STRUCTURAL AND BIOCHEMICAL STUDIES OF EUKARYOTIC REPLICATION INITIATION FACTOR MCM10 FROM XENOPUS LAEVIS

By

Eric Mason Warren Jr.

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

August, 2009

Nashville, Tennessee

Approved:

Brandt Eichman, Ph.D.

Ellen Fanning, Dr. rer. nat.

Al Beth, Ph.D.

Walter Chazin, Ph.D.

Daniel Kaplan, Ph.D.

ACKNOWLEDGEMENTS

First I would like to thank my mentor and Ph.D. advisor, Dr. Brandt Eichman. Brandt has always done everything in his power to make sure I had every resource possible to proceed with my research. Thank you for always pushing and encouraging me to be the best scientist I could. Thank you for spending hours at the bench and in your office making sure I could do the best work possible. I am indebted to the NIH and the Molecular Biophysics Training Program (MBTP) for providing funding for this research.

I would also like to express my deepest gratitude to my graduate committee members: Drs. Ellen Fanning, Daniel Kaplan, Walter Chazin, and Albert Beth. This research would not have been possible without the guidance, encouragement, and feedback of these brilliant scientists and mentors.

I am especially grateful to Ellen as she is the one who first piqued my interest in research and spurred me on to pursue my fascination with the role of DNA replication in cancer. Ellen's love for research is contagious, and I appreciate her investing so much time in undergraduates because I was an undergraduate and that love for research rubbed off on me. Again, thank you for being a great mentor. Last, but not least, I thank you and Steve for being wonderful friends. Sabrina and I thoroughly enjoyed discussing music, politics and life over dinner with the two of you. Thank you!

I extend my great appreciation to so many other friends and lab members. To Dr. Bob Ott who always made science fun and enjoyable. I really value our candid and random conversations as well as your friendship. To Dr. Vitaly Klimovich who always seemed to find ways of helping everyone relax. Things always seemed brighter because

of your jokes, organizing tea times, and laughs at lunch time. To Dr. Steve Gray for his sense of humor and for providing a temporary escape while playing multiplayer LAN games. To Patrick Robertson for constant laughs that helped me to stay sane when I ran into problems with experiments. Your friendship and support was invaluable to me. You are all wonderful friends and I truly value the times we were able to spend together.

I want to extend a very special thanks to my family who helped me immensely in getting to this point. My wife Sabrina has been my pillar of strength and support, without her I could not survive. She has been such a wonderful source of love, strength, happiness and joy. It didn't matter how frustrating my day was, when I came home she always somehow knew what I needed to hear or what needed to be done to make those frustrations vanish away. I would not be who I am today without my: grandmother Marie Warf who always encouraged me to persevere, loved inquiring about my work, and never let me go without; parents Eric and Julie Warren who always pushed me towards my goals and would do anything to help me accomplish them; and my siblings Kenner and Kelley for their love and support. Thank you for being the most encouraging, caring, and loving family in the world! I will always be grateful for what you have done for me.

Last, but certainly not least, I want to take the time to thank our Lord and Creator. No words can do justice to the love You have for us. Thank You for that love that You demonstrated by sending Jesus. And thank You for never leaving me, and for never allowing me to face more than I could handle.

Psalm 139:1-14

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAA+	ATPases associated with a variety of cellular activities
aa	amino acids
ARS	Autonomously replicating sequence
ASU	Asymmetric unit
ATP	Adenosine triphosphate
BME	2-mercaptoethanol
bp	Base pair
C	Celsius
CCCH	Cystine-Cystine-Cystine-Histidine
CD	Circular Dichroism
Cdc	Cell division cycle
CDK	Cyclin dependent kinase
Cdt1	Cdc10-dependent transcript 1
Clb	Cyclin B
CMG	Cdc45/Mcm2-7/GINS
CTD	Carboxyl-terminal domain
DDK	Dbf-4 dependent kinase
Dbf4	Dumb bell forming 4
DNA	Deoxyribonucleic acid
DNA2	DNA replication ATP-dependent helicase
DNA43	DNA synthesis defective 43

Dpb11	DNA polymerase B subunit 11
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediaminotetraacetic acid
EM	Electron Microscopy
FITC	Fluorescein Isothiocyanate
g	Gram
G1	Gap 1 phase
GFAA	Gas Furnace Atomic Absorption
GINS	*Go ichi ni san* (Japanese for 5, 1, 2, 3)
PSF1	Partner of Sld5 1
PSF2	Partner of Sld5 2
PSF3	Partner of Sld5 3
SLD5	Synthetic lethality with dpb11 5
GST	Glutathione-S-transferase
h	hour
HCl	Hydrochloric acid
HSQC	Heteronuclear single quantum coherence
ID	Internal domain
IP	Immunoprecipitation
IPTG	Isopropyl thio-beta-D-galactopyranoside
kb	Kilobase pair
k_d	Dissociation constant

