through two important phosphorylation events by cell-cycle dependent kinases: cyclin-dependent kinase (CDK) and Dbf4-dependent kinase (DDK) (Stillman 2005; Sclafani 2007). In yeast, CDK phosphorylates Sld2 and Sld3, and facilitates their binding to Dpb11 (Tanaka 2007; Zegerman 2007), whereas DDK directly phosphorylates Mcm2 and Mcm4 (Lei 1997; Sheu 2006). In *S. pombe*, Mcm10 interacts with both the Mcm2-7 complex and Dfb1-Hsk1 (homolog of Dbf4-Cdc7 in *S. cerevisiae*). The N-terminus of Mcm10 was found to interact directly with Dfp1-Hsk1 and was essential for phosphorylation of the Mcm complex. Truncated derivatives of Mcm10 that complemented the temperature-sensitive phenotype of *Mcm10* mutant cells also stimulated the phosphorylation of the Mcm2-7 complex (Lee 2003). These phosphorylation events may induce a conformational change to

14

allow for the association of other critical factors to activate a latent helicase activity, or allow for an isomerization of the pre-assembled complexes to achieve such a switch (Ilves 2010). Research has also shown that *Drosophila* Mcm2-7 helicase is activated in complex with Cdc45 and the four GINS proteins (Sld5, Psf1, Psf2, and Psf3) (Bochman 2008; Ilves 2010), forming the CMG complex (Wohlschlegel 2002; Moyer 2006; Pacek 2006). Assembly of the CMG complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins (Im 2009). RecQL4 is a homolog of Sld2, and is required for DNA replication also in *Xenopus* and *Drosophila* (Sangrithi 2005; Wu 2008). Ctf4/And-1 is required for sister chromatid cohesion in yeasts (Hanna 2001), and is also essential for the chromatin binding of DNA polymerase α and for DNA replication initiation in mammalian cells (Zhu 2007). Mcm10 loads at the onset of S phase, and therefore serves an important role to subsequently load other proteins in the process.

Cdc45 and GINS are not only required for the initiation, but also the elongation stage of DNA replication (Aparicio 1997; Tercero 2000; Kanemaki 2003; Kubota 2003; Takayama 2003; Pacek 2004; Aparicio 2006; Labib 2007). The GINS proteins bind specifically to Mcm4, and enhance DNA substrate affinity. DNA unwinding and onset of replication may occur through allosteric changes in Mcm2-7, which is contributed by the association of Cdc45 and GINS (Ilves 2010). Interestingly, depletion of a Dpb11 homolog in human, TopBP1, did not significantly affect CMG complex formation, suggesting that the initiation complex assembly may be different in human cells from in yeast cells (Im 2009).

During this initiation stage, origin DNA is melted, i.e. double-stranded DNA is separated into two single strands. The exact mechanism and timing is not known, but the MCM is thought to remodel to surround the single strand DNA. DNA polymerase α/primase synthesizes a short RNA to serve as a template for DNA replication. Pol α is composed of four essential subunits, named based on their molecular weight: p48, p58, p68, and p180 (Sugino 1995). p48 and p58 are the primase to synthesize short 7-10 nt RNA primers (Arezi 2000). p68 does not have enzymatic activity and is thought to have a regulatory role (Mizuno 1998). The p180 subunit has the catalytic DNA polymerase activity and extends the primer to ~30 nt (Conaway 1982). Pol α has limited processivity and lacks 3' exonuclease activity for proofreading errors (reviewed in Burgers 1998). PCNA, the ring-shaped processive clamp, and its clamp loader RFC are then loaded onto DNA to help engage the processive polymerases (Garg 2005). Pol δ and pol ε, the two processive polymerases, are then loaded to begin bulk DNA synthesis (Burgers 2009). Leading strand synthesis is made mainly through pol δ, whereas the lagging strand requires pol ε (McElhinny 2008; Pursell 2007). This specificity was determined using a mutant polymerase allele whose error rate is higher for one mismatch than for its complement, and the strand-specific mutation rates strongly depended on the orientation of a reporter gene relative to an adjacent replication origin. At this stage, the replisome prepares the DNA to enter the elongation stage for bulk DNA synthesis. This replication initiation process is also reviewed in Chapter II with a schematic figure (Figure 4) highlighting the steps in replisome assembly.

Elongation

DNA replication elongation refers to the process during which the complementary DNA strand to the template is synthesized. Multiple origins on the chromosome will begin the bulk DNA synthesis to ensure the job is done in a timely fashion. As described earlier, the lagging strand is synthesized as a series of Okazaki fragments. As the synthesis of each Okazaki fragment is completed, the RNA primer of the previous fragment must be removed. The RNA primers are excised by RNase H1 and then FEN1 in the more common short flap pathway. FEN1 removes the small 5′-flap generated by strand displacement synthesis; this flap is most often only a mononucleotide in size. If additional initiator RNA is present, Pol δ and FEN1 go through iterative cycles of strand displacement and small flap cutting, a process called nick translation, until all RNA has been degraded. However, FEN1 shows no activity on substrates with long 5′-flaps when the flap is coated with RPA or structured, *e.g.* folded in a hairpin. In this long flap pathway, RNA removal is initiated by the Dna2 nuclease/helicase. DNA ligase then seals the nicks together (reviewed in Burgers 2009). Another important thing is that the supercoil tension generated during DNA synthesis needs to be relieved by topoisomerases. Toposiomerase I cuts a single backbone on the DNA, enabling the strands to swivel around each other to remove the build-up of twists. Topoisomerase II cuts both backbones, enabling one dsDNA to pass through another, thereby removing knots and entanglements that can form within and between DNA molecules (Leppard 2005). Topoisomerase also mediates the ligation process.

Termination

Termination is the final stage of DNA replication. In most cases, termination simply occurs at the site where two replication forks meet. Replication also terminates at the end of the chromosome, the telomere. Telomeres are shortened over multiple cell divisions in somatic cells because those end regions cannot be fully replicated. In germ cells, telomerase extends the repetitive sequence of the telomere region to counteract the loss of DNA that occurs during DNA replication. Errors in telomerase activation can lead to diseases like cancer. Telomerase activation is associated with tumorigenesis because telomere maintenance is required for immortal cellular growth. At the end of DNA replication, there are two identical copies of the genome within the cell. The cell is ready to enter G2 phase and later M phase through the cell cycle.

Archaeal Replication

Archaeal DNA replication machinery is related more closely to eukaryotes than to prokaryotes. A subset of the eukaryotic initiation proteins have been identified in archaea, while several other eukaryotic initiation factors (e.g., Cdt1, Mcm10, and Cdc45) are absent (Grabowski 2003).

Most archaea contain one or two copies of ORC/Cdc6 homologs. It is unclear whether the homolog(s) function through mechanisms similar to ORC or Cdc6, or both. These homologs possess low ATPase activity that can be stimulated by origin DNA, PCNA, the MCM helicase, and subunits of polymerase (Lee 2000; Giraldo 2003). Similar to the eukaryotic proteins, the archaeal

ORC/Cdc6 homologs may also function as helicase loaders. There is at least one MCM homolog in all archaea known. Archaea also have a variety of SSB, which are more similar to eukaryotic RPA than to bacterial SSBs. Archaea also have primase, GINS complex homolog, replicative DNA polymerases, PCNA and RFC (reviewed in Barry 2006). Additional unidentified replication initiation factors may exist and need to be explored in archaea.

The table summarizes the protein machinery in DNA replication across the three kingdoms (Table 1). (reviewed in Grabowski and Kelman 2003)

Table 1. Initiation and elongation proteins in bacteria, eukarya and archaea. (adapted from Grabowski and Kelman 2003)

Initiation			
	Bacteria	Eukarya	Archaea
Origin recognition	DnaA (one subunit)	Origin recognition complex (ORC) (six subunits, Orc 1-6)	ORC/Cdc6[a] (one or two subunits)
Helicase loader	DnaC[b]/DnaI[c] (one subunit)	Cdc6 (one subunit) together with ORC	ORC/Cdc6[a] (one or two subunits)
Replicative helicase	DnaB[b]/DnaC[c] (one subunit)	Minichromosome maintenance (six subunits, Mcm2-7)	MCM (one subunit)
Elongation			
Single stranded DNA binding protein (SSB)	SSB (one subunit)	Replication protein A (RPA) (three subunits)	RPA/SSB[d] (one or three subunits)
Primase	DnaG (one subunit)	Pol α/primase complex (four subunits)	Primase homolog (two subunits)
Polymerase/exonuclease	Pol III core (three subunits)	Pol δ (three or four subunits); Pol ε (five subunits)	Pol B (one subunit); Pol D[e] (two subunits)
Clamp loader	Γ-complex (five subunits)	Replication factor C (RFC) (five subunits, Rfc 1-5)	RFC (two subunits)
Sliding clamp	β	Proliferating cell nuclear antigen (PCNA)	PCNA (one or three subunits)[f]
Removal of primers	Pol I; RNase H	Flap endonuclease I (Fen-1); RNase H	Fen-1; RNase H
Lagging strand maturation	DNA ligase (NAD-dependent)	DNA ligase I (ATP-dependent)	DNA ligase I (ATP-dependent)
Topoisomerase	Type I and II Reverse gyrase[g]	Type I and II	Type I and II Reverse gyrase[g]

a. In most species. **b.** In gram-negative bacteria. **c.** In gram-positive bacteria. **d.** Euryarchaeotal genomes contain one or three RPA homologs while *Crenarchaeota* have a single "SSB-like" protein. **e.** Only in *Euryarchaeota*. **f.** Euryarchaeotal genomes contain one PCNA homolog, and *Crenarchaeota* have three. **g.** Only in hyperthermophilic and some thermophilic organisms.

Mcm protein family

Mcm2-7

The mini-chromosome maintenance (Mcm) proteins were initially identified in genetic screens in yeast for proteins involved in plasmid maintenance, cell cycle progression, and chromosome missegregation (Maine 1984). Mcm proteins are related based on sequence similarity (Figure 3) (reviewed in Dutta 1997). Deleting any one *MCM* gene in *S. cerevisiae* and *S. pombe* causes lethality (Dutta and Bell 1997; Kelly 2000). Hypomorphic mutations in *MCM2*, *MCM3*, and *MCM5* genes cause defective plasmid maintenance (Maine 1984). Other separate screens showed *MCM4* and *MCM7* are important for cell cycle progression (Moir 1982; Hennessy 1991), while *MCM6* is necessary for chromosome segregation (Takahashi 1994). Mcm2-7 form stable sub-complexes: Mcm 2/4/6/7, Mcm4/6/7, Mcm3/7, and Mcm 3/5 (Ishimi 1997; Lee 2000; Maiorano 2000; Prokhorova 2000). Although some sub-complexes were able to associate with the chromatin independent of the rest parts in the Mcm2-7 complex, these sub-complexes are Cdt1-independent and cannot associate with chromatin in a productive manner even with the addition of other parts in Mcm complex in *Xenopus* egg extract (Bell and Dutta 2002). Therefore, all six Mcm proteins are required to load onto DNA for normal replication activity.

All Mcm proteins belong to the AAA^+ protein family (ATPases Associated with various cellular activities) and have conserved ATP binding and hydrolysis activity to provide energy for unwinding duplex DNA (Koonin 1993; Bochman 2007; Bochman and Schwacha 2008). As described above, Mcm2-7 form head-

21

to-head double hexamers during origin licensing. Cdc45 and GINS are loaded to form the Cdc45-Mcm2-7- GINS (CMG) complex, which is believed to be the active helicase to unwind dsDNA. RNA interference knock-down experiments targeting the GINS and Cdc45 components establish that the proteins are required for the S phase transition in *Drosophila* cells (Moyer 2006).
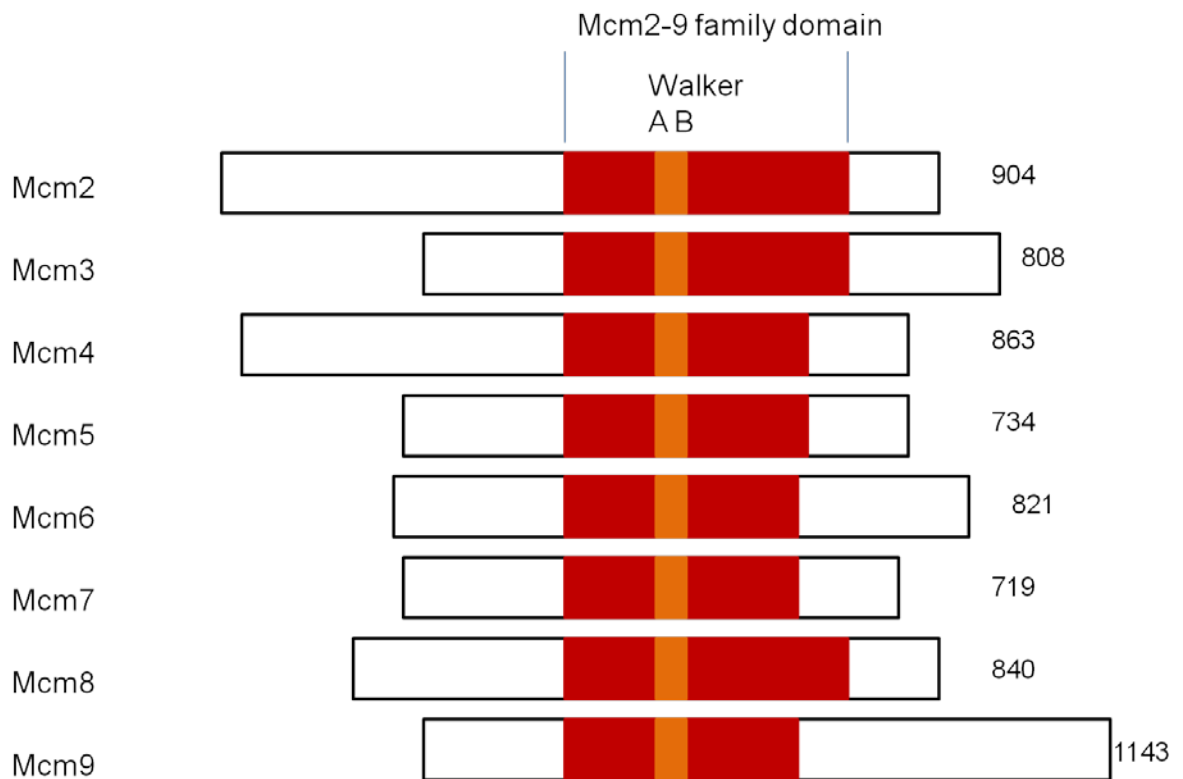

Figure 3. Schematic alignment of human Mcm2-9 proteins. The highly conserved ATPase domain is shown in red. The area corresponding to the Walker A and B motifs is shown in orange. Numbers on the left indicate protein length (adapted from Maiorano 2006).

Archaeal MCM

Archaeal MCM forms a homohexamer and functions as the replicative helicase (Tye 2000). There is high sequence homology between archaeal MCM and eukaryotic helicase Mcm2-7, all of which have ATPase domains. MCM, Mcm4, Mcm6, and Mcm7 have zinc motifs for DNA binding. Archaeal MCM was reported to be the evolutionary predecessor of eukaryotic Mcm2-9 proteins (Liu 2009). MCM unwinds DNA with a robust 3' to 5' helicase activity in the presence of ATP and Mg^{2+}, possibly by oscillating between an open and a closed ring to allow ssDNA loading (Bochman 2009).

The crystal structure of N-terminal *M. thermautotrophicus* MCM (MthMCM) showed a head-to-head double hexamer, which is slightly different from the eukaryotic Mcm2-7 double hexamer, in that MthMCM is connected primarily by interactions near the inner circumference of the N-terminal rings (Fletcher 2003; Remus 2009). The structure of full-length MCM complexes is needed to further understand its function and mechanism in archaeal DNA replication.

Mcm8 and Mcm9

Mcm8 was identified in human in a screen to isolate cancer-related genes associated with Hepatitis B integration sites in hepatocellular carcinomas (Gozuacik 2003). Mcm9 was identified using BLAST searches for protein sequences similar to Mcm2-8 (Lutzmann 2005). Mcm 8 and Mcm9 have now been found in human and mouse. Mcm8 localizes to the nucleus, possibly through binding to a nuclear protein (Gozuacik 2003). Mcm8 was later shown to

co-localize with RPA at replication foci, and it possesses ATP hydrolysis and *in vitro* helicase activity. Maiorano *et al* have further suggested that Mcm8, instead of functioning during initiation, functions in the elongation step of DNA replication as a helicase by facilitating the recruitment of RPA34 and stimulates the processivity of DNA polymerases at replication foci (Maiorano 2005). However, other studies reported the contrary, that Mcm8 has a role in pre-RC formation during replication initiation. These studies showed that Mcm8 interacts with ORC and Cdc6 in G1 phase, and depletion of Mcm8 by RNAi slowed S phase and reduced Mcm2-7 loading onto chromatin (Volkening 2005; Kinoshita 2008). Mcm9 was shown to bind to chromatin in an ORC-dependent manner, and is required for Mcm2-7 loading, as well as forms a complex with Cdct1. The authors then proposed that Mcm9 has a regulatory role in Mcm2-7 loading (Lutzmann 2008). A recent study by Mishimura and colleagues suggested that Mcm8 and Mcm9 work downstream of the Fanconi anemia (FA) and BRCA2/Rad51 pathways, and are required for homologous recombination that promotes sister chromatid exchanges, possibly as a hexameric ATPase/helicase (Nishimura 2012). A more thorough study is needed to better understand the function of Mcm8 and Mcm9, but undoubtedly, they have been proven to be important in DNA replication.

Mcm10

Mcm10 was initially identified in a genetic screen searching for cell division cycle mutants in *S. phombe* (Nasmyth 1981). It is not related to the

Mcm2-7 helicase family. Mcm10 is a DNA binding protein that loads onto chromatin in early S-phase and is required for subsequent loading of additional factors, including Cdc45, RPA (Wohlschlegel 2002), and Pol α (Walter 2000). The precise role of Mcm10 is unknown, and is therefore the basis for the research presented in this thesis. Please refer to Chapter II for a thorough review on Mcm10.

Scope of work

This dissertation presents the structural and biochemical progress made in understanding the role of Mcm10 in DNA replication initiation, specifically focusing on the oligomerization of the vertebrate (*Xenopus laevis*) ortholog. Chapter 2 reviews the domain architecture and structural biology of Mcm10, including a detailed description of the structural features of each domain, and a model for DNA binding and pol α interaction. Chapter 3 describes the structure of the coiled-coil region within the xMcm10 N-terminal domain, the interacting residues responsible for oligomerization, and a model of the oligomeric state of xMcm10 during DNA replication. Finally, Chapter 4 discusses how the work has expanded our knowledge on Mcm10's role in DNA replication, and the future directions for this research.

CHAPTER II

STRUCTURAL BIOLOGY OF REPLICATION INITIATION FACTOR MCM10*

Minichromosome maintenance protein 10 (Mcm10) is a non-enzymatic replication factor required for proper assembly of the eukaryotic replication fork. Mcm10 interacts with single-stranded and double-stranded DNA, polymerase α, and Mcm2-7, and it is important for activation of the pre-replicative complex and recruitment of subsequent proteins to the origin at the onset of S-phase. In addition, Mcm10 has recently been implicated in coordination of helicase and polymerase activities during replication fork progression. The nature of Mcm10's involvement in these activities, whether direct or indirect, remains unknown. However, recent biochemical and structural characterization of Mcm10 from multiple organisms has provided insights into how Mcm10 utilizes a modular architecture to act as a replisome scaffold, which helps to define possible roles in origin DNA melting, pol α recruitment and coordination of enzymatic activities during elongation.

Replication Initiation

DNA replication can be divided into three primary stages: initiation, elongation and termination (Bell and Dutta 2002; Garg 2005). Initiation

commences during the G1 phase of the cell cycle, during which the replisome—

the protein complex responsible for DNA unwinding and synthesis at an active

replication fork—begins to assemble at origins of replication (Figure 4). Initiation

begins with origin licensing, in which the origin recognition complex (ORC),

coupled with Cdc6 and Cdt1, loads minichromosome maintenance (Mcm)

proteins 2-7 onto DNA as a head-to-head double hexamer (Remus 2009). This

event marks the formation of the pre-replicative complex (pre-RC), which

remains inactive in G1-phase.

licensing

ORC

Cdc6, Cdt1

Mcm2-7

(pre-RC)

G_1

activation

Mcm10

DDK

Cdc45   RecQL4   And-1

GINS

RPA   (pre-IC)

S

firing

pol α-primase

pol δ, pol ε
RFC, PCNA

(elongation)
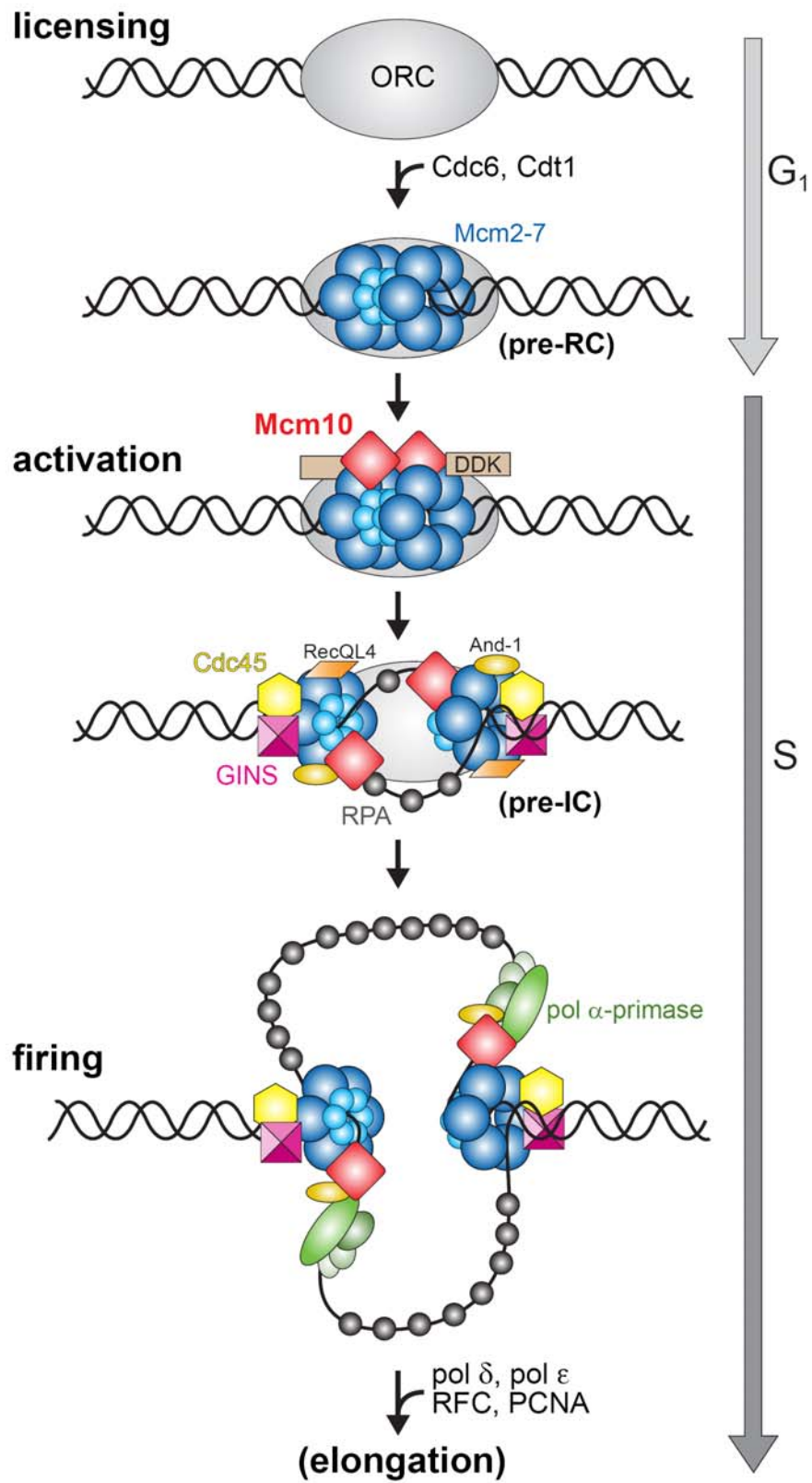
Figure 4. A simplified view of the initiation phase of eukaryotic replication, highlighting key steps involved in replisome assembly. Many replication factors are omitted for clarity. At the end of the G1 phase of the cell cycle, chromatin is licensed for replication at the origin by formation of a pre-replicative complex (pre-RC), which includes an inactive Mcm2-7 helicase. At the onset of S-phase, the pre-RC is activated by Dbf4-dependent kinase (DDK) phosphorylation. Mcm10 loads in early S-phase and is required for loading of Cdc45 and GINS, which form the CMG helicase complex with Mcm2-7 and help constitute a pre-initiation complex (pre-IC). Denaturation of origin DNA at the origin allows for binding of DNA polymerases and the rest of the elongation machinery, stimulating origin firing. Mcm10 and And-1/Ctf4 have been implicated in coupling pol α to the replisome.

The transition to S-phase is accompanied by origin activation. Mcm10 is one of the first proteins loaded onto chromatin at the onset of S-phase, and it is essential for the subsequent recruitment of other replisome proteins (Wohlschlegel 2002). At this point, two phosphorylation events take place to activate the pre-RC. In yeast, cyclin-dependent kinase (CDK) phosphorylates Sld2 and Sld3 and facilitates their binding to Dpb11 (Tanaka 2007; Zegerman and Diffley 2007), and Dbf4-dependent kinase (DDK), composed of Cdc7 and Dbf4, directly phosphorylates Mcm2 and Mcm4 (Lei 1997; Sheu and Stillman 2006). Mcm10 is important for both of these events. It has been shown to stimulate Mcm2-7 phosphorylation by DDK and may, in fact, recruit DDK to the pre-RC (Lee 2003). In addition, the human RecQ4 helicase contains a Sld2-like sequence that is both a phosphorylation target of CDK and a binding site for Mcm10, suggesting that phosphorylation may act as a switch for RecQ4 activity by modulating its interaction with Mcm10 (Xu 2009). These phosphorylation events enable the subsequent loading of two helicase cofactors, Cdc45 and the GINS complex, to form the pre-initiation complex (pre-IC) with the help of Mcm10 and other factors, including And-1/Ctf4 (Tanaka 1998; Zou 2000; Wohlschlegel 2002; Im 2009). Cdc45, GINS, and Mcm2-7 form the CMG complex, which is considered to be the active form of the replicative helicase (Gambus 2006; Moyer 2006; Pacek 2006; Ilves 2010; Costa 2011). Denaturation of origin DNA into single strands forms two bidirectional replication forks and is marked by recruitment of ssDNA binding protein replication protein A (RPA).

The initiation phase concludes upon recruitment of the DNA synthesis machinery to the emerging replication fork. Fork firing requires polymerase α (pol α)-primase to initiate DNA synthesis by generating RNA primers and short stretches of DNA on both leading and lagging strands. Mcm10 and And-1/Ctf4 have been implicated in loading pol α onto chromatin, as well as physical coupling of pol α and Mcm2-7 (Ricke 2004; Zhu 2007; Gambus 2009; Im 2009; Lee 2010). Elongation proceeds through processive DNA synthesis by replicative polymerases δ and ε, which require the sliding clamp, proliferating cell nuclear antigen (PCNA), and the clamp loader, replication factor C (RFC). Fork progression requires concerted DNA unwinding and synthesis through coordination of activities among the CMG complex and polymerases α, δ, and ε.

Role of Mcm10 in Replication

The Mcm10 gene was first identified in genetic screens in yeast. Referred to at the time as *Cdc23*, Mcm10 was shown to be necessary for cell division in *Schizosaccharomyces pombe* (Nasmyth and Nurse 1981; Aves 1998). Bulk DNA synthesis was disrupted in temperature sensitive alleles of *Cdc23*, and thus DNA replication and mitosis were blocked. Similar genes, referred to as *DNA43* and *MCM10*, were identified in screens in *Saccharomyces cerevisiae* and shown to be homologs of *Cdc23* (Dumas 1982; Maine 1984). *DNA43* was found to be essential for entering S-phase and maintaining cell viability (Solomon 1992). Ricke and Bielinsky (2004) showed that the recruitment of *S. cerevisiae* Mcm10 (scMcm10) to replication origins is cell cycle regulated and dependent on pre-RC

assembly, and that scMcm10 is required to maintain pol α on chromatin independently of Cdc45. The importance of Mcm10 to replication initiation in yeast is evident from the number of genetic and physical interactions identified between Mcm10 and proteins involved in origin recognition, replisome assembly, and fork progression (Merchant 1997; Homesley 2000; Kawasaki 2000; Hart 2002; Gregan 2003; Sawyer 2004).

Mcm10 homologs have also been identified and characterized in higher eukaryotes, including humans, *Xenopus*, and *Drosophila* (Izumi 2000; Wohlschlegel 2002; Christensen 2003). Human Mcm10 (hMcm10) interacts with chromatin at the G1/S-phase transition and dissociates in G2-phase (Izumi 2000). It is important for activation of pre-RCs and functional assembly of the replisome and is regulated by phosphorylation-dependent proteolysis during late M- and early G1-phase (Izumi 2001). Studies in *Xenopus* extracts showed that Mcm10 (xMcm10) binds to the pre-RC at the onset of S-phase, with roughly one xMcm10 bound per 5,000 bp of DNA (approximately two Mcm10s per active origin) (Wohlschlegel 2002). These studies also showed Mcm10 to be essential for loading downstream proteins Cdc45 and RPA (Wohlschlegel 2002), which are in turn required for chromatin unwinding and the association of pol α at the origin (Walter and Newport 2000). The *Drosophila* homolog of Mcm10 was able to complement an Mcm10-null strain of *S. cerevisiae* and was shown to interact with many members of the pre-RC in KC cells, including Mcm2, Orc, and Cdc45 (Christensen and Tye 2003). Depletion of Mcm10 from KC cells led to defects in chromosome condensation (Christensen and Tye 2003).

The human, *Xenopus*, and *Drosophila* Mcm10 orthologs have high sequence similarity but are distinct from the yeast protein in several ways. First, the vertebrate proteins have an additional C-terminal domain (Robertson 2008) (Figure 5A). Second, phosphorylated and mono- and diubiquitylated forms of hMcm10 have been identified (Izumi 2001), whereas only diubiquitylated Mcm10 has been shown to be associated with chromatin in yeast (Das-Bradoo 2006). Finally, spMcm10 has been reported to contain primase activity (Fien 2006), a characteristic not observed in other orthologs.

Physical interactions have been observed between Mcm10 and multiple proteins found in the pre-RC and at the replication fork, including ORC (Izumi 2000; Hart 2002), Mcm2-7 (see below), pol α (Ricke and Bielinsky 2004; Ricke 2006; Chattopadhyay 2007), and the recently identified sister chromatid cohesion protein And-1 and the RecQ-like helicase RecQ4 (Zhu 2007; Xu 2009). *S. pombe* Mcm10 (spMcm10) interacts with Mcm4/6/7 and Dfp1p, the *S. pombe* homolog of Dbf4 (Lee 2003). Furthermore, spMcm10 has been reported to stimulate DDK phosphorylation of Mcm2-7 (Lee 2003), and is thus believed to play a role in helicase activation. Studies in *S. cerevisiae* have shown that scMcm10 facilitates assembly of the Cdc45/Mcm2-7/GINS helicase complex (Gambus 2009), and physical interactions have been observed with Mcm2, Mcm4, Mcm5, Mcm6, and Mcm7 subunits (Merchant 1997; Homesley 2000; Hart 2002; Apger 2010). Recent work suggested that scMcm10 serves as a functional linker between the MCM helicase and pol α by coordinating their activities and ensuring their physical stability and integrity at the replication fork (Lee 2010), a role also

identified for Ctf4 (Gambus 2009). *Drosophila* Mcm10 interacts with Mcm2, Dup (Cdt1), Orc2, Cdc45, and Hp1 in yeast two-hybrid assays (Christensen and Tye 2003). xMcm10 interacts with And-1/Ctf4 (Zhu 2007) and with Dna2, a double-strand break repair protein (Wawrousek 2010). hMcm10 interacts with Orc2, Mcm2, and Mcm6 (Izumi 2000), and assembly of the Cdc45/Mcm2-7/GINS complex in human cells requires Mcm10 as well as the And-1/Ctf4 and RecQL4 proteins (Im 2009). hMcm10 also regulates the helicase activity of RecQ4 by direct binding (Xu 2009).

Biochemical studies of Mcm10 have focused on its interactions with DNA and pol α. Mcm10 binds to both single-stranded (ss) and double-stranded (ds) DNA, with about 3-5-fold preference for ssDNA (Fien 2004; Robertson 2008; Eisenberg 2009). Fien *et al.* (2004) showed that spMcm10 can stimulate DNA polymerase activity by interacting with both ssDNA and pol α, leading to the idea that Mcm10 may facilitate the binding of polymerase complexes to primed DNA. Indeed, Mcm10 affects the localization and stability of pol α, further supporting the idea that Mcm10 acts as a molecular chaperone for pol α *in vivo* (Ricke and Bielinsky 2004; Yang 2005; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007). Mcm10 interacts with the p180 subunit of pol α in both yeast and *Xenopus* (Fien 2004; Robertson 2008; Lee 2010). In fact, ssDNA and the N-terminal region of p180 compete for binding to the conserved internal domain of Mcm10 (Warren 2009). Mcm10 can stabilize pol α throughout the cell cycle by preventing its degradation by the proteasome (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007). Moreover, Mcm10

appears to be a cofactor for pol α activity by increasing its affinity for DNA (Fien 2004; Zhu 2007).

Overall Architecture

The Mcm10 protein exists only in eukaryotes; no orthologs have been identified in archaea or bacteria, although loose homology has been observed between regions of Mcm10 and the Mcm2-7 proteins (Robertson 2010). Mcm10 ranges in size from 571 amino acids in yeast to 874 in humans, with regions of sequence homology clustered in the central and extreme N- and C-terminal regions (Figure 5A). The spacing of homologous regions suggests the presence of three distinct structured domains tethered by unstructured linkers. Zinc finger motifs, initially identified from sequence alignments (Homesley 2000; Izumi 2000) and later confirmed by structural analysis (Warren 2008; Warren 2009; Robertson 2010), are present in both the central and C-terminal regions. The yeast homologs lack the C-terminal region altogether (Robertson 2008), suggesting that in lower organisms, the essential functions of Mcm10 reside within its N-terminal and central regions.

A

	30	130	240	430	510	700	
Hs							874
Xl							860
Dm							776
Ce							691
Sc							571
Sp							593
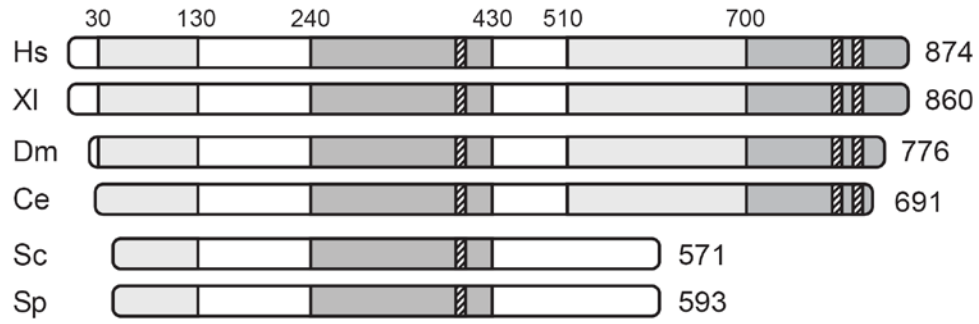
B

species	Oligomeric state	Method used	Reference
S.pombe	Monomer (asymmetric)	GFC, GGS	Lee et al. 2003
	Dimer	GFC, GGS	Fien and Hurwitz 2006
S.cerevisiae	Dodecamer or larger	GFC, GGS	Cook et al. 2003
	Monomer (dsDNA), Trimer (ssDNA)	EMSA, SPR	Eisenberg et al. 2009
X. laevis	FL, Dimer or larger (asymmetric)	AUC	Robertson et al. 2008
	NTD, Dimer (asymmetric)		
H. sapiens	Hexamer	AUC, EM	Okorokov et al. 2007

C

160 Å   120 Å   90°   35 Å

D

1 — NTD → 145   230 — ID → 427   596 — CTD → 860
95 132   280 375 417   690 755 842
CC   OB ZnF   linker   WH Zn Zn
Mcm10   pol α PCNA   DNA   pol α   DNA Mcm2?

36
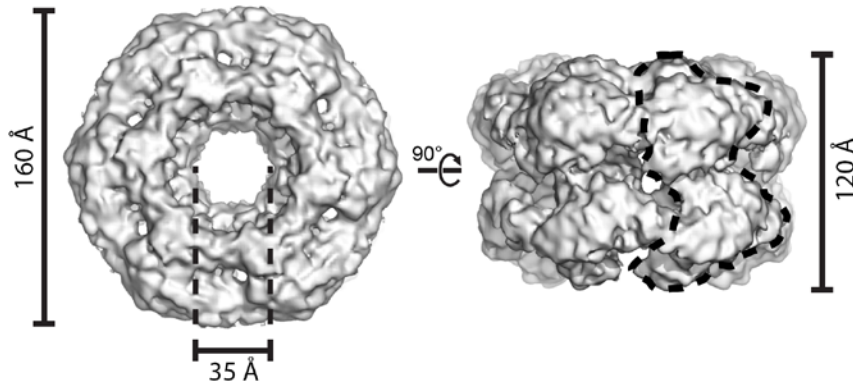
Figure 5. Mcm10 sequence homology, oligomerization, and domain architecture. **A**. A schematic sequence alignment of Mcm10 from *Homo sapiens* (Hs), *Xenopus laevis* (Xl), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce), *Saccharomyces cerevisiae* (Sc), and *Schizosaccharomyces pombe (Sp)*. Light and dark grey bars indicate regi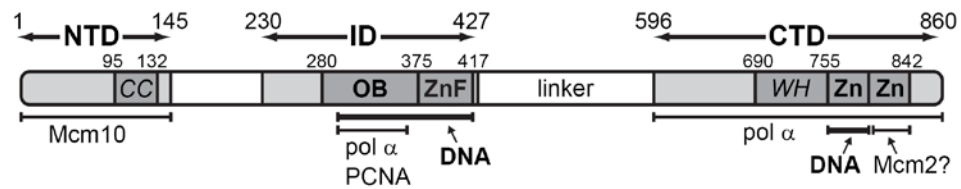ons of moderate and high sequence conservation, respectively, and hatched boxes represent invariant cysteine/histidine clusters involved in zinc coordination. **B**. Various oligomeric states of Mcm10 reported in the literature. GFC, gel filtration chromatography; GGS, glycerol gradient sedimentation; EMSA, electrophoretic mobility shift assay; SPR, surface plasmon resonance; AUC, analytical ultracentrifugation; EM, electron microscopy. **C**. Orthogonal EM reconstruction of human Mcm10 at 16 Å resolution and contoured at 1σ. The dashed line represents one 95-kDa subunit. (Figure adapted from Okorokov 2007) **D**. Domain architecture of *Xenopus laevis* Mcm10. NTD, N-terminal domain; ID, internal domain; CTD, C-terminal domain; CC, predicted coiled coil; OB, oligonucleotide/oligosaccharide binding fold; ZnF, zinc-finger; WH, predicted winged helix; Zn, zinc ribbon. Interactions with proteins and DNA are shown below the schematic.

Biochemical and structural studies using vertebrate and yeast Mcm10 orthologs have been rather controversial with regard to the architecture and oligomeric state of the full-length Mcm10 protein (Figure 5B). scMcm10 has been reported to form large, 800 kDa homocomplexes consisting of ~12 molecules when analyzed by size-exclusion chromatography (Cook 2003), although the shape of the molecule could potentially confound this analysis. Self-association in yeast was shown to be mediated by the central zinc finger-containing domain, and mutations in the zinc-binding residues rendered yeast cells temperature-sensitive, with demonstrable replication defects (Homesley 2000; Cook 2003). A more recent surface plasmon resonance study showed that in the presence of ssDNA, scMcm10 forms complexes with three subunits (Eisenberg 2009). On dsDNA, however, scMcm10 interacted as a monomer with a stoichiometry directly proportional to the length of the DNA (~1 scMcm10 per 21-24 bp). Work from the Hurwitz laboratory has reported highly asymmetric monomeric and dimeric forms of spMcm10 using glycerol gradient centrifugation (Lee 2003; Fien and Hurwitz 2006). Analytical ultracentrifugation of xMcm10 was consistent with self-associated, asymmetric complexes, although the precise oligomeric state could not be determined from the data (Robertson 2008). More recent work using size exclusion chromatography with multi-angle light scattering is indicative of xMcm10 complexes containing 2-3 subunits in the absence of DNA (see Chapter III). This is consistent with the presence of a coiled-coil domain at the N-terminus of the protein (Robertson 2008) and the calculation of two molecules per

replication origin based on the concentration of chromatin-bound Mcm10 in *Xenopus* extracts (Wohlschlegel 2002).

hMcm10 was reported to form a ring-shaped hexameric structure using electron microscopy (EM) and single-particle analysis (Okorokov 2007). The particle has dimensions of 160 Å x 120 Å, a 35 Å central channel (Figure 5C), and a system of smaller lateral channels and inner chambers. The volume of the electron density calculated at the 1 sigma contour level using Chimera (Pettersen 2004) is consistent with a 570 kDa MW particle, or six 95-kDa subunits (unpublished result). From the side, individual subunits appear to adopt two distinct lobes. Model fitting with the structures available at the time suggested that each subunit within the hexamer had the same orientation, with the zinc molecules positioned toward the upper and lower edges of the ring (Okorokov 2007). Subsequent crystal and NMR structures of individual Mcm10 domains, discussed below, cannot be unambiguously positioned into the EM density. The hexameric structure was reportedly consistent with its sedimentation behavior by analytical ultracentrifugation, although the experimental data were not presented (Okorokov 2007).

The authors of the EM structure provided two explanations for hexamerization of Mcm10. The first was that this architecture may enable a topological link with DNA to allow for processive DNA binding like many other ring-shaped DNA-binding proteins. Another explanation was that Mcm10 inherited the hexameric fold from a DNA helicase ancestor, but lost the helicase activity during evolution and instead now serves as a "docking" module to

facilitate protein-protein interactions in DNA replication, such as Mcm2-7 helicase and pol α (Patel 2000; Pape 2003; Chen 2005; Okorokov 2007). It is enticing to speculate that a hexameric Mcm10 structure would provide an extensive binding interface for the six subunits of Mcm2-7, although there are no data to support such a hexamer-hexamer interaction, and Mcm10 does not travel with the helicase that has been uncoupled from the replisome by inhibition of the polymerase with aphidicolin (Pacek 2006). In light of the facts that a hexameric form of Mcm10 has not been reported in non-human orthologs, that other studies identify Mcm10 assemblies composed of 2-3 subunits, and that only two molecules of Mcm10 are likely present at the origin, we offer an additional explanation—that the hexamer is simply one of several states occurring in cellular equilibrium and is needed under specific conditions during the onset of replication. For example, hexamerization may be used for sequestering the molecule at the replication fork or as a compact storage state of the protein during replication inactivity. We note that to be consistent with the available oligomerization data, the Mcm10 hexamer may in fact be a trimer of dimers.

Mcm10 Domain Structure

Biochemical and structural studies have been performed using xMcm10, which has 84% sequence similarity and 58% identity to the human protein. Limited proteolysis and mass spectrometry revealed that full length xMcm10 is composed of three structured domains at the N-terminal (NTD; residues 1-145), internal (ID; 230-427), and C-terminal (CTD; 596-860) regions of the protein

(Figure 5D) (Robertson 2008). The functional significance of the NTD is currently undefined, while the ID and CTD each bind DNA and pol α (Robertson 2008). Interdomain linkers are predicted to be largely unstructured by secondary structure and disorder predictions and by virtue of their extreme proteolytic sensitivity in purified preparations (Robertson 2008).

Mcm10-NTD

Circular dichroism indicates that the NTD is predominantly α helical and random coil, consistent with secondary structure predictions (Robertson 2008). The NTD alone is a dimeric assembly unit as judged by analytical ultracentrifugation, consistent with the presence of a predicted coiled-coil motif comprising residues 93-132 (Robertson 2008). A strong yeast 1-hybrid interaction from the first 100 residues of *Drosophila* Mcm10 was recently reported (Apger 2010), suggesting that the NTD might function as an oligomerization domain for the full-length protein. Interestingly, self-interaction of ID and CTD regions was observed in yeast 2-hybrid assays when the NTD was deleted, suggesting that the NTD may not be the only point of contact between Mcm10 subunits. Nonetheless, the strong NTD self-interaction supports a proposed model in which Mcm10 forms a dimer with two subunits oriented in the same direction, which provides a plausible explanation for interaction of Mcm10 with both leading and lagging strand polymerases at a replication fork. Unlike the ID and CTD, the NTD does not bind to DNA (Robertson 2008).

Mcm10-ID

The ID (residues 230-427) is homologous across all species from vertebrates to yeast and is the most conserved region in the entire protein (Izumi 2000). Mutations in this region were identified in yeast genetic screens to affect minichromosome maintenance and replication *in vivo* (Nasmyth and Nurse 1981; Maine 1984; Grallert 1997; Liang 2001). The ID contains a CCCH-type zinc finger and an oligonucleotide/oligosaccharide binding (OB)-fold (Figure 5D) that are now known to facilitate interactions with a number of proteins and DNA (Izumi 2000; Ricke and Bielinsky 2006; Warren 2008). Specifically, the ID has been shown to interact with ssDNA and the N-terminal 323 residues of DNA polymerase α (Robertson 2008). In addition, a PCNA interacting peptide (PIP) region was identified in the sequence of scMcm10's ID (Das-Bradoo 2006). Mutations within the PIP box abrogated the interaction between diubiquitylated Mcm10 and PCNA (Das-Bradoo 2006; Warren 2009).

The crystal structure of xMcm10-ID revealed that this region forms a globular domain consisting of an α-helical/random coil region (αA-αB, residues 230-283), an OB-fold (β1-β5.2, residues 286-375), and a C-terminal zinc finger motif (βC-αE, residues 378-418) (Figure 6A). The α-helical/coil region packs onto the back of the OB-fold, opposite the canonical DNA-binding cleft of the OB-fold, to form a flat molecular surface (Warren 2008). The zinc finger protrudes sideways relative to the OB-fold cleft and makes extensive electrostatic and van der Waals contacts with both the L23 loop of the OB-fold and the α-helical/coil region. It is interesting to note that the sequential arrangement of the OB-fold and

the zinc finger in Mcm10-ID is different from other DNA processing proteins that contain both structural motifs. In the structures of the archaeal MCM helicase (Fletcher 2003), RPA trimerization core (Bochkareva 2002), T4 gp32 (Shamoo 1995), and NAD^+- dependent DNA ligase (Lee 2000), a zinc ribbon is inserted into the OB-fold L12 loop, whereas in Mcm10-ID the zinc finger is C-terminal to the entire OB-fold (Warren 2008) (Figures 6A,D). The unique arrangement of the OB-fold/zinc finger suggests that, in Mcm10, this domain assembly may have a unique function.
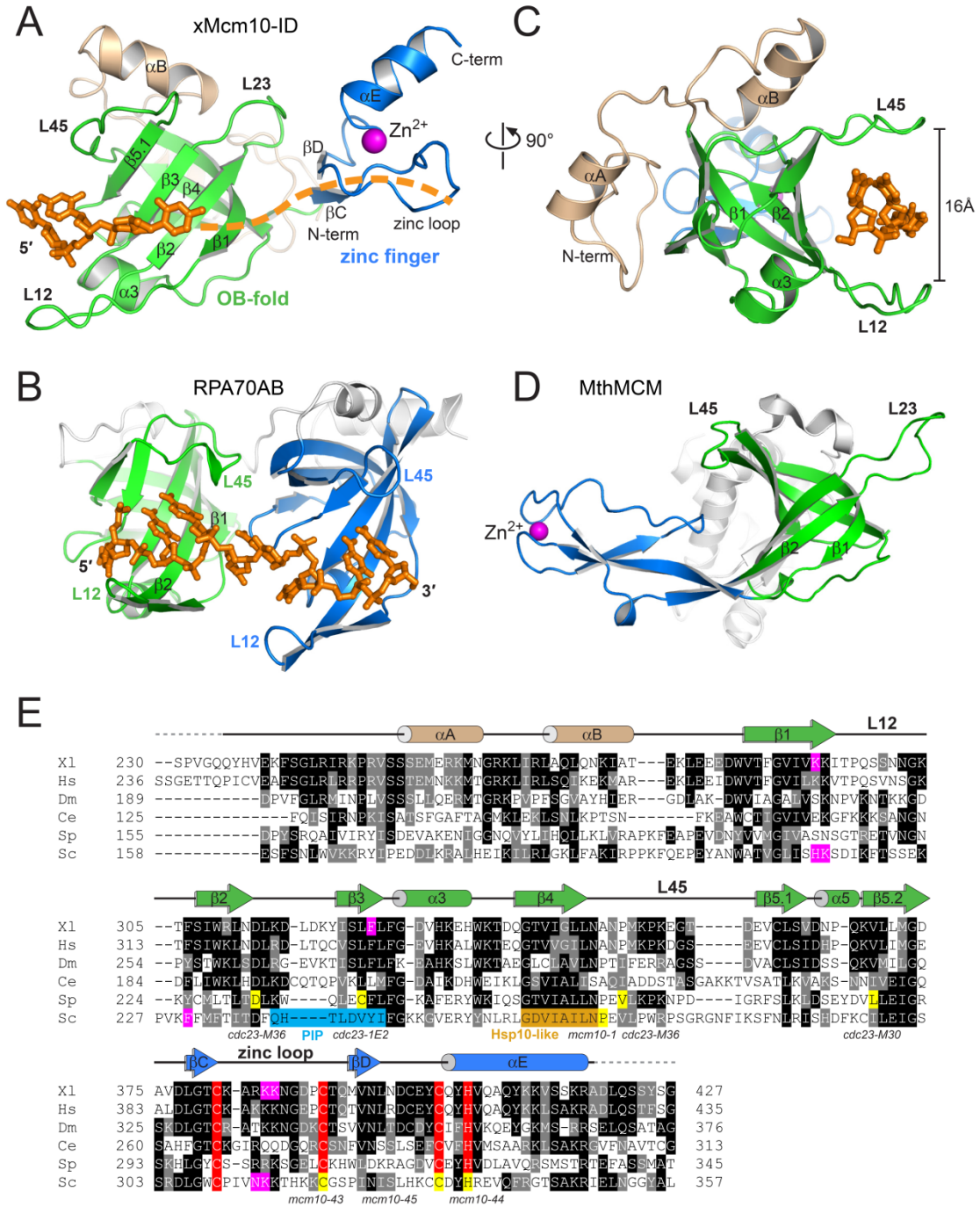
44

Figure 6. The crystal structure of xMcm10-ID bound to ssDNA. **A.** xMcm10-ID (residues 235-419) bound to ssDNA, with the OB-fold in green, the zinc finger in blue, the Zn2+ ion in magenta, and the N-terminal α-helical/coil region in tan. The three nucleotides of ssDNA observed in the crystal structure are shown as orange sticks. ssDNA traverses both the OB-fold cleft and the zinc loop. The trajectory of bound DNA determined by NMR is represented by the orange dashed line. **B**. Crystal structure of RPA70AB subunit bound to ssDNA (PDB ID 1JMC). The OB-folds are colored green and blue, and the ssDNA orange. **C**. Crystal structure of the xMcm10/ssDNA complex viewed 90° with respect to the view shown in panel A. **D**. Crystal structure of *Methanobacterium thermoautotrophicum* MCM (PDB ID 1LTL). **E**. Sequence alignment of Mcm10-ID. Zn^{2+}-coordinating residues are highlighted in red. Mutations identified in yeast genetic screens to affect cell growth and DNA replication are highlighted in yellow. Residues that affect xMcm10 binding to DNA *in vitro* or that increase the sensitivity of *S. cerevisiae* to HU are highlighted in pink. The PIP-box and Hsp10-like motif are highlighted in blue and gold, respectively.