kDa	Kilo Dalton
L	Liter
LB	Luria broth
M	Molar
MADS	Mcm1, AG, DEFA, and SRF
MALDI-TOF	Matrix assisted laser desorption ionization – time of flight
MCM	Mini-chromosome maintenance
Mcm	Mini-chromosome maintenance
hMcm10	*H. sapiens* Mcm10
scMcm10	*S. cerevisiae* Mcm10
spMcm10	*S. pombe* Mcm10
xMcm10	*X. laevis* Mcm10
µg	Mirogram
µM	Miromolar
mg	Milligram
min	Minute
mL	Milliliter
mM	Millimolar
MS	Mass Spectrometry
MTS	Methanethiosulfonate
Ni-NTA	Nickel-nitrilotriacetic acid
ng	Nanogram
NLS	Nuclear localization signal

nm	Nanometers
NMR	Nuclear magnetic resonance
nt	Nucleotide
NTD	Amino-terminal domain
OD	Optical density
OB-fold	Oligonucleotide/oligosaccharide binding-fold
ORC	Origin recognition complex
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDB	Protein databank
PMSF	Phenylmethylsulfonyl fluoride
Pol	DNA polymerase
Pol α	DNA polymerase α-primase
Pre-RC	Pre-replication complex
Pre-IC	Pre-initiation complex
RFC	replication factor C
RPA	Replication protein A
rmsd	Root-mean-square Deviation
RNA	Ribonucleic acid
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	Second
Sld	Synthetic lethality with dpb11
ssDNA	single-stranded DNA

SV40	Simian virus 40
T-ag	SV40 Large T antigen
Tris	Tris-(hydroxymethyl-)aminomethane
TopBP1	Topoisomerase II binding protein 1
U	Unit
WT	Wild-type

CHAPTER I

INTRODUCTION

DNA Replication Overview

Maintenance of the genome of living organisms is paramount to their survival. Errors resulting in overeplication or undereplication of genetic information can have disastrous consequences and can lead to a plethora of diseases such as cancer, birth defects, and many developmental abnormalities (DePamphilis 2006). Thus, this process must be precisely coordinated by an ordered series of proteins at specified locations on chromatin, known as origins of replication. In its most basic form, the unit of DNA replication consists of two regulatory components: a cis-acting element known as a replicator, and a trans-acting element known as the initiator (Figure 1)(Jacob F 1963). The initiator is a sequence-specific DNA binding protein that binds to a defined region in the genome – the replicator. Once bound, it then recruits factors to assemble the functional replication machinery called the replisome. Over evolutionary time, these components have become increasingly complex to compensate for more complex genomes so that the factors that determine an origin of replication vary significantly from bacteria to yeast to humans.

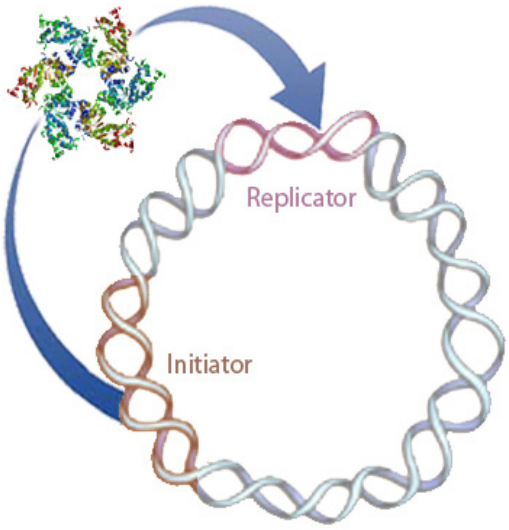


Figure 1. The replicon model. Originally proposed by Jacob, Brenner and Cuzin in 1963. (Initiator image from PDB ID 1SVO (Gai, Zhao et al. 2004)).

In eubacteria, such as *Escherichia coli*, replication is initiated by a single protein called DnaA that recognizes and binds to simple nine base pair AT-rich sequences known as DnaA boxes. The initiator DnaA then distorts the AT-rich sequences to melt the DNA duplex. This is in contrast to most eukaryotes. Although unicellular eukaryotes such as *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* have defined recognition sequences termed autonomously replicating sequences (ARSs), they are typically longer (10-15 base pairs) and have a spatial arrangement of 100-150 and 800-1000 base pairs respectively (Bell 1995; Clyne and Kelly 1995). In metazoan cells, the situation is less clear, there are some origins that act as defined sequences on which to replicate DNA (Paixao, Colaluca et al. 2004), while other times (such as the dihydrofolate reductase origin) initiation seems to occur randomly over broad zones spanning 5-15kb (Altman and Fanning 2001; Blow, Gillespie et al. 2001). In mammalian cells, DNA replication occurs over broad zones, often stretching over 10kb and containing genetic and epigenetic elements (Mendez and Stillman 2003). Despite these complexities, it has been shown that initiation begins at preferred chromosomal locations (Vaughn, Dijkwel et al. 1990).