The tandem OB-fold-zinc finger arrangement in xMcm10-ID is reminiscent of the high affinity ssDNA binding surface created by side-by-side OB-folds in the RPA70AB sub-domain (Bochkarev 1997) (Figure 6B). NMR chemical shift perturbation of xMcm10-ID indicated that ssDNA binds to both the OB-fold cleft and to the extended loop of the highly basic zinc finger (Warren 2008). The nature of the ssDNA interaction with the concave cleft of the OB-fold was revealed by the crystal structure of xMcm10-ID in complex with ssDNA (Warren 2009). A tricytidine oligonucleotide was clearly observed within the OB-fold cleft, traversing β strands β1- β3 and β5.1 (Figure 6A). The channel created by loops L12 and L45 was ~16 Å in diameter (Figure 6C), allowing the ssDNA a degree of flexibility that precluded observation of atomic-level interactions. However, the polarity of the ssDNA was unmistakably defined, with the 5′ end oriented toward β5.1 and the 3′ end toward β1 and the zinc finger, similar to the polarity reported for the RPA70AB structure (Bochkarev 1997) (Figure 6B). Also similar to RPA70AB, the L12 loop was unobservable in the unliganded structure, presumably due to flexibility (Warren 2008), but upon DNA binding, its electron density was readily visible (Warren 2009).

The Mcm10-ID zinc finger extends the ssDNA binding surface of the OB-fold in a manner analogous to the RPA70B subunit (Bochkarev 1997). A crystal lattice contact occluded DNA binding by the zinc finger in the Mcm10-ID/ssDNA complex structure (Warren 2009). Nonetheless, NMR chemical shift perturbation had unequivocally showed both this region and the cleft between it and the OB-fold to be affected by ssDNA binding, and residues in these regions were shown to affect DNA binding by xMcm10-ID and replication in yeast (Warren 2008). A Lys385Glu/Lys386Glu double mutant on the extended zinc loop reduced ssDNA binding affinity by 10-fold, and a Lys293Ala mutant in the cleft reduced it by 5-fold (Warren 2008). Transferring these mutations to yeast for assessment of their functional consequences showed that they increased the sensitivity of yeast cells to treatment with hydroxyurea (Warren 2008). The Lys293Ala mutation (His215Ala/Lys216Ala in scMcm10) caused a 2-fold decline in cell survival, while the Lys385Glu/Lys386Glu mutation (Asn313Glu/Lys314Glu in yeast) led to a striking 7-fold decrease. Cell survival was also significantly compromised (~60%) in cells containing the Phe306Ala mutation (Phe230Ala/Phe231Ala in yeast), which resides in the cleft between the OB-fold and zinc finger. Interestingly, the zinc finger domain was also found to be affected by dsDNA binding (Warren 2008). The presence of an extended loop in the zinc finger renders it structurally distinct from the archetypical Zif268 zinc finger that binds dsDNA in a sequence dependent manner, so it remains to be seen how dsDNA binds to this motif.

The Mcm10-ID crystal structures elucidated the yeast mutations originally identified to affect minichromosome maintenance and DNA replication (Figure

47

6E). The *cdc23-1E2* (Cys239Tyr) (Grallert and Nurse 1997) and *cdc23-M30* (Leu287Pro) (Liang and Forsburg 2001) mutations, which correspond to xMcm10 Leu323 and Leu369, respectively, are located in the interior of the OB-fold's β-barrel, and thus are likely to cause structural perturbations that disrupt protein folding. Other mutations probably affected protein interactions necessary for replisome formation and/or progression. These include *cdc23-M36* (Asp232Gly), corresponding to the invariant xMcm10 Asp313 that lies on the interior of the L23 loop, and *cdc23-M36* (Val265Ile) and *mcm10-1* (Pro269Leu), which map to solvent exposed positions in the L45 loop (Nasmyth and Nurse 1981; Maine 1984). The human counterpart to the xMcm10-ID has been crystallized (Jung 2008), but the structure was never determined and is expected to be virtually identical to the reported *Xenopus* domain on the basis of high sequence homology (58% identity; 84% similarity).

Mcm10-CTD

Vertebrate homologs of Mcm10 contain a CTD that is unique to higher eukaryotes; yeast Mcm10 is not predicted by sequence alignments to have this domain (Figure 5A). Interactions between xMcm10-CTD (residues 596-860) and ssDNA, dsDNA, and pol α have been mapped to a proteolytically stable subdomain (residues 690-842) that consists of a putative winged helix motif (residues 690-755) followed by tandem CCCH- and CCCC-type zinc motifs (residues 756-842) (Figure 7A) (Robertson 2008; Robertson 2010). Heteronuclear NOE experiments on this region showed that the putative winged

helix contains high backbone flexibility while the zinc motif is more rigid (Robertson 2010). The two Zn^{2+} atoms in xMcm10-CTD, originally identified by atomic absorption spectroscopy, likely play a structural role based on the observations that, in the presence of EDTA, the CTD is more proteolytically sensitive and DNA binding affinity decreases (Robertson 2008).
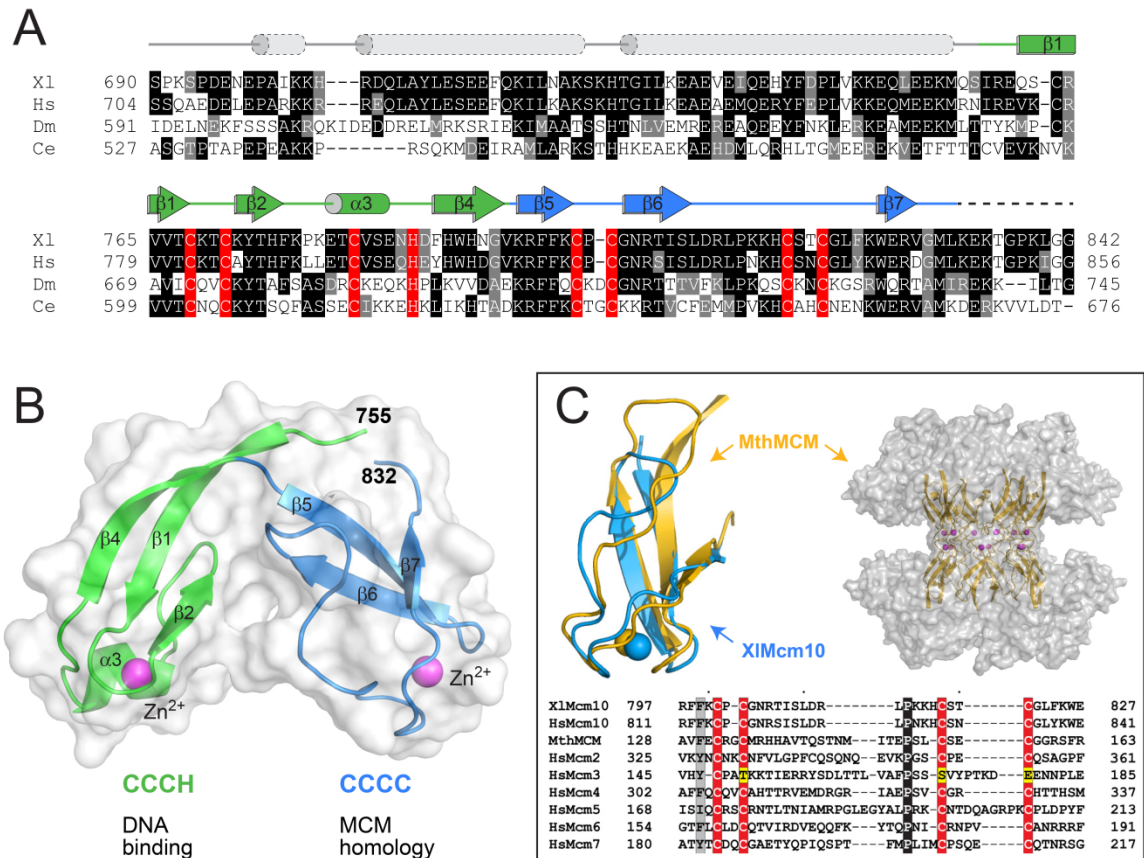
Figure 7. NMR structure of xMcm10-CTD zinc binding region. **A**. Sequence alignment of Mcm10-CTD. Secondary structure elements shown above the sequence are either predicted (grey) or determined from the NMR structure (green, blue). CCCH and CCCC zinc coordinating residues are highlighted in red. **B**. The NMR structure of xMcm10 (756-842) is shown as a ribbon with a transparent grey molecular surface. **C**. Structural and sequence alignment of the CCCC zinc ribbon from xMcm10 (aa 796-830) (blue) and MthMCM (gold, PDB ID 1LTL). In the crystal structure of the MthMCM N-terminal domain (top right of panel), the CCCC zinc ribbon mediates head-to-head double hexamer formation. The sequence alignment of this region (bottom of panel) shows the CCCC motif to be conserved in human Mcm2-7.

The solution NMR structure of the zinc binding region of xMcm10-CTD revealed a V-shaped globular domain in which the two zinc binding motifs are tethered by a hinge and the two zinc atoms bind at the tips of the V (Figure 7B) (Robertson 2010). The N-terminal CCCH zinc motif (residues 756-795) consists of a three-stranded antiparallel β-sheet capped with a short perpendicular α-helix with a Zn^{2+} ion embedded in between. DNA binding maps to the CCCH zinc motif, the structure of which is unique to Mcm10. The residues involved in DNA binding trace a nearly continuous 35 Å path around the CCCH arm (Robertson 2010). The length of DNA required for maximal binding affinity was between 10 and 15 nucleotides, suggesting that all of the residues along that path are involved to some extent in interactions with DNA.

The CCCC zinc motif (residues 796-830) adopts a twisted antiparallel β-sheet with the zinc coordinated between the loops by the four cysteines. This motif is not involved in DNA binding, and, interestingly, is identical in structure to a zinc ribbon motif in the N-terminal domain of *Methanobacterium thermoautotrophicum* MCM helicase (mtMCM) (Figure 7C). This MCM zinc motif mediates the head-to-head double hexamer assembly observed in mtMCM crystals (Fletcher 2003; Fletcher 2005) and in scMcm2-7 loaded onto DNA (Remus 2009). The sequence of the CCCC zinc motif in xMcm10-CTD is highly conserved relative to those in the metazoan Mcm2-7 subunits and in Mcm8 and Mcm9 proteins as well (Robertson 2010). Mcm10 has been shown to interact directly with several subunits of Mcm2-7 helicase (Izumi 2000; Lee 2003; Gambus 2006), and the recent finding that *Drosophila* Mcm10's interaction with

Mcm2 is localized to the CTD (Apger 2010) suggests that the CCCC zinc motif in both proteins may be the point of contact. The Mcm2-7 double hexamer that is loaded onto chromatin in the pre-RC (Remus 2009) is able to separate during DNA unwinding (Yardimci 2010), leading us to predict that if Mcm10-Mcm2-7 interactions are indeed facilitated by the zinc motifs, then this interaction would take place only after fork firing. Of course this is highly speculative, and additional experiments are needed to define this aspect of Mcm10's function. What is known with certainty is that Mcm10 interacts with both the helicase and pol α. Most likely, Mcm10 serves as a scaffold to co-localize the essential players within the replisome during the initiation and elongation phases of replication (Ricke and Bielinsky 2004; Lee 2010; Robertson 2010).

Implications of Modular Architecture for Function

Modular architecture is a common feature of DNA processing proteins that allows for the coordination of distinct biochemical activities (Stauffer 2004). Flexible linkers between structured domains allow those domains to accommodate DNA and protein binding partners simultaneously, by virtue of the fact that the domains retain their structure while distance and angular adjustments are made between them. When tandem domains bind to the same entity, affinity for that entity is often increased relative to the strength of binding by one domain or the other. And many DNA processing proteins contain bifunctional folds, which are known to bind both DNA and other proteins. Thus, when two different entities compete for the same binding site, it promotes

molecular hand-off that facilitates the progression of the DNA processing pathway.

The structural organization of Mcm10 exemplifies each of these general features of DNA processing proteins. First, the attachment of the functional domains of Mcm10 by unstructured linkers, coupled with the spatial separation of protein and DNA binding sites, may allow it to bind both DNA and proteins simultaneously. Robertson et al. reported NMR spectra showing that the ID and the CTD of xMcm10 retain their individual structural properties in the context of a larger "ID+CTD" construct containing both domains and that the interdomain linker region is unstructured and flexible (Robertson 2010). Second, the full-length xMcm10 protein, as well as the ID+CTD construct, binds DNA with 10-fold greater affinity than either the ID or CTD alone (Robertson 2008; Warren 2009), and ID+CTD binds pol α-p180 with 15-fold greater affinity than Mcm10-ID alone (Warren 2009). Lastly, ssDNA and the N-terminal region of p180 compete for binding to the OB-fold cleft of Mcm10-ID (Warren 2009). dsDNA also binds to essentially the same site on Mcm10-ID (Warren 2008). Moreover, the PIP box predicted in the scMcm10 sequence (Ricke and Bielinsky 2006) coincides with the OB-fold β3 strand, suggesting that the OB-fold can bind to PCNA as well. Indeed, xMcm10 Phe324, which corresponds to the residue in scMcm10 that mediates interaction with PCNA (Das-Bradoo 2006), had a modest effect on DNA binding (Warren 2008).

The interaction of multiple binding partners with identical sites on Mcm10 fits with two distinct models of molecular hand-off (Figure 8A, B). In the first,

Mcm10-ID binds to ssDNA while the CTD is used to recruit a protein partner (e.g., pol α p180). In the second model, the CTD binds ssDNA and ID recruits a protein binding partner via its OB-fold. Hand-off would be facilitated in either scenario by competition between the binding partner and Mcm10 for the exposed ssDNA or, alternatively, for the OB-fold in Mcm10-ID. Depending on the oligomeric state of Mcm10 *in vivo*, there could be more than one subunit of the ID and the CTD present at the origin, increasing the number of possible interaction points and competition events (Figure 8C). Future studies of the oligomerization state of Mcm10 and of the order of events at the origin will be needed to clarify which model is in play at each stage of replication.

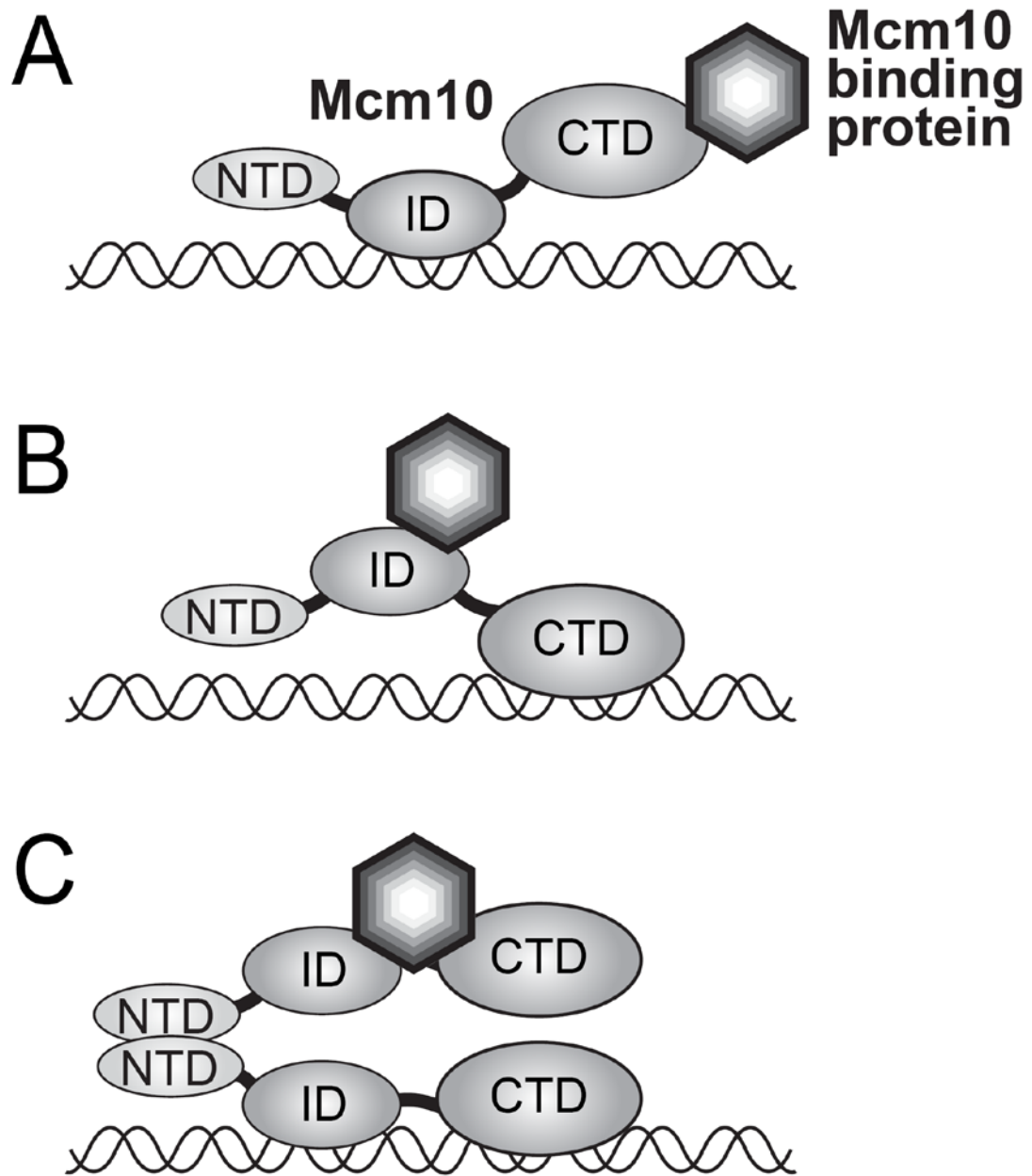**Figure 8.** Three possible models for Mcm10 hand-off of other proteins (e.g., pol α) onto DNA. **A,B.** Either Mcm10-ID or CTD interact with DNA, while the non-DNA bound domain is free to bind protein cargo. **C**. ID+CTD together create a high-affinity DNA binding platform, and Mcm10 self-association through the NTD would present an additional free binding platform to localize proteins to the DNA.

Its modular architecture and lack of enzymatic activity suggest that Mcm10 serves as a scaffold for the coupling of protein and DNA interactions during replication initiation. For example, a head-to-head Mcm10 dimer could couple events on the leading and lagging strands, or physically tether the helicase and pol α, while retaining the polarity necessary for fork progression. Interestingly, recent studies have shown that yeast Mcm10 displays differential packing on ssDNA versus dsDNA (Eisenberg 2009), suggestive of an Mcm10-DNA scaffold during origin melting or helicase unwinding. The authors speculated that a change in Mcm10 conformation or oligomeric state could facilitate strand separation.

Summary and Future Perspectives

Despite decades of work, Mcm10 remains an essential yet mysterious player in DNA replication. As one of the first proteins to load after pre-RC formation, Mcm10 is needed for subsequent protein loading and downstream events in DNA replication initiation. Mcm10 interacts with multiple replisome components and DNA. Its interactions with ssDNA and pol α are mediated by the conserved ID and CTD through OB-fold and zinc finger structural elements. Crystal and NMR structures have elucidated the details of the ID-DNA interactions and have begun to address the binding activity within the CTD. Full-length xMcm10 forms a number of oligomeric species, which may be assembled through coiled coil interactions within the NTD. The functional significance of Mcm10's self-assembly and its interactions within the replisome, the structure of

the NTD, the mechanisms of multi-domain DNA binding activities, and the effects of ubiquitylation and other post-translational modifications on Mcm10 structure and function are all questions that remain unanswered.

The nature of Mcm10 self-assembly is critical for understanding its role at the replication fork, although the structural and functional relationship between the apparent multiple oligomeric states is not at all clear from the literature. A dimerization model best explains the physical and genetic evidence for Mcm10's interaction with both leading and lagging strand polymerases at a replication fork (Fien 2004; Ricke and Bielinsky 2006; Robertson 2008), but this remains to be determined. Detailed structural analyses of the N-terminal domain may help to address this issue. In addition, studies of the configuration(s) of the tandem ID and CTD in complex with DNA, pol α, and other protein binding partners will yield insight into the mechanisms by which Mcm10 acts as a scaffold at the replication origin.

A growing body of research also suggests Mcm10 plays a role in elongation. The first glimpse of a potential role for Mcm10 in fork progression came from the observation in *S. cerevisiae* that Mcm10 mutants delayed completion of DNA synthesis after cells were released from hydroxyurea arrest (Kawasaki 2000). Mcm10 interacted genetically with pol δ and ε (Kawasaki 2000), and a physical interaction with replisome progression complexes, which exist at DNA replication forks, has been observed in yeast (Gambus 2006). In addition, a diubiquitylated form of scMcm10 interacts with PCNA, suggesting that Mcm10 directly participates in DNA elongation (Das-Bradoo 2006). Pacek et al.

showed that Mcm10 travels with the replication fork by inducing specific replication fork pausing on biotin-streptavidin-modified plasmids in *Xenopus* egg extracts (Pacek 2006). In these experiments, Mcm10 was found to localize to the vertebrate DNA replication fork by chromatin immunoprecipitation. Finally, recent work in yeast suggested that Mcm10 coordinates the activities of the Mcm2-7 helicase and pol α and ensures their physical stability at the elongating replication fork (Lee 2010).

Although the majority of work to date has been focused on its role in DNA replication, Mcm10 has also been shown to be important for transcriptional gene silencing (Douglas 2005; Liachko 2005; Liachko 2009; Apger 2010). scMcm10 physically interacts with Sir2 and Sir3, two essential silencing factors in *S. cerevisiae* (Douglas 2005). Moreover, Mcm10 mediates interactions between Sir2 and subunits 3 and 7 of the Mcm2-7 helicase via a ~100-residue segment at its C-terminus. Mutations to this region of Mcm10 caused silencing defects, but had no detrimental effect on replication (Liachko and Tye 2009). The corresponding segment in the *Xenopus* protein resides within an unstructured linker between the ID and CTD (Figure 4B), which suggests that either the yeast protein has an organism-specific function or the vertebrate Mcm10s have an as yet uncharacterized role in gene silencing. The yeast sequence between residues 515 and 523 is predicted to be an amphipathic helix (Liachko and Tye 2009), a finding which warrants further investigation into the analogous segments of its orthologs.

In summary, Mcm10 lies at the heart of the replication initiation pathway. It loads early onto licensed replication origins and is necessary for pre-RC activation and origin melting as a result of its interactions with DNA and many of the enzymes involved in fork progression. In addition to its essential role in establishing active replication forks at the origin, Mcm10 is involved in other aspects of genome utilization. The structures and interactions between Mcm10 and its binding partners are adding to a growing body of knowledge for how multi-conformation scaffolding proteins are used to maintain the integrity of the genome. To further enhance our understanding of the mechanisms involved in replisome assembly and function, the next step is to utilize the existing structures of the Mcm10 DNA binding domains as a foundation to build up larger sub-complexes, taking advantage of the extensive network of Mcm10 interactions. This higher resolution picture of the replisome will be critical to understand the transactions involved at the replication fork, including DNA synthesis, damage response and repair, and cell cycle regulation.

CHAPTER III

MCM10 SELF-ASSOCIATION IS MEDIATED BY AN N-TERMINAL COILED-COIL DOMAIN*

Minichromosome maintenance protein 10 (Mcm10) is an essential eukaryotic DNA-binding replication factor thought to serve as a scaffold to coordinate enzymatic activities within the replisome. Mcm10 appears to function as an oligomer rather than in its monomeric form. However, various orthologs have been found to contain 1, 2, 3, 4, or 6 subunits and thus, this issue has remained controversial. Here, we show that self-association of *Xenopus laevis* Mcm10 (xMcm10) is mediated by a conserved coiled-coil (CC) motif within the N-terminal domain (NTD). Crystallographic analysis of the CC at 2.4 Å resolution revealed a three-helix bundle, consistent with the formation of both dimeric and trimeric Mcm10 CCs in solution. Mutation of the side chains at the subunit interface disrupted *in vitro* dimerization of both the CC and the NTD as monitored by analytical ultracentrifugation. In addition, the same mutations also impeded self-interaction of the full-length protein *in vivo*, as measured by yeast-two hybrid assays. We conclude that Mcm10 likely forms dimers or trimers to promote its diverse functions during DNA replication.

Introduction

DNA replication is carried out by multi-protein factories that in eukaryotes are assembled in stages to regulate the timing of DNA synthesis within the cell cycle (Bell and Dutta 2002; MacNeill 2012; Thu 2013). Pre-replicative complexes (pre-RCs) are assembled at origins during G1 and are composed of origin recognition complex (ORC), Cdc6, Cdt1, and an inactive form of the minichromosome maintenance (Mcm) 2-7 helicase. The pre-RC is activated at the onset of S-phase by Dbf4-dependent kinase (DDK) and cyclin-dependent kinase (CDK) activities (Heller 2011). In yeast, CDK phosphorylates Sld2 and Sld3 and facilitates their binding to Dpb11 (Adachi 1997; Tanaka 2007; Zegerman and Diffley 2007) and DDK phosphorylates Mcm2 and Mcm4 (Lei 1997; Sheu and Stillman 2006) to promote the assembly of additional factors. Ultimately, pre-RC activation leads to the loading of Cdc45 and GINS (Go-Ichi-Nii-San), which form a functional helicase (CMG) complex with Mcm2-7 (Gambus 2006; Moyer 2006; Pacek 2006; Im 2009; Ilves 2010; Costa 2011). Unwinding of the origin is signified by loading of replication protein A (RPA), followed by recruitment of DNA polymerase α (pol α)-primase, which initiates DNA synthesis at the heads of the leading strands and each Okazaki fragment.

Mcm10 is a non-enzymatic protein that aids assembly and activation of the replisome and coordinates helicase and polymerase activities during elongation (Ricke and Bielinsky 2004; Chattopadhyay and Bielinsky 2007; Zhu 2007). Mcm10 interacts with single- (ss) and double-stranded (ds) DNA (Fien 2004; Robertson 2008; Eisenberg 2009), is loaded onto chromatin in early S-phase,

and is essential for helicase activation (Kanke 2012; van Deursen 2012; Watase 2012) and the subsequent recruitment of replisome proteins, including RPA and pol α (Wohlschlegel 2002; Ricke and Bielinsky 2004; Heller 2011). *Saccharomyces cerevisiae* Mcm10 (scMcm10) is required to maintain pol α on chromatin independently of Cdc45 (Ricke and Bielinsky 2004), and both Mcm10 and the sister chromatid cohesion protein And-1/Ctf4 have been implicated in loading pol α onto chromatin and physically coupling pol α and Mcm2-7 (Ricke and Bielinsky 2004; Zhu 2007; Gambus 2009; Im 2009; Lee 2010). Mcm10 from various organisms has been shown to interact physically with key proteins involved in initiation and elongation, including ORC (Izumi 2000; Hart 2002), Mcm2-7 (Merchant 1997; Homesley 2000; Hart 2002; Apger 2010), pol α (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007; Robertson 2008; Warren 2009), proliferating cell nuclear antigen (PCNA) (Das-Bradoo 2006), And-1 (Zhu 2007) and the RecQ-like helicase RecQ4 (Zhu 2007; Xu 2009). Mcm10 binds to the Sld2-like sequence of the human RecQ4 helicase, suggesting that it may regulate the phosphorylation of RecQ4 to facilitate initiation (Xu 2009). Furthermore, loss of Mcm10 from human cells causes chromosome breakage and genomic instability (Chattopadhyay and Bielinsky 2007; Thu and Bielinsky 2013).

Mcm10 contains at least three functional domains (Robertson 2008). An N-terminal coiled-coil (CC) domain (NTD) has been implicated in Mcm10 self-association (Robertson 2008) and the interaction with Mec3, a subunit of the 9-1-1 clamp (Alver and Bielinsky, unpublished results). In addition, the protein has a

highly conserved internal (ID) and vertebrate-specific C-terminal domain (CTD) that bind DNA and the catalytic (p180) subunit of pol α (Robertson 2008; Warren 2009; Du 2012). The yeast orthologs have also been shown to interact with DNA and pol α despite the apparent lack of the CTD (Fien 2004; Ricke and Bielinsky 2004; Yang 2005; Eisenberg 2009). Thus, the ID is likely to mediate these interactions in *S. cerevisiae*. Moreover, recent evidence suggests that acetylation of the ID and CTD in human Mcm10 differentially controls their respective DNA binding and protein-protein interactions (Fatoba 2013). However, the details of this potential mechanism are still unclear.

The oligomeric state of Mcm10 has remained controversial, with reports ranging in size from 1-12 subunits (Du 2012). scMcm10 was shown by size-exclusion chromatography to form large, 800-kDa homocomplexes consisting of ~12 molecules (Cook 2003). Self-association in that case was presumably dependent on the integrity of the zinc-finger (ZnF) motif within the ID, although the purified xMcm10-ID was found to be monomeric (Robertson 2008). Electron microscopy (EM) and single-particle analysis of the human protein showed a hexameric ring-shaped structure (Okorokov 2007). In contrast, asymmetric monomeric and dimeric forms of *S. pombe* Mcm10 (spMcm10) were reported (Lee 2003; Fien and Hurwitz 2006). Similarly, xMcm10 exhibited mass-dependent association into low molecular weight complexes that were presumed to represent Mcm10 dimers solely on the basis of dimerization of the isolated NTD (Robertson 2008). Consistent with NTD-mediated self-assembly, scMcm10 showed a strong yeast two-hybrid interaction that was ablated when one binding

partner carried a truncation of the first 100 amino acids. Moreover, these truncation mutants exhibited a striking sensitivity to the replication inhibitor hydroxyurea that was revealed in the absence of the 9-1-1 checkpoint clamp (Alver and Bielinsky, unpublished results). These observations agree with a report that demonstrates that scMcm10 is monomeric when bound to dsDNA, but capable of forming multi-subunit complexes on ssDNA (Eisenberg 2009).

Here, we studied the role of the NTD on xMcm10 self-assembly using structural, biophysical, and *in vivo* binding assays. We show that the CC region is necessary and sufficient to explain Mcm10-Mcm10 interaction, and is capable of forming both dimers and trimers in solution. The trimeric form of the CC was stabilized in a crystal structure, which revealed the residues at the subunit interface. Specific mutations at this interface disrupted dimerization of the isolated CC, the NTD, and eliminated self-association of the full-length protein by yeast-two hybrid interaction.

Experimental Procedures

Protein Purification

Full-length *Xenopus laevis* Mcm10 (xMcm10) was purified from baculovirus infected insect cells using the Bac-to-Bac expression system (Invitrogen). The gene was subcloned into pFastBac1 vector with a His$_6$ tag added to the C-terminus by PCR. Protein was expressed in Hi-5 insect cells for 48 hr. Cells were resuspended in lysis buffer (50 mM Tris buffer (pH 7.5), 500 mM NaCl, 10%

glycerol) and hand homogenized. Protein was purified by nickel-nitrilotriacetic (NTA) acid affinity chromatography. Pooled Ni-NTA fractions were buffer exchanged into 50 mM Tris buffer (pH 7.5), 150 mM NaCl, and 10% glycerol and purified using Source Q (GE Healthcare) cation exchange, followed by gel filtration on a Superose6 (GE Healthcare) column equilibrated in 25 mM Tris buffer (pH 7.5), 150 mM NaCl, 5% glycerol, and 2mM β-mercaptoethanol (BME). Mcm10ΔN (aa 230-860) and Mcm10-NTD (aa 1-145) were expressed and purified as previously described (Robertson 2008; Warren 2009).

Gene sequences encoding xMcm10 amino acids 95-124 and 95-132 were cloned into a pMALX(E) vector using NotI and BamHI restriction sites to generate coiled-coil (CC) fragments fused to the C-terminal end of a mutant form of maltose binding protein (MBP) with a short, uncleavable peptide linker as previously described (Moon 2010). The recombinant proteins were overexpressed in *E.coli* C41 cells for 3 hrs at 37°C in LB medium supplemented with 100 µg/mL ampicillin and 0.5 mM IPTG at mid-log phase. Cells were resuspended in lysis buffer and lysed under pressure (25,000 p.s.i.) using an EmulsiFlex-C3 homogenizer (Avestin, Inc.). Lysate was centrifuged at 35,000Xg for 20 min. The supernatant was incubated with amylose resin (New England Biolabs) overnight at 4°C and washed with 15 column volumes of lysis buffer. Fusion proteins were eluted with 40 mM maltose in lysis buffer, concentrated, and further purified by size exclusion chromatography on a Superdex S200 column (GE Healthcare) equilibrated in 25 mM Tris buffer (pH 7.5), 150 mM NaCl, 5% glycerol, 4 mM BME, and 40 mM maltose. Purified MBP-CC proteins

were flash frozen and stored at -80°C in 25 mM Tris buffer (pH 7.5), 150 mM NaCl, 0.2 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), and 5 mM maltose.

The 2A (L104A/L108A), 2D (L104D/L108D) and 4A (L104A/L108A/M115A/L118A) mutations were incorporated into full-length, NTD, and MBP-CC constructs using the QuikChange Mutagenesis Kit (Qiagen). Mutant proteins were expressed and purified the same as their corresponding wild-type proteins.

X-ray crystallography

Purified MBP-CC proteins were concentrated to 50 mg/mL using a 10-kDa MWCO Amicon spin concentrator and buffer exchanged into 25 mM Tris buffer (pH 7.5), 150 mM NaCl, 0.2 mM TCEP, and 5 mM maltose for crystallization. Crystals were grown by sitting drop vapor diffusion at 16°C by adding 2 μl protein to 2 μl reservoir solutions containing 0.1 M sodium acetate (pH 4.8), 0.1 M NaCl, 0.1 M $CaCl_2$, 15% PEG 2K, and 5% (w/v) N-dodecyl-beta-D-maltoside (MBP-CC[95-124]), or 0.05 M sodium acetate (pH 4.8), 0.2 M $NH_4H_2PO_4$, and 12% PEG 3350 (MBP-CC[95-132]). Crystals were flash frozen in 0.1 M sodium acetate (pH 4.8), 0.1 M NaCl, 0.1 M $CaCl_2$, 22% PEG 2K, and 15% glycerol (MBP-CC[95-124]) or 0.05 M sodium acetate (pH 4.8), 0.2 M $NH_4H_2PO_4$, 23% PEG 3350, and 15% glycerol (MBP-CC[95-132]) prior to data collection. X-ray diffraction data were collected at the Advanced Photon Source LS-CAT/sector 21 and processed using HKL2000 (Otwinowski 1997).

The structures of MBP-CC$^{95-124}$ (2.4 Å) and MBP-CC$^{95-132}$ (3.1 Å) were determined by molecular replacement using MBPX(E) from PDB ID 3H4Z as a search model (Moon 2010; Mueller 2010). Phases generated from three copies of MBP in the asymmetric unit revealed clear electron density for the Mcm10 coiled-coil in both cases. The models were built in COOT (Emsley 2004) and refined against a maximum likelihood target in PHENIX (Adams 2002). Although one additional turn of the α-helix was visible in the MBP-CC$^{95-132}$ structure, the side chains could not be unambiguously identified, and thus the lower resolution structure was not pursued further. Anisotropic motion was modeled using translation/libration/screw-rotation (TLS) refinement, with each protomer defined as a TLS group. Individual anisotropic B-factors derived from the refined TLS parameters were held fixed during subsequent rounds of refinement. Adjustments to the model and addition of solvent was carried out iteratively through inspection of $2F_o$-F_c, F_o-F_c and composite omit electron density maps. The final MBP-CC$^{95-124}$ model, consisting of MBPX(E) residues 1-367, the five-residue linker (AAAMG), and xMcm10 residues 95-122, was validated using PROCHECK (Laskowski 1996). 97.5% and 2.2% of residues reside in the favored and allowed regions of the Ramachandran plot, respectively. The remaining 0.3% in disallowed regions reside in the MBP-CC linker, MBP loops, or the extreme MBP amino terminus. The final model was deposited in the Protein Data Bank under accession number 4JBZ.

Analytical Ultracentrifugation

Sedimentation velocity experiments were performed using a Beckman ProteomeLab XL-I ultracentrifuge operating at 42,000 rpm and 4°C (Mcm10, Mcm10ΔN) or 20°C (MBP, MBP-CC, and NTD). Full-length Mcm10 was concentrated to 1.6 mg/ml in PBS buffer, 150 mM NaCl, and 0.3 mM TCEP, and Mcm10ΔN was analyzed at 1.0 mg/ml in 25 mM Tris buffer (pH 7.5), 150 mM NaCl, 2 mM $MgCl_2$, 5% glycerol, and 4 mM BME. MBP, MBP-CC, and NTD constructs were analyzed at 0.6 mg/ml in PBS buffer (pH 7.4), 150 mM NaCl, and 0.3 mM TCEP or 25 mM sodium acetate buffer (pH 4.7), 150 mM NaCl, and 0.3 mM TCEP. Buffer viscosity, buffer density and partial specific volume were calculated using SEDNTERP (Laue 1992). Data was processed using c(s) analysis in SEDFIT (Schuck 2000; Schuck 2002).

Multi-angle light scattering

Molecular mass analysis of *Xenopus* Mcm10 and Mcm10ΔN by size exclusion chromatography and multi-angle light scattering (SEC-MALS) was carried out using a Superose6 10/300GL column (GE Healthcare) operating at 0.4 ml/min in 25 mM Tris buffer (pH 7.5), 150 mM NaCl, 2% glycerol, and 2 mM BME. Absorbance, refractive index, and light scattering of the eluants were measured using a DAWN HELEOS II detector (Wyatt Technology) and data analyzed by ASTRA software.

Yeast two-hybrid assay and immunoblotting

Mcm10 self-association by yeast two-hybrid was performed using the GAL4-based MATCHMAKER Two-Hybrid System 3 (CLONTECH Laboratories, INC.). Full-length wild-type and 2D and 4A mutants of *Xenopus laevis* Mcm10 (xMcm10) and a construct lacking the N-terminal region (xMcm10ΔN, residues 230-860) were cloned into pGBKT7 (bait) and pGADT7 (prey) vectors and co-transformed into yeast strain AH109. Positive controls included co-transformation of murine p53/pGBKT7 and SV40 large T-antigen/pGADT7. As negative controls, Mcm10/pGBKT7 was co-transformed with the pGADT7 empty vector, Mcm10/pGADT7 was co-transformed with the pGBKT7 empty vector, and the pGBKT7 and pGADT7 empty vectors were co-transformed. Co-transformed cells were plated on synthetic defined medium lacking leucine and tryptophan (SD/-leu-trp) or on SD lacking leucine, tryptophan, histidine, and adenine (SD/-Leu-Trp-His-Ade). Cells were transformed immediately prior to cell growth and spotting on plates. Plates were incubated at 30°C for two days and photo-documented. For the spotting assay, single colonies were picked from the transformation plates and grown in liquid culture (SD/-Leu-Trp, supplemented with 0.2% adenine) at 30°C for 1 day. All samples were diluted to a starting concentration of 24,625 cells/μL. 5 μl each of 10-fold serial dilutions were spotted on SD/-Leu-Trp and SD/-Leu-Trp-His-Ade selective plates and incubated for four days at 30°C. To verify protein expression, total protein extracts were obtained from yeast cultures by trichloroacetic acid (TCA) preparation as described and separated by SDS-PAGE and subsequently transferred onto nitrocellulose

membrane (Haworth 2010). HA-tagged xMcm10 was visualized using a horseradish peroxidase (HRP)-conjugated anti-HA antibody (Roche, 3F10). Myc-tagged xMcm10 was detected using an anti-Myc antibody (Thermo Scientific, 9E11).

Results

Mcm10 self-associates through its N-terminal domain

Full-length *Xenopus laevis* Mcm10 (xMcm10) self-associates into low molecular mass complexes, which we previously hypothesized to form as a result of NTD dimerization (Robertson 2008). In order to investigate the contribution of the NTD on self-association behavior, we purified a deletion mutant lacking the first 230 residues of xMcm10 (Mcm10ΔN) and analyzed molecular masses of full-length and Mcm10ΔN proteins by sedimentation velocity analytical ultracentrifugation and size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). The full-length protein showed a broad distrubution of low and high sedimenting species indicative of multiple oligomeric states (Figure 9B). The complex nature of the sedimentation profile precluded assignment of precise molecular mass to each peak. Similarly, SEC-MALS analysis of the full-length protein showed a broad elution profile with at least three overlapping but distinct peaks and the majority of the protein existing as the lower molecular weight form (Figure 9C). As with the sedimentation data, the overlapping nature of the peaks only allowed for an estimation – not a definitive assessment – of the respective molecular masses. From the light scattering data,

the three major species were approximately 90.4 kDa (I), 189.3 kDa (II), and 322.7 kDa (III) in size, corresponding to 1, 2, and 3.4 Mcm10 subunits, respectively (the calculated mass from amino acid composition is 95.4 kDa). Although not strictly conclusive, these data are consistent with reports of dimeric and trimeric forms of yeast Mcm10 (Fien and Hurwitz 2006; Eisenberg 2009).

In contrast, Mcm10ΔN formed a single species corresponding unequivocally to a monomeric protein in both experiments. The molecular mass of the major (1.5S) peak from sedimentation velocity (Figure 9B) was calculated to be 68.8 kDa, compared to 70.4 kDa calculated from the amino acid composition. The minor species observed at 2.3S did not increase with the protein concentration (Table 2) and was thus judged to be a contaminant. In addition, Mcm10ΔN eluted as a single, monodispersed species from a size exclusion column with a molecular mass of 75.1 ± 0.8 kDa determined by MALS (Figure 9D). Therefore, deletion of the NTD eliminated self-association of the full-length protein.
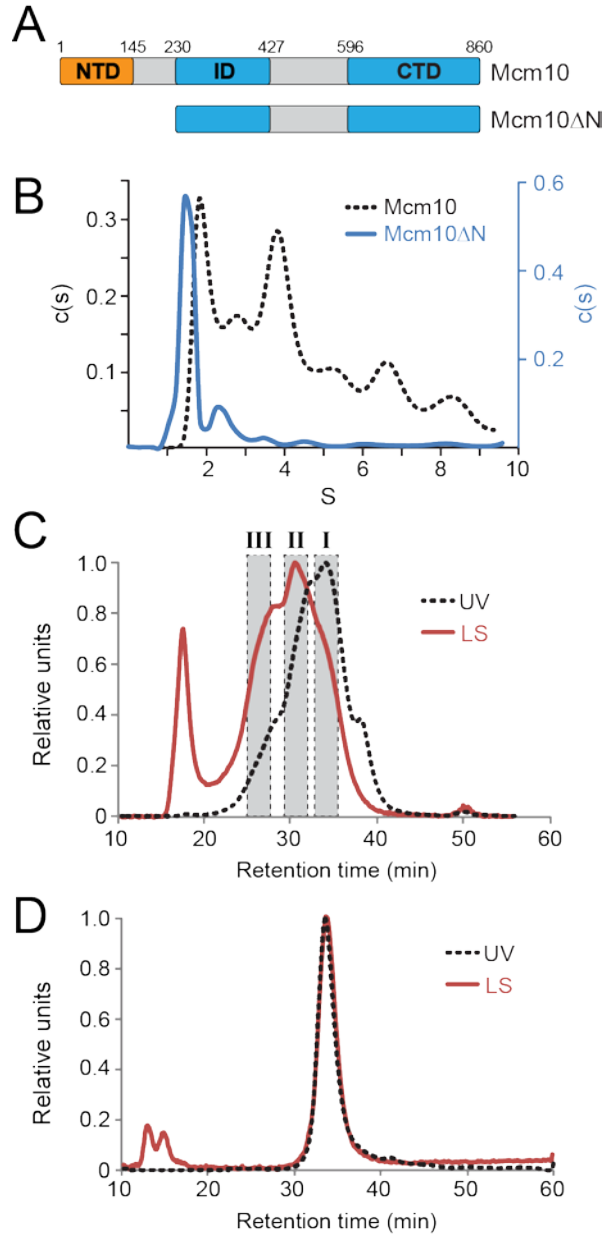
Figure 9. The NTD is necessary for Mcm10 self-association. **A.** Schematic of Xenopus laevis Mcm10 constructs used to study self-association. Theoretical molecular masses are 95.4 kDa (Mcm10) and 70.4 kDa (Mcm10ΔN). **B.** Sedimentation velocity data for full-length Mcm10 (black dotted line) and Mcm10ΔN (blue) as described in Materials and Methods. The molecular mass of the Mcm10ΔN major peak was calculated to be 68.8 kDa. Molecular masses could not be accurately determined from the full-length data. **C,D.** SEC-MALS analysis of Mcm10 (C) and Mcm10ΔN (D). The UV trace is shown as a black dotted line and the light scattering trace is red. **C.** Estimated molecular masses were calculated from the three shaded regions to be 90.4 kDa (I), 189.3 kDa (II), and 322.7 kDa (III). The peak at 18 min corresponds to the void volume. **D.** The molecular mass of Mcm10ΔN was calculated to be 75.1 ± 0.8 kDa.

Table 2. Sedimentation velocity data for Mcm10 constructs.

	pH	Conc (mg/mL)	S	$S_{20,w}$[1]	MW_{calc}[2]	MW_{exp}[3]	# subunits[4]	Peak area
Mcm10								
WT	7.4	1.7	1.98	3.17	95.4	52.8	0.5	19.7
	7.4	1.7	2.87	4.58	95.4	91.6	1.0	13.5
	7.4	1.7	3.96	6.33	95.4	149	1.6	27.9
2D	7.4	1.0	1.76	2.80	95.4	42.2	0.4	30.2
	7.4	1.0	2.66	4.24	95.4	78.8	0.8	28.4
	7.4	1.0	4.19	6.67	95.4	156	1.6	19.7
Mcm10ΔN								
WT	7.4	1.0	1.53	2.99	70.4	68.8	1.0	74.1
	7.4	1.0	2.75	5.38	70.4	166.0	2.4	19.2
	7.4	1.6	1.53	2.99	70.4	71.6	1.0	71.2
	7.4	1.6	2.55	4.99	70.4	152.0	2.2	17.3
	7.4	3.5	1.36	2.66	70.4	71.2	1.0	79.3
	7.4	3.5	2.32	4.55	70.4	160.0	2.3	18.0
NTD								
WT	7.4	0.4	1.62	1.59	16.0	27.5	1.7	76.4
2D	7.4	0.4	1.30	1.26	16.0	13.2	0.8	78.4
4A	7.4	0.4	1.31	1.27	15.8	17.0	1.1	71.4
Coiled-coil								
MBP-CC[95-132]	7.4	0.6	3.48	3.61	45.1	52.4	1.2	55.2
	7.4	0.6	4.41	4.58	45.1	74.8	1.7	42.8
	4.7	0.6	3.48	3.61	45.1	46.0	1.0	7.5
	4.7	0.6	5.17	5.36	45.1	84.1	1.9	77.3
	4.7	0.6	6.84	7.10	45.1	129.0	2.9	14.0
MBP-CC[95-124] WT	7.4	0.7	3.22	3.50	44.3	52.2	1.2	27.3
	7.4	0.7	4.79	5.21	44.3	94.7	2.1	65.8
MBP-CC[95-124] 2D	7.4	0.7	3.15	3.42	44.1	49.1	1.1	96.9
MBP-CC[95-124] 4A	7.4	0.6	3.15	3.41	43.9	45.4	1.0	98.5
MBP	7.4	0.6	3.14	3.46	40.4	43.0	1.1	92.0
	4.7	0.6	3.26	3.55	40.4	42.1	1.0	99.5

[1] Sedimentation coefficient at standard conditions (water at 20 °C).
[2] MW_{calc} is the molecular weight in kDa calculated from the amino acid composition
[3] MW_{exp} is the molecular weight in kDa derived from the ultracentrifugation data
[4] # subunits = MW_{exp}/ MW_{calc}

The structure of the Mcm10 coiled-coil region

The NTD of the vertebrate and *Saccharomyces cerevisiae* Mcm10 orthologs contain a putative coiled-coil (CC) that we previously hypothesized accounts for dimerization of the NTD (Robertson 2008). We tested the ability of this region to dimerize by fusing the peptide corresponding to xMcm10 residues 95-132 to maltose binding protein (MBP) and analyzing by sedimentation velocity ultracentrifugation. MBP alone sedimented as a monomer, in agreement with a previous determination (Marvin 1997). In contrast, two species consistent with monomeric and dimeric forms of MBP-CC were present (Figure 10A,B). The dimeric form of MBP-CC persisted on SDS-PAGE gels even in the presence of high concentrations of reducing agents (Figure 10C), a characteristic of coiled-coils observed in other proteins (Wigge 1998).
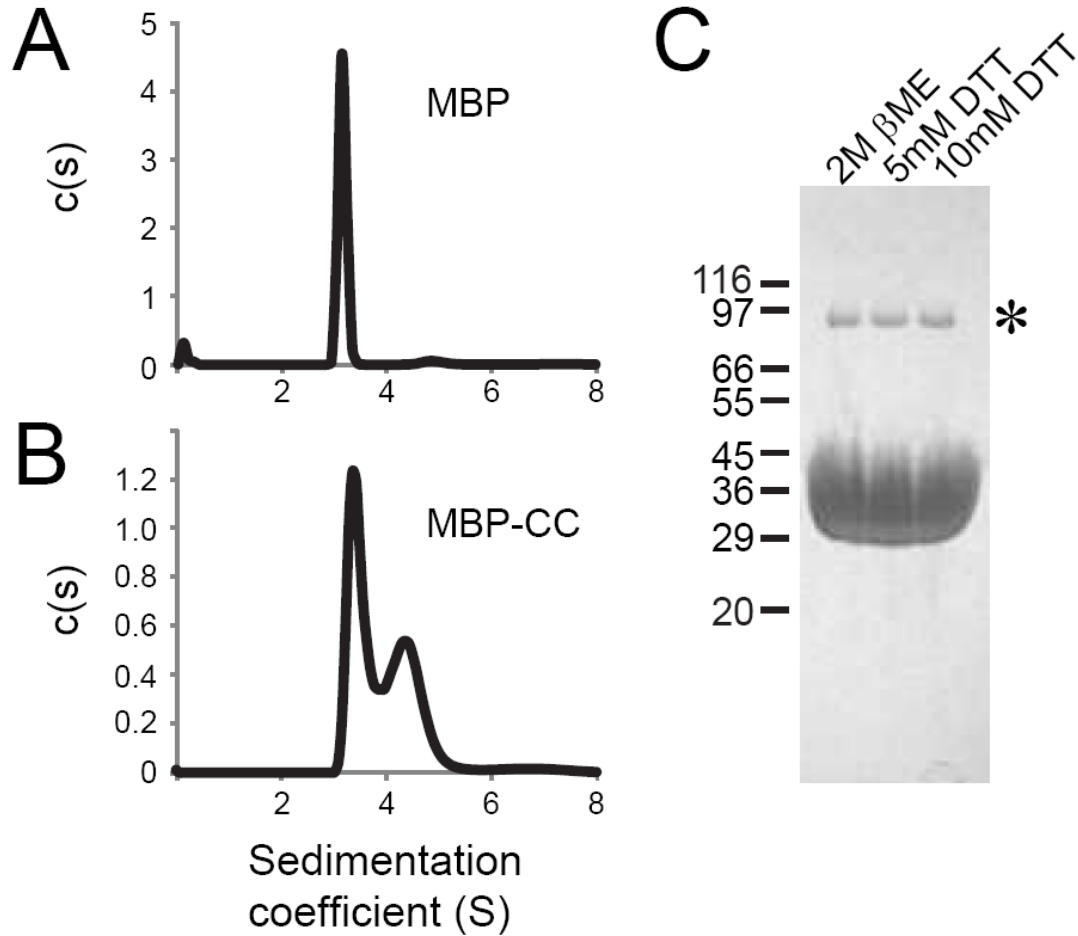
Figure 10. Dimerization of the putative Mcm10 coiled-coil region. **A,B.** Sedimentation velocity profiles of free MBP (A) and MBP-CC$^{95\text{-}132}$ (B) at pH 7.4. Molecular masses derived from the data (Table 2) are 43 kDa (MBP) and 52 and 75 kDa (MBP-CC), corresponding to 1.2 and 1.7 MBP-CC subunits, respectively. **C.** SDS-PAGE of MBP-CC$^{95\text{-}132}$ in the presence of varying amounts of reducing agents. Both bands were confirmed by mass spectrometry to be xMcm10 residues 95-132. The loading buffer in each sample contained 62.5 mM Tris-HCl (pH 6.8), 10% glycerol, 2% SDS (w/v), and bromophenol blue in addition to the reducing agents shown at the top of each lane. The peak marked with an asterisk (*) represents a molecular mass exactly twice that of the calculated mass, and persisted at DTT concentrations as high as 200mM (not shown).