Although the details differ greatly from prokaryotes to eukaryotes, there remain three common steps of DNA replication: initiation, elongation, and termination. This dissertation will focus on examining the intricacies of eukaryotic DNA replication initiation from a structural standpoint. Briefly, initiation is the process by which origins are recognized, DNA is initially melted, and replication proteins are recruited to form competent replication complexes. During the elongation phase, the helicase uses the energy harnessed from ATP hydrolysis to unwind the melted DNA to expose each strand.

Concurrently, DNA polymerases use this exposed template strand to synthesize the complementary strand in a 5' to 3' manner, while topoisomerases ahead of the progressing replication fork relieve helical tension. Termination of DNA replication occurs when replication forks encounter termination sequences or regions of the genome that have already been replicated. These replication forks are signaled to disassemble and more topoisomerases act to separate the sister chromosomes.

Eukaryotic DNA Replication Initiation

In eukaryotes, the initiation of DNA replication is a highly regulated process and is essential for maintenance of genome integrity. Origins of replication direct and choreograph the assembly of several dynamic multiprotein complexes. These complexes must recognize specific DNA elements at origins of replication, unwind the DNA duplex, and assemble into the functional replisome—the multiprotein machinery necessary to synthesize DNA at the replication fork. The replisome consists of a polymerase, a primase, single-stranded binding proteins, a sliding clamp and clamp loader, and a helicase which come together in the formation of two bidirectional active replication forks (Baker and Bell 1998). Although the end result of replication is the same among eukaryotic species, there is variability in the order and regulation of initiation steps between budding yeast, fission yeast, metazoans, and mammals.

Replication initiation begins in the early G1 phase of the cell cycle with the assembly of the pre-replicative complex (pre-RC) around origins of replication (Bell and Dutta 2002). The six-subunit origin recognition complex (ORC) first locates and binds DNA at the origin (Gilbert 2001), followed by loading of Cdc6 (cell division cycle 6) and

Cdt1 (cdc10-dependent transcript 1). Cdc6 is an AAA+ family member and is hypothesized to be a helicase loader due to its sequence homology to RFC (replication factor C, a well characterized replication clamp loader). The pre-RC is formed from the Cdc6- and Cdt1- dependent recruitment of the Mcm2-7 (minichromosome maintenance) complex (Maiorano, Moreau et al. 2000; Nishitani, Lygerou et al. 2000). After recruitment, several copies of Mcm2-7 are loaded onto each origin by the concerted hydrolysis of ATP via the activity of Cdc6 and several ORC subunits (Mendez and Stillman 2003; Blow and Dutta 2005; Randell, Bowers et al. 2006). These proteins are thought to "crack" the hexameric ring open and close the ring back after Mcm2-7 has encircled the dsDNA. The Mcm2-7 heterohexamer has been demonstrated to be the replicative helicase responsible for DNA unwinding during replication fork progression (Pacek and Walter 2004; Takahashi, Wigley et al. 2005; Pacek, Tutter et al. 2006; Bochman and Schwacha 2008). The loading of Mcm2-7 completes formation of the pre-RC. Pre-RCs assembled at the origins signal a "licensed" state of chromatin that must then be activated for initiation to continue.

Although the origins now exist in a licensed state, the action of many kinases (i.e. CDKs (cyclin dependent kinases), DDKs (Dbf4 dependent kinases), and others) are required to activate the pre-RCs. In *S. pombe*, Cdc7 and Cdk2 hyperphosphorylate the N-terminus of Mcm4 in an Mcm10 and Cdc45 dependent manner (Sheu and Stillman 2006), however in *S. cerevisiae* Mcm2, Mcm4, and Mcm6 are phosphorylated by DDK independent of Mcm10 (Francis, Randell et al. 2009). The phosphorylation of Mcm2-7 occurs at the transition from G1 to S phase and serves as a signal for recruitment of another set of initiation factors so that replication can begin. Pre-RC activation by

Mcm10 is the first committed step in S-phase and allows for Cdc45 binding, which in turn results in origin unwinding, followed by the recruitment of the eukaryotic single-stranded DNA binding protein replication protein A (RPA), the helicase co-factor GINS (*go ichi ni san,* Japanese for 5, 1, 2, 3), and DNA polymerase α-primase (pol α) to the origin (Wohlschlegel, Dhar et al. 2002; Gambus, Jones et al. 2006). Several groups have demonstrated that the phosphorylation-dependent interactions between Dpb11 (TopBP1), Sld2, and Sld3 (synthetic lethality with Dpb11) are required to maintain the association of pol α with chromatin (Araki, Leem et al. 1995; Masumoto, Sugino et al. 2000; Kamimura, Tak et al. 2001; Takayama, Kamimura et al. 2003; Tanaka, Tak et al. 2007). The addition of PCNA (proliferating cell nuclear antigen), RFC, and replicative DNA polymerases δ and ε with the origin completes the replisome (for review, see Garg and Burgers 2005).