To verify this region of the protein as a *bona fide* CC, we determined the crystal structure of the MBP-CC fusion protein to a resolution of 2.4 Å (Figure 11). The highest quality diffraction data (Table 3) were obtained from a construct spanning xMcm10 residues 95-124 crystallized under low pH conditions. The final model was refined to crystallographic residuals of 16.4% (R_{work}) and 20.5% (R_{free}). Surprisingly, the asymmetric unit consisted of a trimeric assembly with the Mcm10 residues at the center forming a parallel three-helix CC wrapped in a left-handed superhelix (Figure 11A,B). Trimer formation is not a crystallographic artifact, since we observed trimeric and dimeric MBP-CCs in solution under the same (low pH) conditions used for crystallization (Figure 11C). Similarly, we verified that MBP did not influence trimerization since MBP alone is monomeric in solution at the low pH condition (Figure 11D). Thus, the Mcm10 CC has the propensity to form both dimeric and trimeric helical bundles, consistent with our SEC-MALS analysis of the full-length protein (Figure 9C).
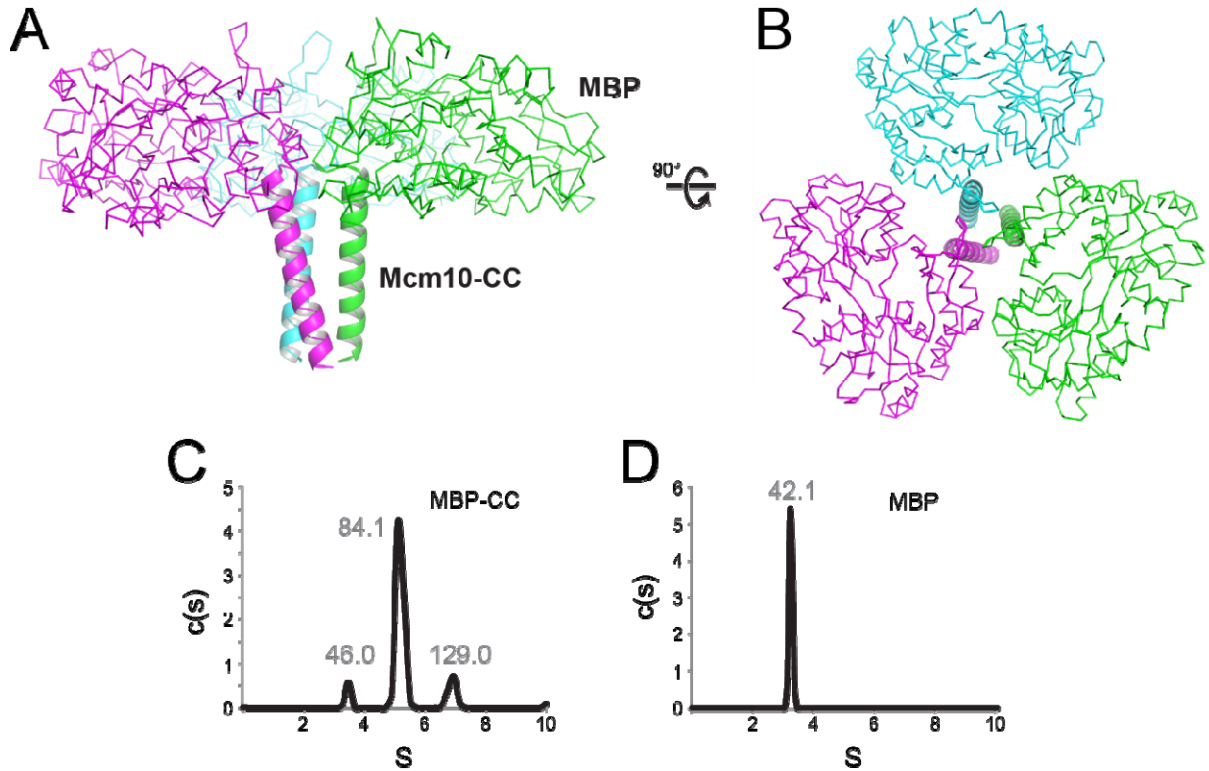
Figure 11. Trimerization of MBP-CC. **A,B.** Crystal structure of the MBP-CC asymmetric unit, with each protomer colored differently. Maltose-binding protein is shown as a C_α-trace, and the xMcm10 coiled coil is depicted as a cartoon ribbon. **C,D.** Sedimentation velocity profiles of MBP-CC[95-132] (C) and free MBP (D) at pH 4.7. Molecular masses (kDa) calculated from the sedimentation data are shown above each peak. The molecular mass of a single polypeptide calculated from the amino acid composition are 45.1 kDa (MBP-CC) and 40.4 kDa (MBP).

Table 3. MBP-CC[95-124] crystallographic data collection and refinement statistics

Data collection	
Wavelength (Å)	0.97872
Space group	C2
Cell dimensions	
a, b, c (Å)	187.4, 116.6, 73.4
α, β, γ (°)	90, 103.5, 90
Resolution (Å)	30.0-2.4 (2.5-2.4)
R_{sym}	0.069 (0.401)
I / σ_I	23.2 (3.6)
Completeness (%)	98.6 (94.4)
Redundancy	7.1 (6.4)
Refinement	
Resolution (Å)	30-2.4
Unique reflections	59,488
R_{work} / R_{free}	0.164 / 0.205
No. of atoms	
Protein/maltose	8,934
Solvent	446
Ave. B-factors	
Protein/maltose	46.4
Solvent	42.8
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.105
PDB ID	4JBZ

Values in parentheses refer to the highest resolution shell.

CCs in other proteins have been shown to exist in multiple oligomeric states, a property largely dependent on the characteristics of the *a* and *d* hydrophobic side chains of the heptad repeat that form the helical interface (Burkhard 2001). For example, two-, three-, and four-stranded CCs in the GCN4 leucine zipper were engineered by mutating the *a* and *d* positions (Harbury 1993). The Mcm10 CC helical region spans Glu98 to Leu122, two invariant residues in the human, frog, mouse, and budding yeast orthologs (Figure 12A), although we did observe the helices to extend to at least Thr125 in lower

resolution structures obtained from a longer 95-132 construct (data not shown).

Most importantly, the high resolution of the structure enabled us to identify the

residues of the CC interface as Leu104, Leu108, Met111, Met115, and Leu118

(Figure 12B). This interface is entirely hydrophobic, with the side chains of each

residue interacting with its equivalents on the other two helices through van der

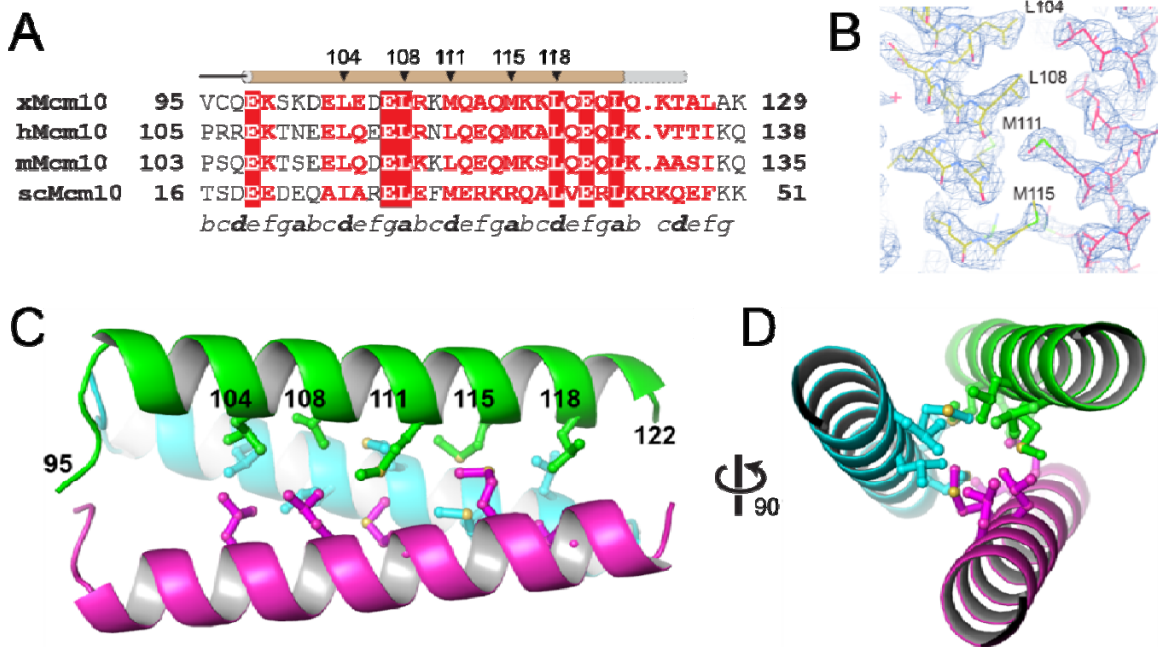Waals packing around a three-fold rotation axis (Figure 12C,D).

Figure 12. Crystal structure of the Mcm10 coiled-coil. **A.** Sequence alignment of the coiled-coil region from *Xenopus laevis* (x), *Homo sapiens* (h), *Mus musculus* (m), and *Saccharomyces cerevisiae* (sc) Mcm10. Red boxes indicate identical amino acids. The predicted heptad repeat pattern is labeled by letters *a-g* at the bottom. The position of the helix is shown schematically at the top (brown, MBP-CC[95-124]; grey, MBP-CC[95-132]). **B.** Composite $2F_o$-F_c omit electron density map (contoured at 1σ) with carbon atoms colored according to protomer. **C.** Crystal structure of *Xenopus laevis* Mcm10-CC[95-124]. Residues at the interface are shown in ball and stick, and the amino acid numbers labeled in black. **D.** View down the helical axis, rotated 90° from the view in *B*.

We expected the same residues to form the interface in a dimeric form of the CC based on other structures with both dimeric and trimeric propensities (Harbury 1993; Harbury 1994; Ciani 2010). The conformation of the Mcm10 CC trimer is virtually identical to the isoleucine zipper variant of the GCN4 CC (Harbury 1994), with only a modest divergence at the N-terminal end (Figure 13B), which likely results from non-hydrophobic heptad repeat *a* and *d* residues (Gln97, Lys101) and/or the MBP tag (Figure 13A). We therefore constructed a model of the dimeric Mcm10 CC using the GCN4 leucine zipper dimer as a template (Figure 13C) (O'Shea 1991). As shown in Figure 13D, the dimer and trimer are related by a simple 60° rotation and 8 Å translation of one helix. The *a* and *d* positions are conserved between the two models, and the conformations of only two side chains (Leu104 and Leu108) needed to be adjusted to avoid steric collision across the dimer interface. Thus, only modest adjustments are required to interconvert between the CC dimer and trimer.

A

```
                                           104     108 111   115 118
Mcm10-CC    VCQEKSKDELEDELRKMQAQMKKLQEQLQK
GCN4 trimer RMKQIEDKIEEILSKIYHIENEIARIKKLIGE
GCN4 dimer  RMKQLEDKVEEILSKNYHLENEVARLKKLVGER
            abcdefgabcdefgabcdefgabcdefgabc
```
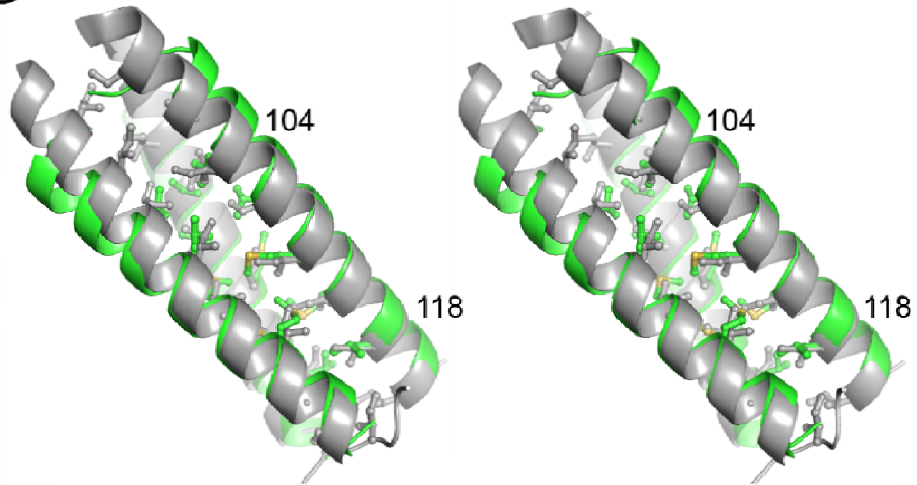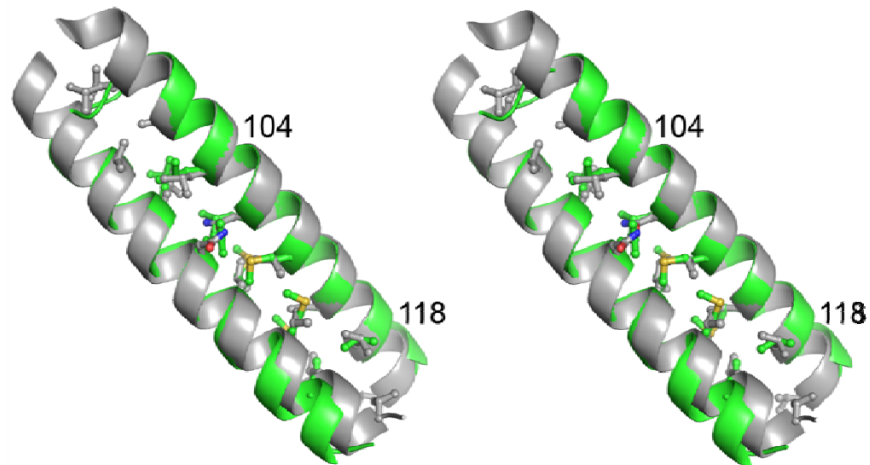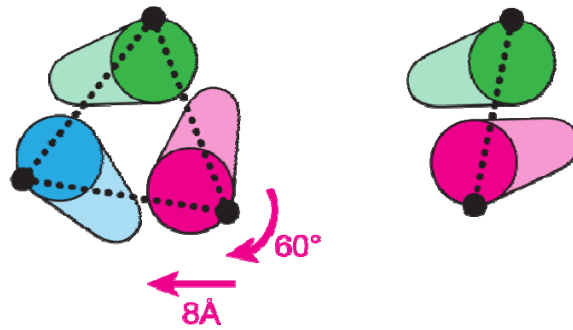
Figure 13. A model of the dimeric Mcm10 CC. **A.** Structure-based sequence alignment between Mcm10 CC and the trimeric and dimeric forms of GCN4 CC. **B.** Stereoview of the Mcm10 CC trimer (green) superimposed onto the isoleucine GCN4 trimer (grey, PDB ID 1GCM). Side chains at the CC interface are shown, with the exception of the N-terminal methionine in GCN4. Two Mcm10 residues are labeled for orientation. The view is rotated 45° clockwise with respect to Figure 3C. **C.** Stereoview of the Mcm10 CC dimer model (green) superimposed on the CGN4 CC (grey, PDB ID 2ZTA). **D.** Schematic showing the relationship between trimer (left) and dimer (right) forms of the CC. The conformation of the dimer can be constructed from the trimer by a 60° rotation and 8Å translation of the magenta helix. The interhelical distances (dashed line) are 15.5 Å (trimer) and 10 Å (dimer).

Mutations in the coiled-coil motif disrupt Mcm10 oligomerization

To validate the crystal structure as representative of a functional CC, we designed mutations aimed at disrupting self-interaction. We introduced electrostatic repulsion at the interface by substituting Leu104 and Leu108 with aspartate to create a L104D/L108D double (2D) mutant. In addition, we eliminated side chains at positions 104, 108, 115, and 118 by alanine substitution to create a L104A/L108A/M115A/M118A quadruple (4A) mutant. Mutations were introduced into the MBP-CC[95-124] and NTD protein constructs and tested for dimerization using sedimentation velocity (Figure 14). Both 2D and 4A mutants disrupted dimerization of the wild-type CC and NTD (Figure 14A,B). Interestingly, replacing only Leu104 and Leu108 with alanine (2A mutant) was not enough to disrupt dimerization (Figure 15A), suggesting that the remainder of the interface is sufficient to hold the CC together.
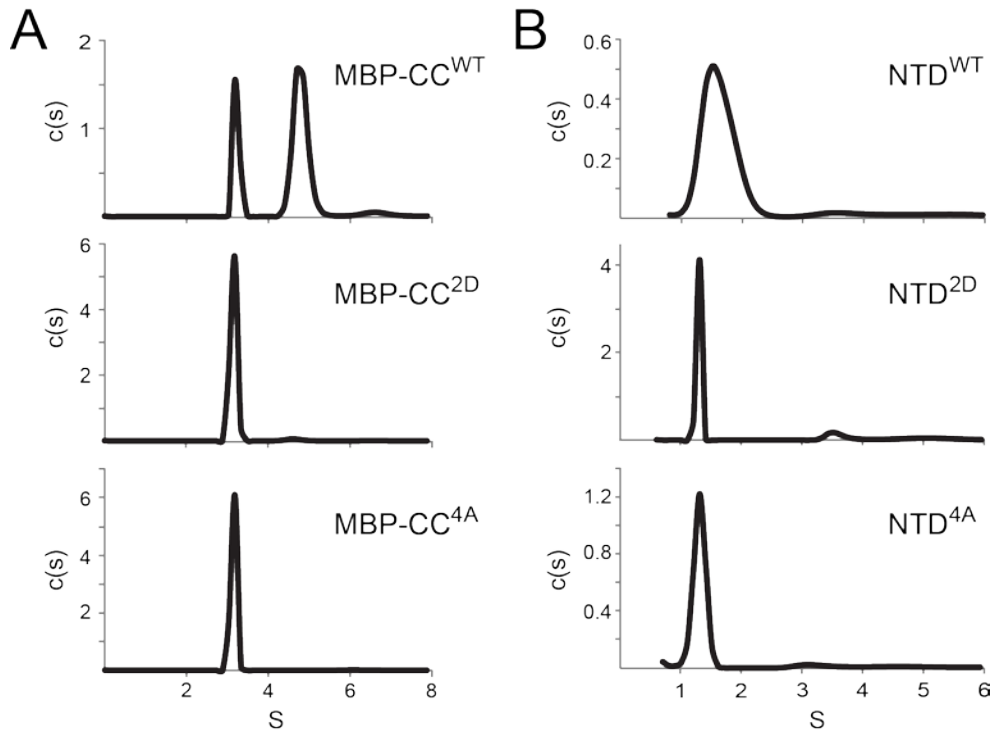
Figure 14. Coiled-coil mutations disrupt CC and NTD dimerization. Sedimentation velocity data for MBP-CC$^{95\text{-}124}$ (**A**) and NTD (**B**) constructs as wild-type (WT) or containing L104D/L108D (2D) or L104A/L108A/M115A/L118A (4A) mutations. Molecular masses corresponding to each peak are reported in Table 2.
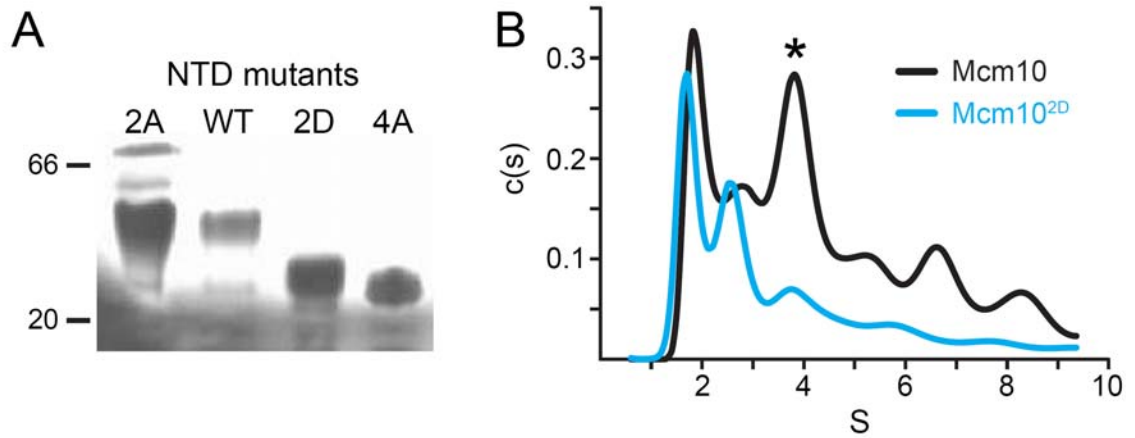
Figure 15. Effect of coiled-coil point mutants on Mcm10 self-association. **A.** Native gel electrophoresis (4-16% Bis-Tris) of the NTD as wild-type (WT), 2A (L104A/L108A), 2D (L104D/L108D), or 4A (L104A/L108A/M115A/L118A). Size markers in kDa are shown to the left. **B.** Sedimentation velocity analytical ultracentrifugation of full-length Mcm10 (black, wild-type; blue, 2D mutant). Data were collected at 4 °C and 42,000 rpm in PBS buffer, 150 mM NaCl, and 0.3 mM TCEP at protein concentrations of 1.6 mg/ml (WT) and 1.0 mg/ml (2D). The estimated massess of these peaks are shown in Table S1. Although the precise masses cannot be accurately determined due to the complex nature of the sedimentation profile, the reduction of the 4S peak (marked with an asterisk) in the monomeric 2D mutant represents a significant difference from the WT.

To confirm that the 2D and 4A mutations impeded Mcm10 dimerization *in vivo*, we conducted a yeast two-hybrid analysis (Figure 16). Full-length xMcm10 as well as the 2D, 4A and Mcm10ΔN mutants were each fused to either a Gal4-binding or -activation domain. The interaction between T-antigen (T-ag) and p53 served as a positive control, whereas combinations of the respective activation domain fusions combined with an empty vector served as negative controls. Plasmid retention was evaluated by spotting cells onto double selection plates lacking leucine and tryptophan. The ability to interact was scored on quadruple selection plates. Full-length xMcm10 displayed strong self-interaction, almost at the level as the binding between T-ag and p53. In contrast, the interaction was eliminated by the 2D and 4A point mutations and the N-terminal deletion construct (Figure 16A). As expected, empty vector controls did not show any viable colonies (Figure 16A). Importantly, the lack of self-association between the respective mutants or the N-terminally truncated protein and full-length xMcm10 was not due to differences in protein expression, as analyzed by Western blot (Figure 16B,C). Since we also observed a significant difference in the sedimentation velocity profile of purified full-length wild-type protein and the 2D mutant (Figure 15B), we conclude that the CC is the primary oligomerization motif in xMcm10.
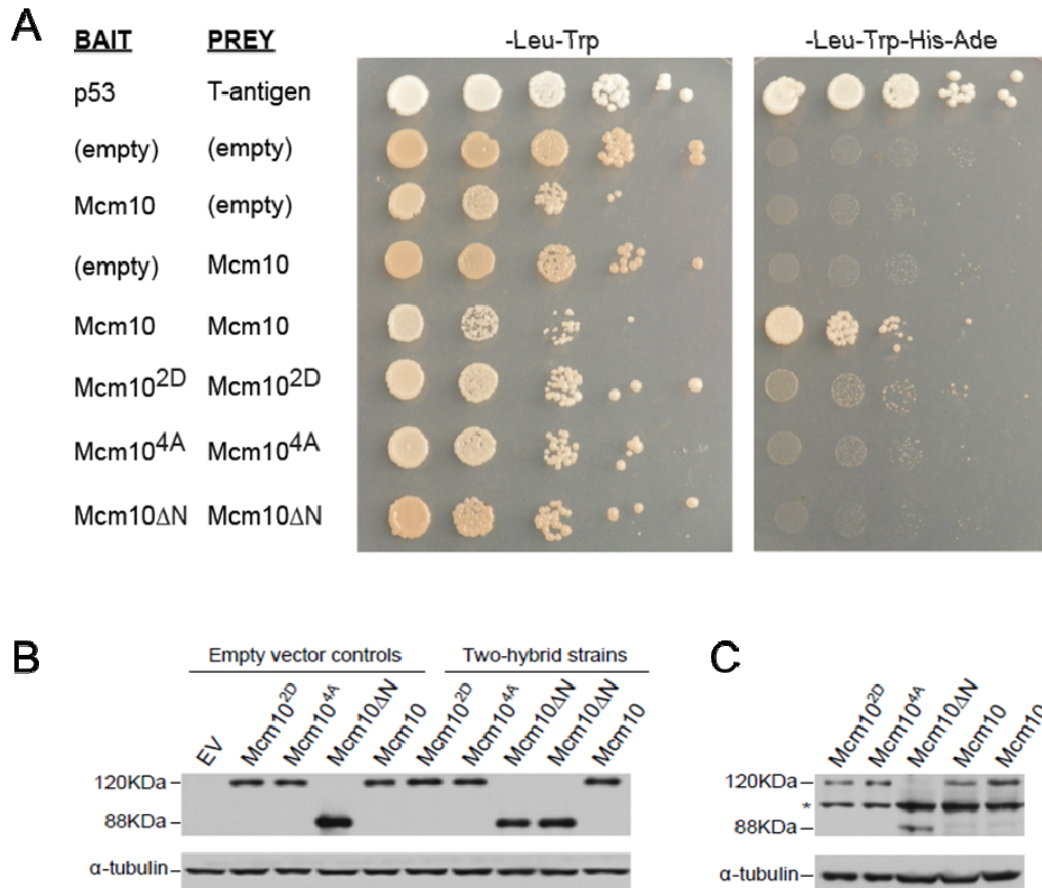
Figure 16. Coiled-coil mutations disrupt Mcm10 self-association. **A.** Bait and Prey proteins were cloned into the yeast-2-hybrid system and co-transformed into AH109 cells. Media lacking leucine and tryptophan (-Leu-Trp) select for maintenance of the both bait and prey plasmids. Media lacking leucine, tryptophan, histidine and adenine (-Leu-Trp-His-Ade) select for the 2-hybrid interaction. Ten-fold serial dilutions were shown. WT Mcm10 self-associates while mutants do not interact. The p53/T-antigen pair is a positive control for interaction, and empty vectors were used as negative controls. **B.** Western blot showing wild-type and mutant xMcm10 protein expression in representative strains. Gal4-AD fusions were detected by a HA-specific antibody. Strains carrying empty vector controls, or Gal4-AD fusion genes and Gal4-BD empty vectors are shown on the left (Empty vector controls). Strains expressing pair-wise combinations of theGal4-AD and Gal4-BD fusion genes are shown on the right (Two-hybrid strains). Full-length xMcm10 and the 2D and 4A mutants ran at an approximate size of 140 kDa, whereas the truncated form of xMcm10 ran at an approximate size of 94 kDa. Tubulin served as a loading control. **C.** Gal4-BD fusions were detected by a Myc-specific antibody. Extracts from the identical two-hybrid strains shown in (B) were loaded in the same order. The asterisk denotes a non-specific band.

Discussion

This work identifies an evolutionarily conserved CC motif in the N-terminus of xMcm10 and provides evidence that it is required for self-association. Our data also strongly suggest that Mcm10 exists in a dynamic equilibrium between multiple oligomeric states, which helps to explain the disagreement in the literature regarding the number of subunits. We observed a broad distribution of states of the full-length protein using two quantitative approaches, and consistently found the presence of both dimeric and trimeric species. It is striking that in addition to estimates of dimer and trimer formation of full-length xMcm10 by SEC-MALS, we detected a mixture of dimeric and trimeric forms of the isolated CC region, with dimers predominating in solution and a trimer in crystals.

The propensity of the Mcm10 CC to form multiple states can be explained by the particular CC sequence, since the rules governing the number of CC subunits as a function of the amino acids at the *a* and *d* positions within the heptad repeat is well understood (Harbury 1993; Burkhard 2001; Woolfson 2005; Ciani 2010). Inclusion of methionine at the *a* position in the human, mouse, and frog Mcm10 CC (Met115 in our structure) likely destabilizes the dimer and would even favor parallel tetramers and pentamers (Woolfson 2005). This raises the possibility that the Mcm10 CC could accommodate higher order oligomers, formed either as a simple association between the helices or as more complex patterns such as a trimer of dimers (Chan 1997; Moutevelis 2009; Spinola-Amilibia 2011). Regardless of the oligomeric state, the residues lining the supercoil interface would remain the same. Indeed, our data indicate that the *a*

and *d* residues identified in the trimer crystal structure are important for dimerization of the CC and the NTD *in vitro* and the self-association of the full-length protein *in vivo*.

The existence of dimers and trimers implies that the Mcm10 CC is metastable and therefore its oligomeric state is sensitive to environmental factors. In support of this, the Mcm10 CC trimer is stabilized at lower pH. Interestingly, pH dependent CC switches are important biological mechanisms by which proteins change conformation to drive various processes (Burkhard 2001). For example, viral glycoproteins adopt trimeric CCs in response to pH as a mechanism to fuse viral and cellular membranes (Carr 1993; Chan 1997; Gibbons 2000). In fact, the crystal structure of the human T cell leukemia virus type 1 transmembrane ectodomain, determined as a MBP fusion, formed a parallel trimeric CC required for proper function (Center 1998; Kobe 1999), further validating the importance of our trimeric MBP-Mcm10-CC structure. In addition, CC folding and remodeling in response to other environmental factors, including temperature and effector molecules is a general phenomenon (Noelken 1964; Lowey 1965; O'Shea 1992; O'Shea 1993; Dutta 2001; Dutta 2003; Spinola-Amilibia 2013). On the basis of these examples and consistent with our data, we speculate that the Mcm10 CC exists mainly as an intrinsically disordered monomer or as a CC dimer, and has the propensity to attain other multimeric configurations in response to its environment.

It is intriguing to speculate that Mcm10 may adopt different oligomeric states to perform multiple roles during DNA replication. For example, higher-order

oligomers may be used for sequestering Mcm10 at the replication fork. Upon pre-RC activation and origin melting, Mcm10 may reform as a dimer or trimer as DNA is denatured and replication factors recruited to the emerging fork. scMcm10 was reported to exhibit differential packing on ssDNA versus dsDNA (Eisenberg 2009), suggesting that a change in the Mcm10 conformation or its oligomeric state could facilitate strand separation. In this context, it is noteworthy that Mcm10 binds ssDNA with a 3-5-fold preference over dsDNA (Fien 2004; Robertson 2008; Eisenberg 2009). Oligomerization on ssDNA might thus assist in the initial unwinding step and aid what has been termed "helicase activation" (Thu and Bielinsky 2013) but may very well just be the coordinated stabilization of short stretches of unwound DNA after the separation of Mcm2-7 dimers (Gambus 2006; Remus 2009).

Mcm10's modular architecture and lack of enzymatic activity suggest that it serves as a scaffold to orchestrate protein and DNA interactions within the replisome. Self-association would provide multiple points of contact between replication factors and DNA (Du 2012). Indeed, Mcm10 is involved in multiple interactions, including but not limited to the replication and checkpoint clamps, PCNA and 9-1-1 (unpublished results and ref. Das-Bradoo 2006), and pol α (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007; Robertson 2008; Warren 2009). Protein-protein interactions could be mediated by the CC directly, similar to the interaction between Cdt1 and geminin (Lee 2004; Saxena 2004; Thepaut 2004). Alternatively, dimerization could facilitate molecular interactions and recruiting proteins to the origin simply

by increasing the number of possible binding sites on Mcm10 (Du 2012). For example, the ID and CTD each bind DNA and pol α and could therefore be involved in a molecular hand-off, whereby Mcm10 is anchored to DNA via the ID while binding pol α at the CTD, and *vice versa* (Warren 2009). Additional Mcm10 subunits would enhance these interactions by increasing the number of ID and CTD present. Similarly, a parallel Mcm10 dimer could couple events on the leading and lagging strands or physically tether the helicase and pol α (Ricke and Bielinsky 2004; Zhu 2007; Lee 2010) while retaining the polarity necessary for fork progression. This would also explain why loss of the first 100 residues of scMcm10 confers such a strong sensitivity to hydroxyurea in the absence of the 9-1-1 clamp (Alver and Bielinsky, unpublished results).

Taken together, the dimerization or trimerization of xMcm10 in the absence of DNA reported here is consistent with previous work on spMcm10 (Fien and Hurwitz 2006), and the observation that three subunits of scMcm10 are bound to short ssDNA oligonucleotides, although these latter complexes were not shown to have the three-fold symmetry revealed in our crystal structure (Eisenberg 2009). We did not find any evidence for Mcm10 hexamers, which were previously observed by EM of the human protein (Okorokov 2007). As discussed above we do not rule out a trimer of dimers, although this would not be consistent with the six-fold symmetrical EM structure reported.

CHAPTER IV

DISCUSSION AND FUTURE DIRECTIONS

Role of Mcm10 in DNA replication

Despite decades of research, the precise function of Mcm10 is not well understood. The research presented here brings us one step closer towards understanding its role. The oligomeric state of Mcm10 is likely to be important to its function. Here I propose several models for how Mcm10 self-association helps with its function. First, Mcm10 might self-associate to recruit proteins to DNA using a molecular hand-off mechanism. Second, Mcm10 might use different oligomeric states to regulate Mcm10-protein interaction. Third, the multiple changing states of Mcm10 may help with replisome assembly.

Our data and other existing research support the hypothesis that Mcm10 functions as a scaffold protein to orchestrate protein and DNA interactions within the replisome during DNA replication. As an example of its role in protein recruitment to DNA, Mcm10 is required for the activation of the CMG complex at origins. Without Mcm10, the origin unwinding is blocked (van Deursen 2012); RPA and pol α cannot be recruited (Kanke 2012). Mcm10 is also required for chromatin loading of And-1, an essential DNA replication initiation factor, and the stability of p180 in mammalian cells (Zhu 2007). Mcm10 might use a molecular hand-off mechanism in which the ID is anchored to DNA while the CTD binds to pol α, and *vice versa*. But more importantly, self-association might help protein

recruitment to DNA using the molecular hand-off mechanism. For example, an Mcm10 dimer could use one copy of Mcm10 to primarily interact with DNA to provide an anchor, while the other copy binds to Mcm10-interacting proteins such as pol α. In this scenario, when two copies of Mcm10 self-associate through the NTD, one copy of ID+CTD binding to DNA would provide a high-affinity DNA-binding platform, while the other copy of ID+CTD exposes a free binding platform to recruit proteins to DNA. A dimer model could also be compatible with the idea that one copy of Mcm10 is responsible for activities on the leading strand (such as DNA binding), and the other copy on the lagging strand performs other activities (such as protein recruitment). Such coordination could help recruit important proteins to DNA during replication, while maintaining the polarity of the fork in DNA replication.

A second possibility is that Mcm10 uses different oligomeric states to regulate its interaction with other proteins. Mcm10 has an extensive network of interacting proteins, such as pol α (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007), ORC (Izumi 2000; Hart 2002), Mcm2-7 (Merchant 1997; Homesley 2000; Hart 2002; Apger 2010), RecQL4 (Xu 2009), And-1 (Zhu 2007), and PCNA (Das-Bradoo 2006). These proteins may need to interact with Mcm10 only during a certain stage during DNA replication initiation. Their timely dissociation from Mcm10 could be tightly regulated to prevent replication errors. Some proteins might interact with a certain state of oligomeric Mcm10, and then disassociate from another oligomeric Mcm10 form. Mcm10 might therefore oscillate between monomer, dimer, and trimer forms, for

example, to regulate protein binding. These on and off states could ensure efficient transitions between different Mcm10-protein interactions.

A third model for the biological significance of Mcm10 self-association involves multiple changing oligomeric states as needed to help with replisome assembly. Our results showed a trimeric crystal structure of the NTD, and also further showed that both dimer and trimer forms of NTD exist in solution. The exact oligomeric state of full-length Mcm10 could not be determined, because of the poor protein sample quality despite multiple attempts to improve protein purity. Although not conclusive, the SEC-MALS and AUC data (see Chapter III) suggested that multiple oligomeric states might exist. If this is the case, Mcm10 could adopt different oligomeric forms to suit different requirements in replisome assembly. For example, Mcm10 might first use a higher oligomerization state (e.g. dimer of trimers) to have each copy of Mcm10 interacting with one subunit of Mcm2-7 to facilitate DNA unwinding at the origin. Alternatively, an Mcm10 dimer might use specific structural features to help replisome formation. In a model in which the Mcm2-7 double hexamers dissociate upon origin firing to form two bidirectional replication forks (Remus 2009), Mcm10 might mediate this process. The interaction between the two hexamers of Mcm2-7 is partly mediated by a zinc ribbon motif, which is structurally homologous to the CCCC motif in Mcm10-CTD (Robertson 2010). Mcm10 might mimic the zinc ribbon of Mcm2-7 double hexamer and facilitate its dissociation. In yeast Mcm10 that lacks CTD (and thus CCCC motif), another mechanism might be needed to facilitate Mcm2-7 double hexamer dissociation. At a later step in replication initiation,

Mcm10 might then adapt to a lower oligomeric state (e.g. dimer or trimer) to recruit proteins (such as pol α) to DNA using a molecular hand-off mechanism (Du 2012). Mcm10 might change between monomer, dimer, trimer and other forms to regulate Mcm10-protein interactions during various stages of replication initiation. It is also conceivable that higher order Mcm10 oligomers could be inactive as scaffolds, and that lower order Mcm10 oligomers or Mcm10 monomers are released to perform their individual recruiting and scaffolding functions. These speculative scenarios are derived from the current *in vitro* study. It will be ideal to investigate this question again with *in vivo* studies.

It is also important to note that evidence has suggested that Mcm10 also has a role in DNA replication elongation (Pacek 2006). Mcm10 interacts with pol α and prevents its degradation by the proteosome to stabilize it through the cell cycle. Mcm10 increases the binding of pol α to DNA and stimulates its primase activity, and thus seems to be a co-factor for pol α. Pol α undergoes cycles of chromatin association and dissociation during Okazaki fragment synthesis, and therefore Mcm10 might also bind only transiently to the fork. It is interesting to speculate that self-association of Mcm10 might help this transient binding of pol α to DNA by changing between different oligomeric states, while pol α might only be recruited by a certain Mcm10 oligomeric form. It is also possible that pol α is bound to Mcm10 throughout the process, but its enzymatic activity is regulated by different Mcm10 oligomers, i.e. pol α is active when bound to a certain Mcm10 oligomeric form, but not others. Mcm10 also travels with the replication fork during elongation. A *Xenopus* egg extract study targeted replication factors

including Mcm10 (Pacek 2006). They used ChIP assay of plasmid templates that contained a biotinylated nucleotide. Adding streptavidin caused replication fork arrest without uncoupling DNA polymerases from the helicase. Mcm10 co-localized with DNA polymerases and Mcm2-7 subunits at the biotin-streptavidin site, which supports that Mcm10 is a component of the elongation machinery (reviewed in Thu and Bielinsky 2013).

It is interesting to speculate how Mcm10 is able to perform these multiple functions during replication, from DNA and protein binding to coordinating enzymatic activities. One possibility is that there are different "pools" of Mcm10, each subset performing one specific function. It will be very informative to investigate how various Mcm10 pools are formed and how the cell recognizes the difference. It can be a result of the coordination of conformational changes, protein expression levels through the cell-cycle, and the overall cellular spatial and temporal environment. Another possibility is that there are some other factors to trigger the changes between its different roles. This could be different oligomeric states, post-translational modifications, or the existence of other upstream trigger proteins. Sophisticated regulatory mechanism will have to exist to maintain the proper balance and accuracy.

DNA replication and diseases

Chromosome instability syndromes are a group of diseases developed from DNA replication defects. Known diseases include Ataxia telangiectasia (caused by lack or inactivation of the ATM protein kinase), Ataxia telangiectasia-

like disorder (deficiency of the human Mre11 protein), Bloom syndrome (mutations in the *BLM* gene), Fanconi anaemia (15 genes related, including BRCA1/BRCA2, FANCA, etc.) (D'Andrea 2010), and Nijmegen breakage syndrome (defective Nbs1 protein) (Taylor 2004). These diseases result in predisposition to cancer because of various gene rearrangements or mutations. Moreover, when depleted of replication factors, normal cells and cancer cells respond differently. Cancer cells are more sensitive and usually undergo cell death, while normal cells do not. This difference makes replication factors possible targets for cancer therapy. Recent studies suggest that cell cycle and checkpoint control may be coupled to the regulation of other biological cycles such as the circadian cycle or metabolic cycle, proposing that replication factors might have other functions besides DNA replication (reviewed in Masai 2010). Better characterization of Mcm10, an important factor in DNA replication, will likely improve our understanding of related diseases.

Coiled-coil

Coiled coils are major oligomerization motifs in proteins (Burkhard 2001). Coiled coil motifs are highly versatile in form, consisting of 2-7 alpha helices (Liu 2006); and often have "tailored" oligomeric states that correspond to biological functions of proteins (Burkhard 2001). Coiled coils are composed of heptad repeats with hydrophobic residues in the *a* and *d* positions. Coiled coils are mainly stabilized by a distinctive "knobs-into-holes" packing of the apolar side chains into a hydrophobic core (reviewed in Walshaw 2003). The characteristics

of *a* and *d* hydrophobic residues side chains have significant impact on the oligomeric states of coiled coil. In a classic example, the GCN4 leucine zipper could be switched between two-, three-, and four-strands by mutating these residues (Harbury 1993).

The stability of coiled-coil is pH-dependent; low pH stabilizes the coiled coil structure (Noelken 1964; Lowey 1965; O'Shea 1992; Zhou 1992; Carr and Kim 1993; Gibbons 2000; Dutta 2001). This has important biological implications. For example, enveloped viruses such as influenza, Ebola or HIV have coiled-coil-containing surface glycoproteins to fuse their membrane coats with cellular membranes to import genomes to cells. During influenza infection, hemagglutinin refolds at low pH and the membrane lipid bilayers are rearranged (Burkhard 2001). Besides pH and the hydrophobic residues at *a* and *d* positions, another study showed that the *b* and *c* heptad positions are also important to the stability of coiled-coil (Fairman 1996).

The coiled-coil region of xMcm10 formed a trimer in the crystallization condition at pH 4.6. Our AUC analysis further showed the existence of the monomer, dimer, and trimer forms at low pH; while at physiological pH 7.4, only monomer and dimer were observed. The trimerization therefore is a low-pH effect (Figure 10, 11). Interestingly, another study on a *de novo*-designed synthesized peptide showed that dimer was preferred over the trimer at lower pH. There are two explanations proposed. One is that the dimeric state has a larger surface-to-volume ratio and thus lower electrostatic repulsions than the trimer. The other possibility is the preferential packing of protonated Glu side-

chains at *a* and *d* positions against the apolar residues (Dieckmann 1998). The pH-dependent oligomeric states of xMcm10 lead us to speculate that although the exact oligomeric form was not determined for FL protein, Mcm10 might be prone to form a dimer and stabilized by replication proteins at physiological pH in cells. This would be consistent with several dimeric Mcm10 models outlined at the beginning of this chapter.

It is also interesting to note that this coiled-coil interaction observed within NTD in Mcm10 could be at least partially disrupted with high salt in solutions *in vitro.* An example comparison of the effect of salt concentration is shown in Figure 17. Oligomerization of MBP-CC[95-124] and MBP-CC[95-132] were each tested by AUC at both pH 4.7 (crystallization condition) and pH 7.4 (physiological pH). To keep the protein stable during overnight dialysis at a pH (4.7) close to their pI (5.1), high salt (400mM NaCl) was added to keep the protein from precipitating out. Results showed that low pH tends to stabilize the trimer, and high salt weakens the oligomerization.
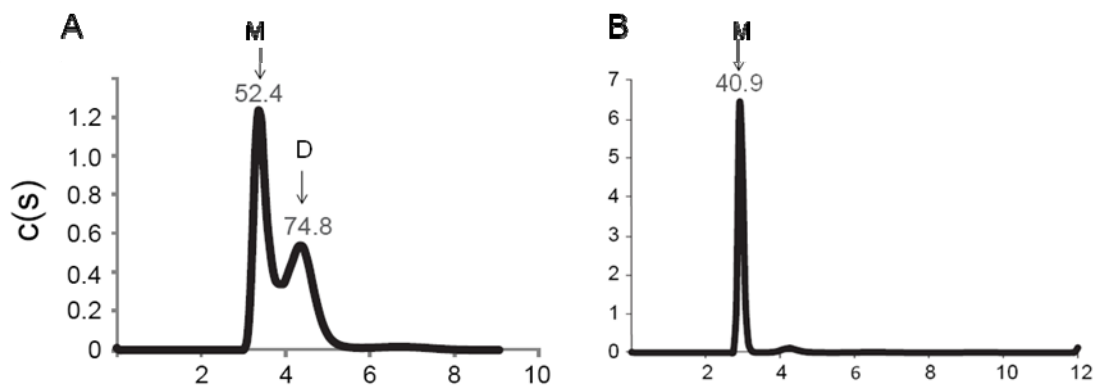
Figure 17. High salt concentration disrupts coiled-coil oligomerization. **A.** AUC of MBP-CC$^{95-132}$ in pH 7.4 buffer with 150mM NaCl. **B.** AUC of MBP-CC$^{95-132}$ in pH 7.4 buffer with 400mM NaCl. The arrows indicate monomer and dimer. *M*, monomer. *D*, dimer.

Future directions

The research presented in this thesis expanded our knowledge of the structural basis for oligomerization of Mcm10 and the interactions of Mcm10 with other players in replication initiation. However, more work needs to be done to understand the precise function of Mcm10 in DNA replication. To test the functional significance of Mcm10 oligomerization, future experiments should be focused on introducing the mutants studied here *in vivo*. There are several model systems that can be used for this purpose. First, the xMcm10 mutants could be introduced into *Xenopus* egg extract. Proteins interacting with Mcm10 can be identified in a pull-down assay. If Mcm10 oligomerizes to regulate protein binding, some interacting proteins may not be pulled-down with the Mcm10 mutants. A possible complication may be that if the interactions are weak or transient, as found in some cases between interacting proteins in DNA replication initiation (Zou and Stillman 2000; Bauerschmidt 2007; Bruck 2011), the protein-

protein interaction would not be easily identified biochemically. Another experiment in the *Xenopus* egg extract would be a ChIP assay to investigate whether Mcm10 oligomerization mutants bind to the chromatin. Since Mcm10 ID and CTD bind DNA, we would expect mutated Mcm10 could also bind DNA. However, if oligomerization Mcm10 mutants do not bind DNA, it suggests that Mcm10 loads onto the chromatin as an oligomer. If protein-protein interactions are not detectable biochemically in *Xenopus* egg extract, we would perform a complementation assay in human cells given the high sequence similarity between *Xenopus* and human Mcm10. Since Mcm10 is essential, we would first introduce *Xenopus* Mcm10 into human cells, and then use RNAi to knock-down expression of endogenous human Mcm10. We would then use pull-down and ChIP assays as above to investigate the consequences of Mcm10 oligomerization mutants. A third system we can use is yeast. Yeast mutations designed based on my *Xenopus* data did not disrupt oligomerization by yeast-2-hybrid (Anja Bielinsky, unpublished), suggesting yeast Mcm10 may utilize additional domains to self-associate. However, one experiment would be to replace endogenous yeast Mcm10 with either wild-type or mutant forms of the *Xenopus* Mcm10 by plasmid shuffling in an Mcm10 deletion strain or through use of a temperature-sensitive or degron allele of *MCM10*, followed by assaying cell viability under normal conditions. For plasmid shuffling, we may want to start with a high-copy yeast vector to achieve high yield of xMcm10 protein, but may switch to a low-copy vector if too much *Xenopus* Mcm10 is toxic to yeast cells. The temperature-sensitive allele is also convenient in that we can raise the

temperature to inactivate yeast Mcm10 and test the functions of mutant *Xenopus* Mcm10. Yeast cells with wild-type Mcm10 should show normal growth. If the *Xenopus* Mcm10 can complement the yeast Mcm10, those cells should also show normal growth. This could serve as a tool for studying the functional consequences of Mcm10 oligomerization mutants. We can then perform the ChIP or pull-down experiments outlined above after expressing the oligomerization-defective *xMCM10* alleles in yeast. We can also synchronize the wildtype and mutants, and study the DNA replication process on those mutant cells. Those might fail to enter S phase, or have prolonged S phase.