Pol α, even with its many accessory factors, lacks processivity and thus must be replaced by more processive polymerases. The two DNA polymerases pol δ and pol ε replace pol α on the lagging and leading strands, respectively (Pursell, Isoz et al. 2007; Kunkel and Burgers 2008; Nick McElhinny, Gordenin et al. 2008). However, since DNA polymerases cannot synthesize DNA *de novo,* they must rely on the template made by the primase component of pol α. Pol α creates and RNA/DNA hybrid duplex ranging 20-30 nt in length (Eliasson and Reichard 1978; Hubscher, Maga et al. 2002). Pol δ and pol ε are loaded by a process that has been dubbed polymerase switching, whereby the clamp loader RFC loads PCNA, a ring-shaped clamp that interacts with the polymerases. PCNA serves to increase the processivity of the polymerases until they encounter a termination sequence or previously synthesized Okazaki fragments (Tsurimoto, Melendy

et al. 1990; Garg and Burgers 2005). Finally, the RNA primers are excised by RNase HI and FEN1 nucleases, the gap is filled by pol δ, and the nick is re-ligated by DNA ligase I (Waga and Stillman 1994).

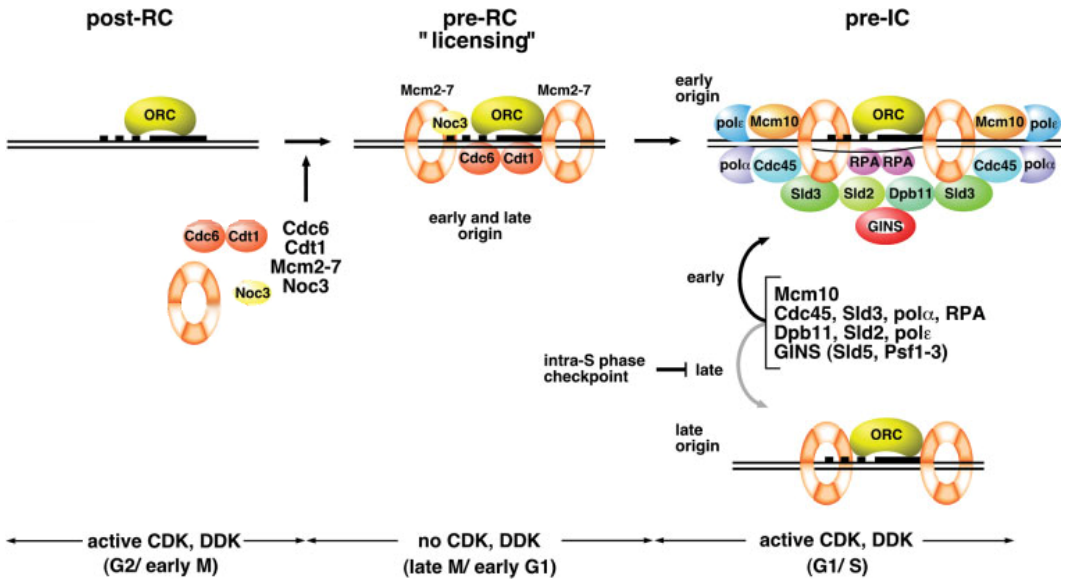


Figure 2. Two-step activation of an eukaryotic origin of replication. Multiple initiator proteins assemble at the origins of replication during the cell cycle. In the post-RC (non-competent) state, ORC is bound to the replicator sequences (darker boxes represent the conserved A, B1, B2, B3 elements identified in well-characterized yeast origins of replication (Kelly and Brown 2000)). Origin licensing occurs during a window of the cell cycle with little or no CDK activity, when Cdc6, Cdt1, Noc3 and MCM proteins are incorporated and forma pre-RC structure. Upon activation of CDKs and DDK, the pre-RC is activated to a pre-IC by the removal of Cdc6 and Cdt1 and the successive incorporation of multiple initiators. The architecture of the pre-IC shown is tentative. Formation of pre-ICs at the different origins follows a temporal program. Pre-IC formation at late origins could be prevented by the action of intra-S phase checkpoint mechanisms. Figure adapted from (Mendez and Stillman 2003).

The MCM Class of Proteins

Originally identified in *S. cerevisiae* by a screen for mutants defective in maintenance of minichromosomes (Merchant, Kawasaki et al. 1997), the MCM class of proteins is composed of ten factors conserved from yeast to mammals functioning in DNA replication. The most well known members of this class are the MCM2-7 proteins which form a hexameric complex that serves as a helicase. Mcm1 and Mcm10 do not belong to the same family as Mcm2-7 but are still involved in replication. Mcm1 is a

MADS box transcription factor that regulates the expression of Cdc6 and some MCM2-7 genes (Tye 1999), while Mcm10 is required for DNA replication (Cook, Kung et al. 2003). This family was recently expanded after the identification of Mcm8 and Mcm9 (Gozuacik, Chami et al. 2003; Yoshida 2005) which are closely related to Mcm2-7. The MCM2-9 family is identified by a conserved sequence that has become known as the MCM box (Figure 3). This sequence contains an AAA+ motif, a motif that includes Walker A and Walker B motifs. This motif is common in ATPases that facilitate DNA melting at promoters (Wedel and Kustu 1995). As expected, Mcm2-7 and Mcm8 have been shown to possess helicase activity (Kearsey and Labib 1998; Labib and Diffley 2001; Maiorano, Cuvier et al. 2005; Bochman and Schwacha 2008). The newly discovered Mcm9 however, has not been shown to contain helicase activity despite the fact that it contains the MCM box and is the sister paralogue of Mcm8 (Liu, Richards et al. 2009).

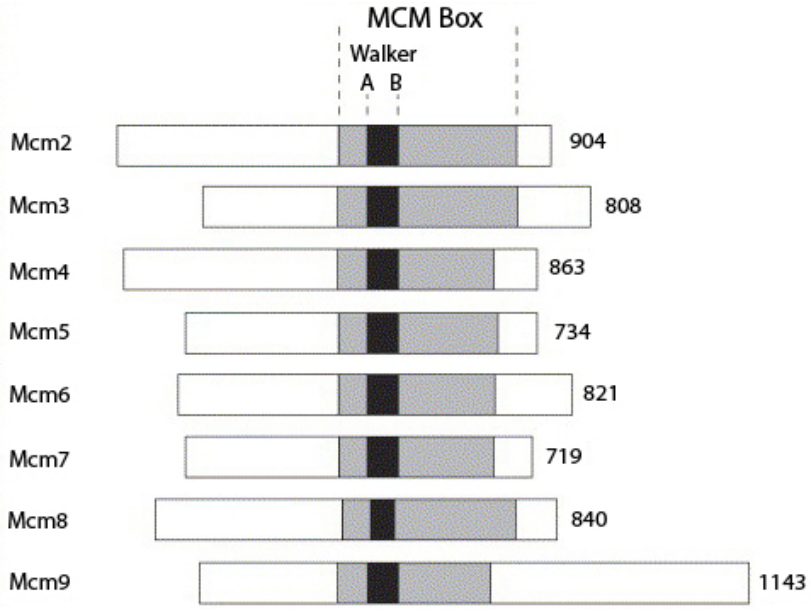


Figure 3. The Mcm2-9 superfamily. Alignment of human Mcm2-9 proteins. Bars represent the indicated proteins. The MCM Box is shown in grey, and the region encompassing the Walker A and B motif is shown in black. Numbers indicate amino acids. Figure adapted from (Maiorano, Lutzmann et al. 2006).

Mcm2-7 was presumed to be the replicative helicase because no one was able to observe *in vitro* helicase activity. The Mcm4,6,7 dimer of trimers was the only complex shown to have helicase activity *in vitro* (Ishimi 1997), but it was not highly processive. Mcm2-7 was finally demonstrated to have *in vitro* helicase activity but only in the presence of certain anions (Bochman and Schwacha 2008). It was also shown to have *in vitro* helicase activity when in complex with Mcm10 and Cdc45 (the CMG complex) (Moyer, Lewis et al. 2006). Later, the related Mcm8 protein was discovered and demonstrated to possess robust *in vitro* Mcm2-7-independent helicase activity at replication forks (Maiorano, Cuvier et al. 2005). This finding aroused intriguing possibilities for Mcm8 operating during elongation, after Mcm2-7 initiates DNA replication.

Mcm10 physically interacts with many of the subunits of Mcm2-7 and Cdc45 (Homesley, Lei et al. 2000; Christensen and Tye 2003; Ramachandran, Hainsworth et al. 2004). The best insights into these interactions came from early observations that the mcm10-1 mutation is suppressed by mutations in the genes that encode two subunits (Mcm5 and Mcm7) of Mcm2-7 (Homesley, Lei et al. 2000). These mutations (mcm7-1 and mcm5-461) were initially identified as suppressors of a Cdc45 mutant isolated from an unrelated screen (Moir and Botstein 1982; Hennessy, Lee et al. 1991). Furthermore, characterization of the mcm10-1 mcm7-1 double mutant showed that all of the defects had been corrected. These results suggest that an interaction between Mcm10 and Mcm7 is required for replication initiation, as well as the elongation phase. In addition, the type of mutual suppression observed in these studies suggests that Mcm10 and Cdc45 function in the same pathway. This information coupled with previous studies showing Mcm10

interacts with Mcm2-7 in late G1 phase while Cdc45 interacts with Mcm2-7 in S phase (Zou and Stillman 1998; Aparicio, Stout et al. 1999) suggests that there must be some hand-off mechanism among Mcm10, Mcm2-7, and Cdc45 when Cdc45 disengages Mcm2-7 from its anchor Mcm10 at the critical point when Mcm2-7 is being converted into an active helicase (Lei and Tye 2001).