Besides oligomerization and function of Mcm10, another interesting question is how the individual domains coordinate binding to DNA. Previous studies identified both ID and CTD as DNA binding domains. In fact, a construct spanning both the ID and CTD (lacking the NTD) has a tighter binding to DNA by 10-fold than either domain alone. Therefore, it is very likely that ID and CTD bind to each other in a coordinated manner. Yeast Mcm10 binding data showed Mcm10s cooperatively bind to dsDNA in an end-to-end manner, supporting our hypothesis (Eisenberg 2009). There are three possible mechanisms for ID and CTD to bind DNA (Figure 18). One is that ID and CTD both bind to the same region of ssDNA, which is sandwiched in between the two binding sites (globular). Another model would be that ID and CTD stack end to end on the ssDNA, or even to the point of overlapping. The third possibility is that ID and CTD are separated from each other upon DNA binding. If ID and CTD coordinate DNA binding, there could be a significant conformational change of Mcm10 upon

DNA binding. Whether ID and CTD restructure upon DNA binding could be investigated by measuring hydrodynamic properties of the protein using biophysical techniques such as SEC-MALS and small angle X-ray scattering (SAXS). High-resolution structures of ID+CTD with and without DNA can be pursued by X-ray crystallography. Critical residues in DNA binding could be examined by mutagenesis, and corresponding *in vivo* functional analysis of the mutants needs to be performed.
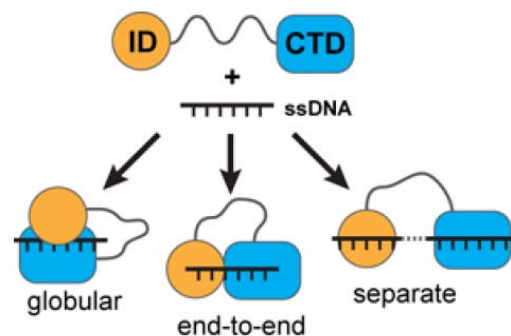

Figure 18. Possible ssDNA binding modes by ID+CTD.

It would also be interesting to know about the preferred substrate in binding. Fluorescence anisotropy and electrophoretic mobility shift assays (EMSA) have determined that the preferred substrate length for ID was 5-10 nt. And for CTD, ~15 nt would be sufficient for ssDNA binding with maximal binding affinity achieved at 20-25 nt (Robertson 2008; Warren 2008). Having characterized the DNA binding length for the individual domains, it would be very informative to investigate the optimal DNA binding length for the full DNA binding site as a whole. Moreover, Mcm10 interacts with pol α, which contains a primase function. Pol α synthesizes a small RNA as the template for DNA replication, which leads to the possibility of Mcm10 also interacting with RNA. Fluorescence

anisotropy and EMSA can be used for optimal substrate determination with various lengths of fluorescein-labeled ssDNA (e.g. 15-45 nt), labeled RNA primer and ssDNA overhang. This RNA substrate study will shed light on the details of Mcm10 interacting with pol α, as well as its cellular function.

To further define the structural basis for the function of Mcm10 in the replication machinery, structures of full-length Mcm10 and possibly Mcm10-pol α complex will be very informative. Efforts on obtaining the structure of Mcm10 by X-ray crystallography have not been successful, possibly due to the flexibility within the protein. Because Mcm10 binds to other proteins such as pol α, co-crystallization could possibly provide more stability. Another technique can be used is SAXS, which requires no crystal- only proteins in solution and is suitable to analyze big protein complexes, although the resolution is lower than crystallography. SAXS can provide initial low-resolution models, high-resolution Mcm10 domain structures can be then fit into the molecular envelop.

Summary

This work examines the role of the N-terminal domain (NTD) in Mcm10 structure and self-assembly. Specifically, NTD was identified to be responsible for full-length xMcm10 oligomerization. The structure of the coiled-coil motif within NTD was determined to be a homotrimer by X-ray crystallography. Both dimer and trimer were observed in solution by AUC. Mutations of the interacting residues within NTD disrupted oligomerization of NTD, as well as full-length Mcm10. A dimer/trimer coiled coil conversion model was proposed. This study

puts us one step further in understanding Mcm10 oligomerization, and thus its role in DNA replication initiation. Future studies to better characterize the coordination of ID and CTD in DNA binding, the complex structure of FL Mcm10 and its interacting partners, and the investigation of the functional significance of Mcm10 will help us better understand Mcm10 and DNA replication. This research will shed light on drug discovery and treatment of DNA-replication related diseases, including cancer.

REFERENCES

Adachi, Y., Usukura, J. and Yanagida, M. (1997) A globular complex formation by Nda1 and the other five members of the MCM protein family in fission yeast. *Genes Cells* **2**(7): 467-79.

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**(Pt 11): 1948-54.

Altman, A. L. and Fanning, E. (2004) Defined sequence modules and an architectural element cooperate to promote initiation at an ectopic mammalian chromosomal replication origin. *Mol Cell Biol* **24**(10): 4138-50.

Anglana, M., Apiou, F., Bensimon, A. and Debatisse, M. (2003) Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. *Cell* **114**(3): 385-94.

Aparicio, O. M., Weinstein, D. M. and 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**(1): 59-69.

Aparicio, T., Ibarra, A. and Mendez, J. (2006) Cdc45-MCM-GINS, a new power player for DNA replication. *Cell Div* **1**: 18.

Apger, J., Reubens, M., Henderson, L., Gouge, C. A., Ilic, N., Zhou, H. H. and Christensen, T. W. (2010) Multiple Functions for Drosophila Mcm10 Suggested Through Analysis of Two Mcm10 Mutant Alleles. *Genetics* **185**: 1151-1165.

Arezi, B. and Kuchta, R. D. (2000) Eukaryotic DNA primase. *Trends Biochem Sci* **25**(11): 572-6.

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* **8**(1): 84-90.

Arias, E. E. and Walter, J. C. (2007) Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev* **21**(5): 497-518.

106

Aves, S. J., Tongue, N., Foster, A. J. and Hart, E. A. (1998) The essential schizosaccharomyces pombe cdc23 DNA replication gene shares structural and functional homology with the Saccharomyces cerevisiae DNA43 (MCM10) gene. *Curr Genet* **34**(3): 164-71.

Barry, E. R. and Bell, S. D. (2006) DNA replication in the archaea. *Microbiol Mol Biol Rev* **70**(4): 876-87.

Bauerschmidt, C., Pollok, S., Kremmer, E., Nasheuer, H. P. and Grosse, F. (2007) Interactions of human Cdc45 with the Mcm2-7 complex, the GINS complex, and DNA polymerases delta and epsilon during S phase. *Genes Cells* **12**(6): 745-58.

Bell, S. P. and Dutta, A. (2002) DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**: 333-74.

Blow, J. J. and Dutta, A. (2005) Preventing re-replication of chromosomal DNA. *Nat Rev Mol Cell Biol* **6**(6): 476-86.

Bochkarev, A., Pfuetzner, R. A., Edwards, A. M. and Frappier, L. (1997) Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385**(6612): 176-81.

Bochkareva, E., Korolev, S., Lees-Miller, S. P. and Bochkarev, A. (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *Embo J* **21**(7): 1855-63.

Bochman, M. L. and Schwacha, A. (2007) Differences in the single-stranded DNA binding activities of MCM2-7 and MCM467: MCM2 and MCM5 define a slow ATP-dependent step. *J Biol Chem* **282**(46): 33795-804.

Bochman, M. L. and Schwacha, A. (2008) The Mcm2-7 complex has in vitro helicase activity. *Mol Cell* **31**(2): 287-93.

Bochman, M. L. and Schwacha, A. (2009) The Mcm complex: unwinding the mechanism of a replicative helicase. *Microbiol Mol Biol Rev* **73**(4): 652-83.

Brewer, B. J. and Fangman, W. L. (1993) Initiation at closely spaced replication origins in a yeast chromosome. *Science* **262**(5140): 1728-31.

Bruck, I. and Kaplan, D. L. (2011) Origin single-stranded DNA releases Sld3 protein from the Mcm2-7 complex, allowing the GINS tetramer to bind the Mcm2-7 complex. *J Biol Chem* **286**(21): 18602-13.

Bullock, P. A., Seo, Y. S. and Hurwitz, J. (1989) Initiation of simian virus 40 DNA replication in vitro: pulse-chase experiments identify the first labeled

species as topologically unwound. *Proc Natl Acad Sci U S A* **86**(11): 3944-8.

Bullock, P. A., Seo, Y. S. and Hurwitz, J. (1991) Initiation of simian virus 40 DNA synthesis in vitro. *Mol Cell Biol* **11**(5): 2350-61.

Burgers, P. M. (1998) Eukaryotic DNA polymerases in DNA replication and DNA repair. *Chromosoma* **107**(4): 218-27.

Burgers, P. M. (2009) Polymerase dynamics at the eukaryotic DNA replication fork. *J Biol Chem* **284**(7): 4041-5.

Burkhard, P., Stetefeld, J. and Strelkov, S. V. (2001) Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol* **11**(2): 82-8.

Carr, C. M. and Kim, P. S. (1993) A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* **73**(4): 823-32.

Caruthers, J. M. and McKay, D. B. (2002) Helicase structure and mechanism. *Curr Opin Struct Biol* **12**(1): 123-33.

Center, R. J., Kobe, B., Wilson, K. A., Teh, T., Howlett, G. J., Kemp, B. E. and Poumbourios, P. (1998) Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein. *Protein Sci* **7**(7): 1612-9.

Chan, D. C., Fass, D., Berger, J. M. and Kim, P. S. (1997) Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**(2): 263-73.

Chattopadhyay, S. and Bielinsky, A. K. (2007) Human Mcm10 regulates the catalytic subunit of DNA polymerase-alpha and prevents DNA damage during replication. *Mol Biol Cell* **18**(10): 4085-95.

Chen, Y. J., Yu, X., Kasiviswanathan, R., Shin, J. H., Kelman, Z. and Egelman, E. H. (2005) Structural polymorphism of Methanothermobacter thermautotrophicus MCM. *J Mol Biol* **346**(2): 389-94.

Christensen, T. W. and Tye, B. K. (2003) Drosophila MCM10 interacts with members of the prereplication complex and is required for proper chromosome condensation. *Mol Biol Cell* **14**(6): 2206-15.

Chuang, R. Y. and Kelly, T. J. (1999) The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc Natl Acad Sci U S A* **96**(6): 2656-61.

Ciani, B., Bjelic, S., Honnappa, S., Jawhari, H., Jaussi, R., Payapilly, A., Jowitt, T., Steinmetz, M. O. and Kammerer, R. A. (2010) Molecular basis of coiled-coil oligomerization-state specificity. *Proc Natl Acad Sci U S A* **107**(46): 19850-5.

Cimprich, K. A. and Cortez, D. (2008) ATR: an essential regulator of genome integrity. *Nat Rev Mol Cell Biol* **9**(8): 616-27.

Conaway, R. C. and Lehman, I. R. (1982) Synthesis by the DNA primase of Drosophila melanogaster of a primer with a unique chain length. *Proc Natl Acad Sci U S A* **79**(15): 4585-8.

Cook, C. R., Kung, G., Peterson, F. C., Volkman, B. F. and Lei, M. (2003) A novel zinc finger is required for mcm10 homocomplex assembly. *J Biol Chem* **278**(38): 36051-8.

Corn, J. E. and Berger, J. M. (2006) Regulation of bacterial priming and daughter strand synthesis through helicase-primase interactions. *Nucleic Acids Res* **34**(15): 4082-8.

Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M. R. and Berger, J. M. (2011) The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nature structural & molecular biology* **18**: 471-477.

Cozzarelli, N. R. (1980) DNA gyrase and the supercoiling of DNA. *Science* **207**(4434): 953-60.

D'Andrea, A. D. (2010) Susceptibility pathways in Fanconi's anemia and breast cancer. *N Engl J Med* **362**(20): 1909-19.

Dahl, J., You, J. and Benjamin, T. L. (2005) Induction and utilization of an ATM signaling pathway by polyomavirus. *J Virol* **79**(20): 13007-17.

Dalgaard, J. Z., Eydmann, T., Koulintchenko, M., Sayrac, S., Vengrova, S. and Yamada-Inagawa, T. (2009) Random and site-specific replication termination. *Methods Mol Biol* **521**: 35-53.

Das-Bradoo, S., Ricke, R. M. and Bielinsky, A. K. (2006) Interaction between PCNA and diubiquitinated Mcm10 is essential for cell growth in budding yeast. *Mol Cell Biol* **26**(13): 4806-17.

Dayn, A., Malkhosyan, S. and Mirkin, S. M. (1992) Transcriptionally driven cruciform formation in vivo. *Nucleic Acids Res* **20**(22): 5991-7.

Debatisse, M., Toledo, F. and Anglana, M. (2004) Replication initiation in mammalian cells: changing preferences. *Cell Cycle* **3**(1): 19-21.

Dieckmann, G. R., McRorie, D. K., Lear, J. D., Sharp, K. A., DeGrado, W. F. and Pecoraro, V. L. (1998) The role of protonation and metal chelation preferences in defining the properties of mercury-binding coiled coils. *J Mol Biol* **280**(5): 897-912.

Doublie, S., Tabor, S., Long, A. M., Richardson, C. C. and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. *Nature* **391**(6664): 251-8.

Douglas, N. L., Dozier, S. K. and Donato, J. J. (2005) Dual roles for Mcm10 in DNA replication initiation and silencing at the mating-type loci. *Mol Biol Rep* **32**(4): 197-204.

Du, W., Stauffer, M. E. and Eichman, B. F. (2012) Structural biology of replication initiation factor mcm10. *Subcell Biochem* **62**: 197-216.

Dumas, L. B., Lussky, J. P., McFarland, E. J. and Shampay, J. (1982) New temperature-sensitive mutants of Saccharomyces cerevisiae affecting DNA replication. *Mol Gen Genet* **187**(1): 42-6.

Dunn, J. J. and Studier, F. W. (1983) Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol* **166**(4): 477-535.

Dutta, A. and Bell, S. P. (1997) Initiation of DNA replication in eukaryotic cells. *Annu Rev Cell Dev Biol* **13**: 293-332.

Dutta, K., Alexandrov, A., Huang, H. and Pascal, S. M. (2001) pH-induced folding of an apoptotic coiled coil. *Protein Sci* **10**(12): 2531-40.

Dutta, K., Engler, F. A., Cotton, L., Alexandrov, A., Bedi, G. S., Colquhoun, J. and Pascal, S. M. (2003) Stabilization of a pH-sensitive apoptosis-linked coiled coil through single point mutations. *Protein Sci* **12**(2): 257-65.

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* **277**(36): 33049-57.

Eisenberg, S., Korza, G., Carson, J., Liachko, I. and Tye, B. K. (2009) Novel DNA binding properties of the Mcm10 protein from Saccharomyces cerevisiae. *J Biol Chem* **284**(37): 25412-20.

Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1): 2126-32.

Fairman, R., Chao, H. G., Lavoie, T. B., Villafranca, J. J., Matsueda, G. R. and Novotny, J. (1996) Design of heterotetrameric coiled coils: evidence for increased stabilization by Glu(-)-Lys(+) ion pair interactions. *Biochemistry* **35**(9): 2824-9.

Fanning, E. and Zhao, K. (2009) SV40 DNA replication: from the A gene to a nanomachine. *Virology* **384**(2): 352-9.

Fatoba, S. T., Tognetti, S., Berto, M., Leo, E., Mulvey, C. M., Godovac-Zimmermann, J., Pommier, Y. and Okorokov, A. L. (2013) Human SIRT1 regulates DNA binding and stability of the Mcm10 DNA replication factor via deacetylation. *Nucleic Acids Res* **41**(7): 4065-79.

Fien, K., Cho, Y. S., Lee, J. K., Raychaudhuri, S., Tappin, I. and Hurwitz, J. (2004) Primer utilization by DNA polymerase alpha-primase is influenced by its interaction with Mcm10p. *J Biol Chem* **279**(16): 16144-53.

Fien, K. and Hurwitz, J. (2006) Fission yeast Mcm10p contains primase activity. *J Biol Chem* **281**(31): 22248-60.

Fletcher, R. J., Bishop, B. E., Leon, R. P., Sclafani, R. A., Ogata, C. M. and Chen, X. S. (2003) The structure and function of MCM from archaeal M. Thermoautotrophicum. *Nat Struct Biol* **10**(3): 160-7.

Fletcher, R. J., Shen, J., Gomez-Llorente, Y., Martin, C. S., Carazo, J. M. and Chen, X. S. (2005) Double hexamer disruption and biochemical activities of Methanobacterium thermoautotrophicum MCM. *J Biol Chem* **280**(51): 42405-10.

Fujita, M. (2006) Cdt1 revisited: complex and tight regulation during the cell cycle and consequences of deregulation in mammalian cells. *Cell Div* **1**: 22.

Funnell, B. E., Baker, T. A. and Kornberg, A. (1987) In vitro assembly of a prepriming complex at the origin of the Escherichia coli chromosome. *J Biol Chem* **262**(21): 10327-34.

Gambus, A., Jones, R. C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R. D. and Labib, K. (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol* **8**(4): 358-66.

Gambus, A., van Deursen, F., Polychronopoulos, D., Foltman, M., Jones, R. C., Edmondson, R. D., Calzada, A. and 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**(19): 2992-3004.

Garg, P., Stith, C. M., Majka, J. and Burgers, P. M. (2005) Proliferating cell nuclear antigen promotes translesion synthesis by DNA polymerase zeta. *J Biol Chem* **280**(25): 23446-50.

Garrett, R. and Grisham, C. M. (2013) Biochemistry. Belmont, CA, Brooks/Cole, Cengage Learning.

Gibbons, D. L., Ahn, A., Chatterjee, P. K. and Kielian, M. (2000) Formation and characterization of the trimeric form of the fusion protein of Semliki Forest Virus. *J Virol* **74**(17): 7772-80.

Gilbert, D. M. (2007) Replication origin plasticity, Taylor-made: inhibition vs recruitment of origins under conditions of replication stress. *Chromosoma* **116**(4): 341-7.

Giraldo, R. (2003) Common domains in the initiators of DNA replication in Bacteria, Archaea and Eukarya: combined structural, functional and phylogenetic perspectives. *FEMS Microbiol Rev* **26**(5): 533-54.

Gozuacik, D., Chami, M., Lagorce, D., Faivre, J., Murakami, Y., Poch, O., Biermann, E., Knippers, R., Brechot, C. and Paterlini-Brechot, P. (2003) Identification and functional characterization of a new member of the human Mcm protein family: hMcm8. *Nucleic Acids Res* **31**(2): 570-9.

Grabowski, B. and Kelman, Z. (2003) Archeal DNA replication: eukaryal proteins in a bacterial context. *Annu Rev Microbiol* **57**: 487-516.

Grallert, B. and Nurse, P. (1997) An approach to identify functional homologues and suppressors of genes in fission yeast. *Current genetics* **32**(1): 27-31.

Gregan, J., Lindner, K., Brimage, L., Franklin, R., Namdar, M., Hart, E. A., Aves, S. J. and Kearsey, S. E. (2003) Fission yeast Cdc23/Mcm10 functions after pre-replicative complex formation to promote Cdc45 chromatin binding. *Mol Biol Cell* **14**(9): 3876-87.

Haase, S. B., Heinzel, S. S. and Calos, M. P. (1994) Transcription inhibits the replication of autonomously replicating plasmids in human cells. *Mol Cell Biol* **14**(4): 2516-24.

Hanna, J. S., Kroll, E. S., Lundblad, V. and Spencer, F. A. (2001) Saccharomyces cerevisiae CTF18 and CTF4 are required for sister chromatid cohesion. *Mol Cell Biol* **21**(9): 3144-58.

Harbury, P. B., Kim, P. S. and Alber, T. (1994) Crystal structure of an isoleucine-zipper trimer. *Nature* **371**(6492): 80-3.

Harbury, P. B., Zhang, T., Kim, P. S. and Alber, T. (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**(5138): 1401-7.

Hart, E. A., Bryant, J. A., Moore, K. and Aves, S. J. (2002) Fission yeast Cdc23 interactions with DNA replication initiation proteins. *Curr Genet* **41**(5): 342-8.

Haworth, J., Alver, R. C., Anderson, M. and Bielinsky, A. K. (2010) Ubc4 and Not4 regulate steady-state levels of DNA polymerase-alpha to promote efficient and accurate DNA replication. *Mol Biol Cell* **21**(18): 3205-19.

Heller, R. C., Kang, S., Lam, W. M., Chen, S., Chan, C. S. and Bell, S. P. (2011) Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. *Cell* **146**(1): 80-91.

Hendrickson, H. and Lawrence, J. G. (2007) Mutational bias suggests that replication termination occurs near the dif site, not at Ter sites. *Mol Microbiol* **64**(1): 42-56.

Hennessy, K. M., Lee, A., Chen, E. and Botstein, D. (1991) A group of interacting yeast DNA replication genes. *Genes Dev* **5**(6): 958-69.

Homesley, L., Lei, M., Kawasaki, Y., Sawyer, S., Christensen, T. and Tye, B. K. (2000) Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and to release replication factors from origins. *Genes Dev* **14**(8): 913-26.

Hyrien, O., Marheineke, K. and Goldar, A. (2003) Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. *Bioessays* **25**(2): 116-25.

Ilves, I., Petojevic, T., Pesavento, J. J. and Botchan, M. R. (2010) Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell* **37**(2): 247-58.

Im, J. S., Ki, S. H., Farina, A., Jung, D. S., Hurwitz, J. and Lee, J. K. (2009) Assembly of the Cdc45-Mcm2-7-GINS complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins. *Proc Natl Acad Sci U S A* **106**(37): 15628-32.

Ishimi, Y. (1997) A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J Biol Chem* **272**(39): 24508-13.

Izumi, M., Yanagi, K., Mizuno, T., Yokoi, M., Kawasaki, Y., Moon, K. Y., Hurwitz, J., Yatagai, F. and 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**(23): 4769-77.

Izumi, M., Yatagai, F. and Hanaoka, F. (2001) Cell cycle-dependent proteolysis and phosphorylation of human Mcm10. *J Biol Chem* **276**(51): 48526-31.

Jung, N. Y., Bae, W. J., Chang, J. H., Kim, Y. C. and Cho, Y. (2008) Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of the central zinc-binding domain of the human Mcm10 DNA-replication factor. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**(Pt 6): 495-7.

Kanemaki, M., Sanchez-Diaz, A., Gambus, A. and Labib, K. (2003) Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. *Nature* **423**(6941): 720-4.

Kanke, M., Kodama, Y., Takahashi, T. S., Nakagawa, T. and Masukata, H. (2012) Mcm10 plays an essential role in origin DNA unwinding after loading of the CMG components. *EMBO J* **31**(9): 2182-94.

Kawasaki, Y., Hiraga, S. and Sugino, A. (2000) Interactions between Mcm10p and other replication factors are required for proper initiation and elongation of chromosomal DNA replication in Saccharomyces cerevisiae. *Genes Cells* **5**(12): 975-89.

Kelly, T. J. and Brown, G. W. (2000) Regulation of chromosome replication. *Annu Rev Biochem* **69**: 829-80.

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. and Daniel, D. C. (2008) Colocalization of MCM8 and MCM7 with proteins involved in distinct aspects of DNA replication. *Microsc Res Tech* **71**(4): 288-97.

Kobe, B., Center, R. J., Kemp, B. E. and Poumbourios, P. (1999) Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. *Proc Natl Acad Sci U S A* **96**(8): 4319-24.

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**(11): 2541-7.

Kubota, Y., Takase, Y., Komori, Y., Hashimoto, Y., Arata, T., Kamimura, Y., Araki, H. and Takisawa, H. (2003) A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication. *Genes Dev* **17**(9): 1141-52.

Labib, K. and Gambus, A. (2007) A key role for the GINS complex at DNA replication forks. *Trends Cell Biol* **17**(6): 271-8.

Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**(4): 477-86.

Laue, T. M., Shah, B., Ridgeway, T. M. and Pelletier, S. L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins. Analytical ultracentrifugation in biochemistry and polymer science. S. E. Harding. Cambridge, UK, Royal Society of Chemistry**:** 90–125.

Lavin, M. F. (2008) Ataxia-telangiectasia: from a rare disorder to a paradigm for cell signalling and cancer. *Nat Rev Mol Cell Biol* **9**(10): 759-69.

Lebofsky, R., Heilig, R., Sonnleitner, M., Weissenbach, J. and Bensimon, A. (2006) DNA replication origin interference increases the spacing between initiation events in human cells. *Mol Biol Cell* **17**(12): 5337-45.

Lee, C., Hong, B., Choi, J. M., Kim, Y., Watanabe, S., Ishimi, Y., Enomoto, T., Tada, S. and Cho, Y. (2004) Structural basis for inhibition of the replication licensing factor Cdt1 by geminin. *Nature* **430**(7002): 913-7.

Lee, C., Liachko, I., Bouten, R., Kelman, Z. and Tye, B. K. (2010) Alternative mechanisms for coordinating polymerase alpha and MCM helicase. *Mol Cell Biol* **30**(2): 423-35.

Lee, D. G. and Bell, S. P. (2000) ATPase switches controlling DNA replication initiation. *Curr Opin Cell Biol* **12**(3): 280-5.

Lee, J. K. and Hurwitz, J. (2000) Isolation and characterization of various complexes of the minichromosome maintenance proteins of Schizosaccharomyces pombe. *J Biol Chem* **275**(25): 18871-8.

Lee, J. K., Seo, Y. S. and Hurwitz, J. (2003) The Cdc23 (Mcm10) protein is required for the phosphorylation of minichromosome maintenance complex by the Dfp1-Hsk1 kinase. *Proc Natl Acad Sci U S A* **100**(5): 2334-9.

Lee, J. Y., Chang, C., Song, H. K., Moon, J., Yang, J. K., Kim, H. K., Kwon, S. T. and Suh, S. W. (2000) Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications. *Embo J* **19**(5): 1119-29.

Lei, M., Kawasaki, Y., Young, M. R., Kihara, M., Sugino, A. and Tye, B. K. (1997) Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis. *Genes Dev* **11**(24): 3365-74.

Leppard, J. B. and Champoux, J. J. (2005) Human DNA topoisomerase I: relaxation, roles, and damage control. *Chromosoma* **114**(2): 75-85.