The functions of the two newest members of this family, Mcm8 and Mcm9, still remain elusive. However, recent work has shed some light onto the possible roles of each protein. Mcm8 was shown to interact with Cdc6, ORC2, and Mcm7. Disruption of the interaction between Cdc6 and Mcm8 reduced the amount of Mcm2-7 that was loaded onto chromatin at G1, suggesting a possible role of Mcm8 in loading Mcm2-7 (Volkening and Hoffmann 2005; Kinoshita, Johnson et al. 2008). Additionally, Mcm8 possesses *in vitro* helicase activity, and colocalizes with RPA, suggesting that Mcm8 plays a role during replication elongation (Maiorano, Cuvier et al. 2005). While the data for Mcm8 is contradictory, the data for Mcm9 is not. Recently Mcm9 was demonstrated to have an opposing role to the inhibitory effects of Geminin. Mcm9 interacts with Cdt1 and serves to recruit Cdt1 to origins, and allows Mcm2-7 to be recruited (Lutzmann and Mechali 2008). When Mcm9 is depleted, Cdt1 is bound and inhibited by Geminin, and pre-RCs do not assemble. The tight binding between Cdt1 and Mcm9 suggests a role for Mcm9 in pre-RC assembly and beyond.

The Role of Initiation Factor Mcm10

Mcm10 was first identified by its effect on DNA synthesis and plasmid stability in yeast, (Maine, Sinha et al. 1984; Solomon, Wright et al. 1992) and was subsequently

found to be required for efficient initiation of DNA replication (Merchant, Kawasaki et al. 1997). Mcm10 has been shown in various organisms to localize to origins through interactions with pre-RC components ORC and Mcm2-7 (Izumi, Yanagi et al. 2000; Kawasaki, Hiraga et al. 2000). Mcm10 is an abundant chromatin-binding protein that interacts with many of the proteins intimately involved in replication initiation. Chromatin immunoprecipitation experiments suggest that Mcm10 is localized at replication origins during S-phase (Homesley, Lei et al. 2000, W. H. Chai and B. K. Tye, unpublished). Importantly, Mcm10 interacts genetically with Mcm2-7, DNA pol δ and ε, ORC, and Dpb11 (Spruck, Won et al. 1999; Labib and Diffley 2001; Arcus 2002; Tanaka and Diffley 2002). *In vitro*, interactions of Mcm10 with initiation factors ORC, Mcm2-7, Cdc45, and Cdc7/Dbf4 have been observed by co-immunoprecipitation from cell extracts (Arcus 2002; Tanaka and Diffley 2002; Wohlschlegel, Dhar et al. 2002; Vaziri, Saxena et al. 2003). Importantly, Cdc45 and RPA cannot load onto chromatin in Mcm10-depleted Xenopus egg extracts and prevent DNA unwinding (Kawasaki, Hiraga et al. 2000). Thus, the essential role of Mcm10 in initiation links the pre-replicative complexes with origin unwinding.

The role of Mcm10 in the transition from the pre-RC to the elongating state was first demonstrated in Xenopus egg extracts. These experiments demonstrated that Mcm10 loading onto chromatin requires the replicative helicase Mcm2-7, and that Mcm10 stimulates phosphorylation of Mcm2-7 by the Dbf4 (dumb bell forming 4)-dependent kinase Cdc7 (Wohlschlegel, Dhar et al. 2002; Lee, Seo et al. 2003). Interestingly, removal of Mcm10 from chromatin after assembly of the pre- RC results in the dissociation of Mcm2-7 from chromatin without affecting the association of ORC

(Lei and Tye 2001), suggesting Mcm10 functions to do more than stimulating phosphorylation of Mcm2-7. Mcm10 loading and Mcm2-7 phosphorylation are required for DNA unwinding and recruitment of other essential replication factors. Of these, Cdc45 (Mimura and Takisawa 1998; Zou and Stillman 1998) and the tetrameric GINS complex (Kubota, Takase et al. 2003; Takayama, Kamimura et al. 2003) are helicase co-factors that are required for Mcm2-7 catalyzed DNA unwinding at the replication fork (Pacek and Walter 2004; Gambus, Jones et al. 2006; Moyer, Lewis et al. 2006; Pacek, Tutter et al. 2006). However, a direct Mcm10-Cdc45 interaction has not been observed (B.F.E., unpublished), and thus the requirement of Mcm10 in Cdc45 loading may be the result of Mcm2-7 activation by Mcm10.

The point at which origin DNA is initially denatured or "melted" is undefined but is likely to occur between Mcm2-7 and Cdc45/GINS loading onto chromatin. Duplex unwinding by the Mcm2-7/Cdc45/GINS complex results in loading RPA, and pol α onto chromatin. It is possible that Mcm10 plays a direct role in DNA unwinding, since the Xenopus protein binds to both double- and single-stranded DNA (Robertson, Warren et al. 2008). S. pombe Mcm10 interacts with and enhances the priming activity of DNA pol α *in vitro* (Fien, Cho et al. 2004), and S. cerevisiae Mcm10 has recently been shown to associate with and regulate the stability of the catalytic subunit of DNA pol α *in vivo* (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007). In addition, Mcm10 interacts with other proteins involved in DNA synthesis, including ORC, Mcm7, Dpb11 (DNA polymerase B possible subunit 11), and DNA polymerases δ and ε (Merchant, Kawasaki et al. 1997; Homesley, Lei et al. 2000; Izumi, Yanagi et al. 2000; Kawasaki, Hiraga et al. 2000). The association of Mcm10 with DNA

and with components of the pre-RC and the replication fork suggests that it has multiple functions which may include recruiting other proteins to the replication fork during DNA unwinding. A recent study of the human homolog of Mcm10 suggests that Mcm10 dissociates from chromatin after pre-RC activation (Izumi, Yatagai et al. 2004).

The activation of Mcm2-7 and DNA binding by Mcm10 suggest a role in DNA manipulation. It was recently determined that Mcm10 is not part of the replicative helicase which excludes it from playing a role in DNA unwinding during fork progression (Pacek, Tutter et al. 2006). However, a role for Mcm10 in local origin melting cannot be ruled out. Another possible rationale for DNA binding by Mcm10 is to recruit downstream factors directly onto DNA. It is clear that a more thorough description of the Mcm10-DNA interaction is needed to clarify the significance of this function during the emergence of an active replisome.

In addition to replisome assembly, several lines of evidence suggest that Mcm10 is also required for replication fork progression through its association with DNA polymerases and DNA. The importance of Mcm10 in elongation is exemplified by its physical interaction with pol α (described below), and its genetic interactions with pol δ, pol ε, and DNA2 (Kawasaki, Hiraga et al. 2000; Liu, Choe et al. 2000). Mutations in yeast Mcm10 show defects in completion of S phase after origins have fired (Merchant, Kawasaki et al. 1997; Kawasaki, Hiraga et al. 2000). Furthermore, elongating forks pause at unfired pre-RCs in a *S. cerevisiae* Mcm10 mutant, suggesting that pre-RCs may present a barrier to fork progression that is overcome by Mcm10 action (Homesley, Lei et al. 2000). Although Cdc45 is required for fork progression, Mcm10's role in elongation is independent of Cdc45, since studies in *Xenopus* extracts have demonstrated that when

an elongating fork stalls, Mcm10 and DNA polymerases α, δ, and ε are uncoupled from the Cdc45/Mcm2-7/GINS helicase (Pacek, Tutter et al. 2006). Additionally, an interaction between diubiquitinated scMcm10 and PCNA is essential for replication in budding yeast (Das-Bradoo, Ricke et al. 2006).

In recent years Mcm10 has emerged as co-factor for DNA synthesis by pol α. Mcm10 physically interacts with p180, the large catalytic subunit in both *S. pombe* and *S. cerevisiae* systems (Fien, Cho et al. 2004; Ricke and Bielinsky 2004). Consequently, Mcm10 stimulates the *in vitro* DNA polymerase activity of pol α in *S. pombe* (Fien, Cho et al. 2004) and regulates the *in vivo* stability of the catalytic subunit in *S. cerevisiae* (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007). *In vitro*, spMcm10 interacts with and stimulates the catalytic activity of thep180 subunit of pol α (Fien, Cho et al. 2004), and has been shown to contain primase activity (Fien and Hurwitz 2006). Whether Mcm10 functions to recruit pol α to origins or to aid DNA synthesis requires further investigation of the Mcm10-pol α interaction.