Liachko, I. and Tye, B. K. (2005) Mcm10 is required for the maintenance of transcriptional silencing in Saccharomyces cerevisiae. *Genetics* **171**(2): 503-15.

Liachko, I. and Tye, B. K. (2009) Mcm10 mediates the interaction between DNA replication and silencing machineries. *Genetics* **181**(2): 379-91.

Liang, D. T. and Forsburg, S. L. (2001) Characterization of Schizosaccharomyces pombe mcm7(+) and cdc23(+) (MCM10) and interactions with replication checkpoints. *Genetics* **159**(2): 471-86.

Liu, J., Zheng, Q., Deng, Y., Cheng, C. S., Kallenbach, N. R. and Lu, M. (2006) A seven-helix coiled coil. *Proc Natl Acad Sci U S A* **103**(42): 15457-62.

Liu, Y., Richards, T. A. and Aves, S. J. (2009) Ancient diversification of eukaryotic MCM DNA replication proteins. *BMC Evol Biol* **9**: 60.

Lowey, S. (1965) Comparative Study of the Alpha-Helical Muscle Proteins. Tyrosyl Titration and Effect of Ph on Conformation. *J Biol Chem* **240**: 2421-7.

Lutzmann, M., Maiorano, D. and Mechali, M. (2005) Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2-8 protein family. *Gene* **362**: 51-6.

Lutzmann, M. and Mechali, M. (2008) MCM9 binds Cdt1 and is required for the assembly of prereplication complexes. *Mol Cell* **31**(2): 190-200.

MacNeill, S. (2012) The eukaryotic replisome : a guide to protein structure and function. New York, Springer.

Maine, G. T., Sinha, P. and Tye, B. K. (1984) Mutants of S. cerevisiae defective in the maintenance of minichromosomes. *Genetics* **106**(3): 365-85.

Maiorano, D., Cuvier, O., Danis, E. and Mechali, M. (2005) MCM8 is an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. *Cell* **120**(3): 315-28.

Maiorano, D., Lutzmann, M. and Mechali, M. (2006) MCM proteins and DNA replication. *Curr Opin Cell Biol* **18**(2): 130-6.

Maiorano, D., Moreau, J. and Mechali, M. (2000) XCDT1 is required for the assembly of pre-replicative complexes in Xenopus laevis. *Nature* **404**(6778): 622-5.

Marahrens, Y. and Stillman, B. (1992) A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* **255**(5046): 817-23.

Marvin, J. S., Corcoran, E. E., Hattangadi, N. A., Zhang, J. V., Gere, S. A. and Hellinga, H. W. (1997) The rational design of allosteric interactions in a monomeric protein and its applications to the construction of biosensors. *Proc Natl Acad Sci U S A* **94**(9): 4366-71.

Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N. and Oda, M. (2010) Eukaryotic chromosome DNA replication: where, when, and how? *Annu Rev Biochem* **79**: 89-130.

Matsumoto, T., Eki, T. and Hurwitz, J. (1990) Studies on the initiation and elongation reactions in the simian virus 40 DNA replication system. *Proc Natl Acad Sci U S A* **87**(24): 9712-6.

McGarry, T. J. and Kirschner, M. W. (1998) Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell* **93**(6): 1043-53.

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* **9**(3): 481-91.

Merchant, A. M., Kawasaki, Y., Chen, Y., Lei, M. and Tye, B. K. (1997) A lesion in the DNA replication initiation factor Mcm10 induces pausing of elongation forks through chromosomal replication origins in Saccharomyces cerevisiae. *Mol Cell Biol* **17**(6): 3261-71.

Mizuno, T., Ito, N., Yokoi, M., Kobayashi, A., Tamai, K., Miyazawa, H. and 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**(6): 3552-62.

Moir, D., Stewart, S. E., Osmond, B. C. and Botstein, D. (1982) Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. *Genetics* **100**(4): 547-63.

Moon, A. F., Mueller, G. A., Zhong, X. and Pedersen, L. C. (2010) A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. *Protein Sci* **19**(5): 901-13.

Mott, M. L. and Berger, J. M. (2007) DNA replication initiation: mechanisms and regulation in bacteria. *Nat Rev Microbiol* **5**(5): 343-54.

Moutevelis, E. and Woolfson, D. N. (2009) A periodic table of coiled-coil protein structures. *J Mol Biol* **385**(3): 726-32.

Moyer, S. E., Lewis, P. W. and Botchan, M. R. (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**(27): 10236-41.

Mueller, G. A., Edwards, L. L., Aloor, J. J., Fessler, M. B., Glesner, J., Pomes, A., Chapman, M. D., London, R. E. and Pedersen, L. C. (2010) The structure of the dust mite allergen Der p 7 reveals similarities to innate immune proteins. *J Allergy Clin Immunol* **125**(4): 909-917 e4.

Murakami, Y., Eki, T. and Hurwitz, J. (1992) Studies on the initiation of simian virus 40 replication in vitro: RNA primer synthesis and its elongation. *Proc Natl Acad Sci U S A* **89**(3): 952-6.

Nasmyth, K. and Nurse, P. (1981) Cell division cycle mutants altered in DNA replication and mitosis in the fission yeast Schizosaccharomyces pombe. *Mol Gen Genet* **182**(1): 119-24.

Nick McElhinny, S. A., Gordenin, D. A., Stith, C. M., Burgers, P. M. and Kunkel, T. A. (2008) Division of labor at the eukaryotic replication fork. *Mol Cell* **30**(2): 137-44.

Nishimura, K., Ishiai, M., Horikawa, K., Fukagawa, T., Takata, M., Takisawa, H. and Kanemaki, M. T. (2012) Mcm8 and Mcm9 form a complex that functions in homologous recombination repair induced by DNA interstrand crosslinks. *Mol Cell* **47**(4): 511-22.

Nishitani, H., Sugimoto, N., Roukos, V., Nakanishi, Y., Saijo, M., Obuse, C., Tsurimoto, T., Nakayama, K. I., Nakayama, K., Fujita, M., Lygerou, Z. and Nishimoto, T. (2006) Two E3 ubiquitin ligases, SCF-Skp2 and DDB1-Cul4, target human Cdt1 for proteolysis. *EMBO J* **25**(5): 1126-36.

Nishitani, H., Taraviras, S., Lygerou, Z. and Nishimoto, T. (2001) The human licensing factor for DNA replication Cdt1 accumulates in G1 and is destabilized after initiation of S-phase. *J Biol Chem* **276**(48): 44905-11.

Noelken, M., and Holtzer, A. (1964). Biochemistry of Muscle Contraction. J. Gergely. Boston, Little, Brown and Co.**:** 374–378.

O'Shea, E. K., Klemm, J. D., Kim, P. S. and Alber, T. (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254**(5031): 539-44.

O'Shea, E. K., Lumb, K. J. and Kim, P. S. (1993) Peptide 'Velcro': design of a heterodimeric coiled coil. *Curr Biol* **3**(10): 658-67.

O'Shea, E. K., Rutkowski, R. and Kim, P. S. (1992) Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. *Cell* **68**(4): 699-708.

Ohta, S., Tatsumi, Y., Fujita, M., Tsurimoto, T. and Obuse, C. (2003) The ORC1 cycle in human cells: II. Dynamic changes in the human ORC complex during the cell cycle. *J Biol Chem* **278**(42): 41535-40.

Okorokov, A. L., Waugh, A., Hodgkinson, J., Murthy, A., Hong, H. K., Leo, E., Sherman, M. B., Stoeber, K., Orlova, E. V. and Williams, G. H. (2007) Hexameric ring structure of human MCM10 DNA replication factor. *EMBO Rep* **8**(10): 925-30.

Oliva, A., Rosebrock, A., Ferrezuelo, F., Pyne, S., Chen, H., Skiena, S., Futcher, B. and Leatherwood, J. (2005) The cell cycle-regulated genes of Schizosaccharomyces pombe. *PLoS Biol* **3**(7): e225.

Ott, R. D., Rehfuess, C., Podust, V. N., Clark, J. E. and Fanning, E. (2002) Role of the p68 subunit of human DNA polymerase alpha-primase in simian virus 40 DNA replication. *Mol Cell Biol* **22**(16): 5669-78.

Otwinowski, Z. and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**: 307-326.

Pacek, M., Tutter, A. V., Kubota, Y., Takisawa, H. and Walter, J. C. (2006) Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol Cell* **21**(4): 581-7.

Pacek, M. and Walter, J. C. (2004) A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. *Embo J* **23**(18): 3667-76.

Paixao, S., Colaluca, I. N., Cubells, M., Peverali, F. A., Destro, A., Giadrossi, S., Giacca, M., Falaschi, A., Riva, S. and Biamonti, G. (2004) Modular structure of the human lamin B2 replicator. *Mol Cell Biol* **24**(7): 2958-67.

Palzkill, T. G. and Newlon, C. S. (1988) A yeast replication origin consists of multiple copies of a small conserved sequence. *Cell* **53**(3): 441-50.

Pape, T., Meka, H., Chen, S., Vicentini, G., van Heel, M. and Onesti, S. (2003) Hexameric ring structure of the full-length archaeal MCM protein complex. *EMBO Rep* **4**(11): 1079-83.

Patel, S. S. and Picha, K. M. (2000) Structure and function of hexameric helicases. *Annu Rev Biochem* **69**: 651-97.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**(13): 1605-12.

Pospiech, H., Kursula, I., Abdel-Aziz, W., Malkas, L., Uitto, L., Kastelli, M., Vihinen-Ranta, M., Eskelinen, S. and Syvaoja, J. E. (1999) A neutralizing antibody against human DNA polymerase epsilon inhibits cellular but not SV40 DNA replication. *Nucleic Acids Res* **27**(19): 3799-804.

Prokhorova, T. A. and Blow, J. J. (2000) Sequential MCM/P1 subcomplex assembly is required to form a heterohexamer with replication licensing activity. *J Biol Chem* **275**(4): 2491-8.

Pursell, Z. F., Isoz, I., Lundstrom, E. B., Johansson, E. and Kunkel, T. A. (2007) Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* **317**(5834): 127-30.

Raghunathan, S., Kozlov, A. G., Lohman, T. M. and Waksman, G. (2000) Structure of the DNA binding domain of E. coli SSB bound to ssDNA. *Nat Struct Biol* **7**(8): 648-52.

Raghunathan, S., Ricard, C. S., Lohman, T. M. and Waksman, G. (1997) Crystal structure of the homo-tetrameric DNA binding domain of Escherichia coli single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-A resolution. *Proc Natl Acad Sci U S A* **94**(13): 6652-7.

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* **139**(4): 719-30.

Ricke, R. M. and Bielinsky, A. K. (2004) Mcm10 Regulates the Stability and Chromatin Association of DNA Polymerase-alpha. *Mol Cell* **16**(2): 173-85.

Ricke, R. M. and Bielinsky, A. K. (2006) A conserved Hsp10-like domain in Mcm10 is required to stabilize the catalytic subunit of DNA polymerase-alpha in budding yeast. *J Biol Chem* **281**(27): 18414-25.

Robertson, P. D., Chagot, B., Chazin, W. J. and Eichman, B. F. (2010) Solution NMR structure of the C-terminal DNA binding domain of MCM10 reveals a conserved MCM motif. *J Biol Chem* **285**(30): 22942-22949.

Robertson, P. D., Warren, E. M., Zhang, H., Friedman, D. B., Lary, J. W., Cole, J. L., Tutter, A. V., Walter, J. C., Fanning, E. and Eichman, B. F. (2008) Domain architecture and biochemical characterization of vertebrate Mcm10. *J Biol Chem* **283**(6): 3338-48.

Saha, S., Shan, Y., Mesner, L. D. and Hamlin, J. L. (2004) The promoter of the Chinese hamster ovary dihydrofolate reductase gene regulates the activity of the local origin and helps define its boundaries. *Genes Dev* **18**(4): 397-410.

Sangrithi, M. N., Bernal, J. A., Madine, M., Philpott, A., Lee, J., Dunphy, W. G. and Venkitaraman, A. R. (2005) Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. *Cell* **121**(6): 887-98.

Sawyer, S. L., Cheng, I. H., Chai, W. and Tye, B. K. (2004) Mcm10 and Cdc45 cooperate in origin activation in Saccharomyces cerevisiae. *J Mol Biol* **340**(2): 195-202.

Saxena, S. and Dutta, A. (2005) Geminin-Cdt1 balance is critical for genetic stability. *Mutat Res* **569**(1-2): 111-21.

Saxena, S., Yuan, P., Dhar, S. K., Senga, T., Takeda, D., Robinson, H., Kornbluth, S., Swaminathan, K. and Dutta, A. (2004) A dimerized coiled-coil domain and an adjoining part of geminin interact with two sites on Cdt1 for replication inhibition. *Mol Cell* **15**(2): 245-58.

Schaarschmidt, D., Baltin, J., Stehle, I. M., Lipps, H. J. and Knippers, R. (2004) An episomal mammalian replicon: sequence-independent binding of the origin recognition complex. *EMBO J* **23**(1): 191-201.

Schaper, S. and Messer, W. (1995) Interaction of the initiator protein DnaA of Escherichia coli with its DNA target. *J Biol Chem* **270**(29): 17622-6.

Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**(3): 1606-19.

Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. and Schubert, D. (2002) Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys J* **82**(2): 1096-111.

Sclafani, R. A. and Holzen, T. M. (2007) Cell cycle regulation of DNA replication. *Annu Rev Genet* **41**: 237-80.

Seki, T., Akita, M., Kamimura, Y., Muramatsu, S., Araki, H. and Sugino, A. (2006) GINS is a DNA polymerase epsilon accessory factor during chromosomal DNA replication in budding yeast. *J Biol Chem* **281**(30): 21422-32.

Senga, T., Sivaprasad, U., Zhu, W., Park, J. H., Arias, E. E., Walter, J. C. and Dutta, A. (2006) PCNA is a cofactor for Cdt1 degradation by CUL4/DDB1-mediated N-terminal ubiquitination. *J Biol Chem* **281**(10): 6246-52.

Shamoo, Y., Friedman, A. M., Parsons, M. R., Konigsberg, W. H. and Steitz, T. A. (1995) Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature* **376**(6538): 362-6.

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* **24**(1): 101-13.

Shi, Y., Dodson, G. E., Shaikh, S., Rundell, K. and Tibbetts, R. S. (2005) Ataxia-telangiectasia-mutated (ATM) is a T-antigen kinase that controls SV40 viral replication in vivo. *J Biol Chem* **280**(48): 40195-200.

Shikata, K., Sasa-Masuda, T., Okuno, Y., Waga, S. and Sugino, A. (2006) The DNA polymerase activity of Pol epsilon holoenzyme is required for rapid and efficient chromosomal DNA replication in Xenopus egg extracts. *BMC Biochem* **7**: 21.

Snyder, M., Sapolsky, R. J. and Davis, R. W. (1988) Transcription interferes with elements important for chromosome maintenance in Saccharomyces cerevisiae. *Mol Cell Biol* **8**(5): 2184-94.

Solomon, N. A., Wright, M. B., Chang, S., Buckley, A. M., Dumas, L. B. and Gaber, R. F. (1992) Genetic and molecular analysis of DNA43 and DNA52: two new cell-cycle genes in Saccharomyces cerevisiae. *Yeast* **8**(4): 273-89.

Spinola-Amilibia, M., Rivera, J., Ortiz-Lombardia, M., Romero, A., Neira, J. L. and Bravo, J. (2011) The structure of BRMS1 nuclear export signal and SNX6 interacting region reveals a hexamer formed by antiparallel coiled coils. *J Mol Biol* **411**(5): 1114-27.

Spinola-Amilibia, M., Rivera, J., Ortiz-Lombardia, M., Romero, A., Neira, J. L. and Bravo, J. (2013) BRMS1 and BRMS1 Are Crystal Oligomeric Coiled Coils with Different Oligomerization States, Which Behave as Disordered Protein Fragments in Solution. *J Mol Biol* **in press**: (10.1016/j.jmb.2013.03.005).

Stahl, H., Droge, P. and Knippers, R. (1986) DNA helicase activity of SV40 large tumor antigen. *EMBO J* **5**(8): 1939-44.

Stauffer, M. E. and Chazin, W. J. (2004) Structural mechanisms of DNA replication, repair, and recombination. *J Biol Chem* **279**(30): 30915-8.

Stillman, B. (2005) Origin recognition and the chromosome cycle. *FEBS Lett* **579**(4): 877-84.

Studier, F. W. and Dunn, J. J. (1983) Organization and expression of bacteriophage T7 DNA. *Cold Spring Harb Symp Quant Biol* **47 Pt 2**: 999-1007.

Sugino, A. (1995) Yeast DNA polymerases and their role at the replication fork. *Trends Biochem Sci* **20**(8): 319-23.

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* **3**(2): 107-13.

Tadokoro, R., Fujita, M., Miura, H., Shirahige, K., Yoshikawa, H., Tsurimoto, T. and Obuse, C. (2002) Scheduled conversion of replication complex architecture at replication origins of Saccharomyces cerevisiae during the cell cycle. *J Biol Chem* **277**(18): 15881-9.

Takahashi, K., Yamada, H. and Yanagida, M. (1994) Fission yeast minichromosome loss mutants mis cause lethal aneuploidy and replication abnormality. *Mol Biol Cell* **5**(10): 1145-58.

Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A. and Araki, H. (2003) GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev* **17**(9): 1153-65.

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* **280**(24): 23416-23.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y. and Araki, H. (2007) CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* **445**(7125): 328-32.

Tanaka, T. and Nasmyth, K. (1998) Association of RPA with chromosomal replication origins requires an Mcm protein, and is regulated by Rad53, and cyclin- and Dbf4-dependent kinases. *Embo J* **17**(17): 5182-91.

Taylor, A. M., Groom, A. and Byrd, P. J. (2004) Ataxia-telangiectasia-like disorder (ATLD)-its clinical presentation and molecular basis. *DNA Repair (Amst)* **3**(8-9): 1219-25.

Taylor, J. H. (1977) Increase in DNA replication sites in cells held at the beginning of S phase. *Chromosoma* **62**(4): 291-300.

Tercero, J. A., Labib, K. and Diffley, J. F. (2000) DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. *Embo J* **19**(9): 2082-93.

Theis, J. F. and Newlon, C. S. (1997) The ARS309 chromosomal replicator of Saccharomyces cerevisiae depends on an exceptional ARS consensus sequence. *Proc Natl Acad Sci U S A* **94**(20): 10786-91.

Thepaut, M., Maiorano, D., Guichou, J. F., Auge, M. T., Dumas, C., Mechali, M. and Padilla, A. (2004) Crystal structure of the coiled-coil dimerization motif of geminin: structural and functional insights on DNA replication regulation. *J Mol Biol* **342**(1): 275-87.

Thu, Y. M. and Bielinsky, A. K. (2013) Enigmatic roles of Mcm10 in DNA replication. *Trends Biochem Sci* **38**(4): 184-94.

Tye, B. K. (2000) Insights into DNA replication from the third domain of life. *Proc Natl Acad Sci U S A* **97**(6): 2399-401.

van Deursen, F., Sengupta, S., De Piccoli, G., Sanchez-Diaz, A. and Labib, K. (2012) Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation. *EMBO J* **31**(9): 2195-206.

Volkening, M. and Hoffmann, I. (2005) Involvement of human MCM8 in prereplication complex assembly by recruiting hcdc6 to chromatin. *Mol Cell Biol* **25**(4): 1560-8.

Waga, S. and Stillman, B. (1998) The DNA replication fork in eukaryotic cells. *Annu Rev Biochem* **67**: 721-51.

Walshaw, J. and Woolfson, D. N. (2003) Extended knobs-into-holes packing in classical and complex coiled-coil assemblies. *J Struct Biol* **144**(3): 349-61.

Walter, J. and Newport, J. (2000) Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. *Mol Cell* **5**(4): 617-27.

Wang, L., Lin, C. M., Brooks, S., Cimbora, D., Groudine, M. and Aladjem, M. I. (2004) The human beta-globin replication initiation region consists of two modular independent replicators. *Mol Cell Biol* **24**(8): 3373-86.

Warren, E. M., Huang, H., Fanning, E., Chazin, W. J. and Eichman, B. F. (2009) Physical interactions between Mcm10, DNA, and DNA polymerase alpha. *J Biol Chem* **284**(36): 24662-72.

Warren, E. M., Vaithiyalingam, S., Haworth, J., Greer, B., Bielinsky, A. K., Chazin, W. J. and Eichman, B. F. (2008) Structural basis for DNA binding by replication initiator Mcm10. *Structure* **16**(12): 1892-901.

Watase, G., Takisawa, H. and Kanemaki, M. T. (2012) Mcm10 plays a role in functioning of the eukaryotic replicative DNA helicase, Cdc45-Mcm-GINS. *Curr Biol* **22**(4): 343-9.

Wawrousek, K. E., Fortini, B. K., Polaczek, P., Chen, L., Liu, Q., Dunphy, W. G. and Campbell, J. L. (2010) Xenopus DNA2 is a helicase/nuclease that is found in complexes with replication proteins And-1/Ctf4 and Mcm10 and DSB response proteins Nbs1 and ATM. *Cell Cycle* **9**(6).

Weigel, C. and Seitz, H. (2006) Bacteriophage replication modules. *FEMS Microbiol Rev* **30**(3): 321-81.

Wigge, P. A., Jensen, O. N., Holmes, S., Soues, S., Mann, M. and Kilmartin, J. V. (1998) Analysis of the Saccharomyces spindle pole by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. *J Cell Biol* **141**(4): 967-77.

Wohlschlegel, J. A., Dhar, S. K., Prokhorova, T. A., Dutta, A. and Walter, J. C. (2002) Xenopus Mcm10 binds to origins of DNA replication after Mcm2-7 and stimulates origin binding of Cdc45. *Mol Cell* **9**(2): 233-40.

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* **290**(5500): 2309-12.

Woolfson, D. N. (2005) The design of coiled-coil structures and assemblies. *Adv Protein Chem* **70**: 79-112.

Wu, L. (2008) Wrestling off RAD51: a novel role for RecQ helicases. *Bioessays* **30**(4): 291-5.

Wu, X., Avni, D., Chiba, T., Yan, F., Zhao, Q., Lin, Y., Heng, H. and Livingston, D. (2004) SV40 T antigen interacts with Nbs1 to disrupt DNA replication control. *Genes Dev* **18**(11): 1305-16.

Xouri, G., Dimaki, M., Bastiaens, P. I. and Lygerou, Z. (2007) Cdt1 interactions in the licensing process: a model for dynamic spatiotemporal control of licensing. *Cell Cycle* **6**(13): 1549-52.

Xouri, G., Squire, A., Dimaki, M., Geverts, B., Verveer, P. J., Taraviras, S., Nishitani, H., Houtsmuller, A. B., Bastiaens, P. I. and Lygerou, Z. (2007) Cdt1 associates dynamically with chromatin throughout G1 and recruits Geminin onto chromatin. *EMBO J* **26**(5): 1303-14.

Xu, X., Rochette, P. J., Feyissa, E. A., Su, T. V. and Liu, Y. (2009) MCM10 mediates RECQ4 association with MCM2-7 helicase complex during DNA replication. *Embo J* **28**(19): 3005-14.

Yanagi, K., Mizuno, T., You, Z. and Hanaoka, F. (2002) Mouse geminin inhibits not only Cdt1-MCM6 interactions but also a novel intrinsic Cdt1 DNA binding activity. *J Biol Chem* **277**(43): 40871-80.

Yang, X., Gregan, J., Lindner, K., Young, H. and Kearsey, S. E. (2005) Nuclear distribution and chromatin association of DNA polymerase alpha-primase is affected by TEV protease cleavage of Cdc23 (Mcm10) in fission yeast. *BMC Mol Biol* **6**: 13.

Yardimci, H., Loveland, A. B., Habuchi, S., van Oijen, A. M. and Walter, J. C. (2010) Uncoupling of sister replisomes during eukaryotic DNA replication. *Mol Cell* **40**(5): 834-40.

You, Z., Ishimi, Y., Mizuno, T., Sugasawa, K., Hanaoka, F. and Masai, H. (2003) Thymine-rich single-stranded DNA activates Mcm4/6/7 helicase on Y-fork and bubble-like substrates. *EMBO J* **22**(22): 6148-60.

Zegerman, P. and Diffley, J. F. (2007) Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* **445**(7125): 281-5.

Zhao, X., Madden-Fuentes, R. J., Lou, B. X., Pipas, J. M., Gerhardt, J., Rigell, C. J. and Fanning, E. (2008) Ataxia telangiectasia-mutated damage-signaling

kinase- and proteasome-dependent destruction of Mre11-Rad50-Nbs1 subunits in Simian virus 40-infected primate cells. *J Virol* **82**(11): 5316-28.

Zhou, N. E., Kay, C. M. and Hodges, R. S. (1992) Synthetic model proteins. Positional effects of interchain hydrophobic interactions on stability of two-stranded alpha-helical coiled-coils. *J Biol Chem* **267**(4): 2664-70.

Zhu, W., Ukomadu, C., Jha, S., Senga, T., Dhar, S. K., Wohlschlegel, J. A., Nutt, L. K., Kornbluth, S. and Dutta, A. (2007) Mcm10 and And-1/CTF4 recruit DNA polymerase alpha to chromatin for initiation of DNA replication. *Genes Dev* **21**(18): 2288-99.

Zlotkin, T., Kaufmann, G., Jiang, Y., Lee, M. Y., Uitto, L., Syvaoja, J., Dornreiter, I., Fanning, E. and Nethanel, T. (1996) DNA polymerase epsilon may be dispensable for SV40- but not cellular-DNA replication. *EMBO J* **15**(9): 2298-305.

Zou, L. and Stillman, B. (2000) Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase. *Mol Cell Biol* **20**(9): 3086-96.