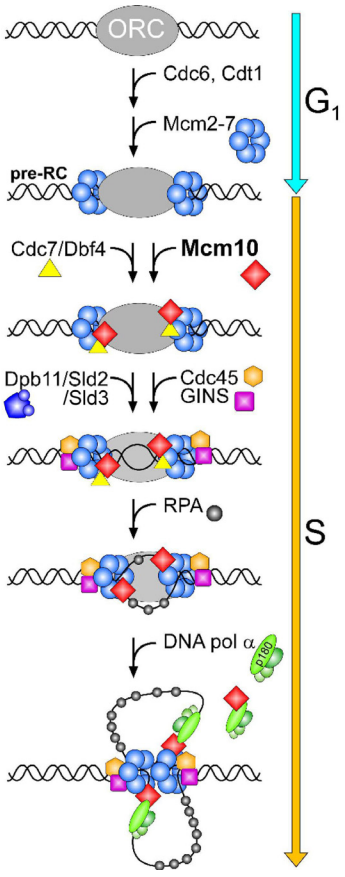


Figure 4. **Schematic of eukaryotic DNA replication initiation.** A schematic drawing depicting a generalized order of assembly of the replisome, starting with origin recognition, proceeding through origin melting, and leading up to replication initiation.

Polymerase α –primase

Pol α is a four-subunit DNA polymerase that is unique because it also contains a primase activity that allows it to synthesize nascent DNA strands *de novo* (Figure 5). Pol α initiates DNA synthesis by first synthesizing a short RNA primer and then extending it ~20 nucleotides using the polymerase activity of its large p180 subunit (Conaway and Lehman 1982; Conaway and Lehman 1982). All four subunits of pol α are conserved among eukaryotes and are necessary for cell viability in yeast (Sugino 1995). Each subunit of pol α is named according to its molecular weight: p180, p68, p58, and p48.

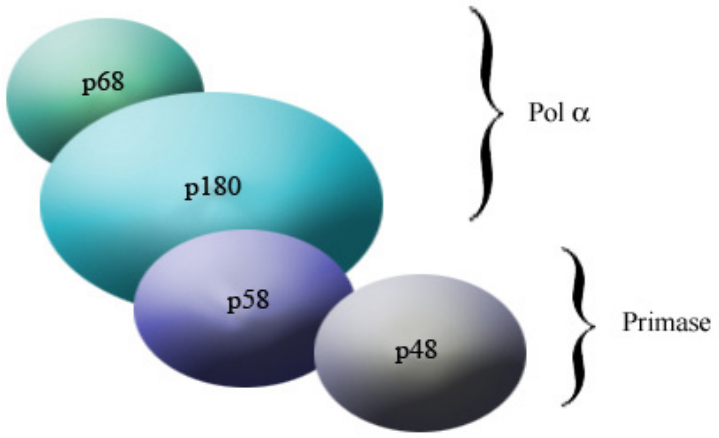


Figure 5. **Schematic of the subunits of pol α.** A depiction of the subunit organization of the two activities: polymerase and primase, that comprise DNA polymerase α. Adapted from http://www.reactome.org/figures/2.4.3a.jpg.

The largest (180 kDa) subunit, p180, contains the polymerase activity that is responsible for elongating the short 8-12 ribonucleotide primers synthesized by its primase subunit (p48) (Plevani, Foiani et al. 1985). Structural studies performed on mouse pol α identified three functional domains: an N-terminal domain (aa 1-329) of regulatory function, a core domain (aa 330-1279) which contains the polymerase function, and the carboxyl-terminal domain (aa 1235-1465) which is required for subunit assembly through interactions with p68 and p58 (Mizuno, Yamagishi et al. 1999). p68 is

an accessory subunit that has no known enzymatic activity. It is however required for import of pol α into the nucleus (Mizuno, Ito et al. 1998) and it is phoshorylated by Cyclin A-Cdk2 and Cyclin E-Cdk2 in a cell cycle dependent manner. Additionally, the phosphorylation of p68 by Cyclin A-Cdk2 has been demonstrated to affect the catalytic activity of p180 and its ability to interact with ORC and Cdc45 (Kukimoto, Igaki et al. 1999; Voitenleitner, Rehfuess et al. 1999; Uchiyama and Wang 2004; Takemura, Yoshida et al. 2006).

The p48 subunit of pol α contains the catalytic primase function. The primase starts by assembling a dinucleotide and then uses the template strand to extend its RNA primer to 8-12 ribonucleotides in length. The p48 subunit is then responsible for handing off the primed substrate to the polymerase active site on the p180 subunit for extension (Sheaff and Kuchta 1994; Sheaff, Kuchta et al. 1994). P48 interacts with a regulatory subunit p58 similarly to the way p180 interacts with p68. p58, like p68, has no known enzymatic function, but does stimulate p48's primase activity *in vitro* (Copeland 1997). Interestingly, p58 interacts with both the primer and the template, suggesting it might help to regulate the length of the primer and facilitate with the handoff to p180 (Arezi, Kirk et al. 1999).

The SV40 Replication System

Much of our knowledge about the initiation of DNA replication in eukaryotes stems from early studies in viral replication systems that set the foundation for understanding DNA replication (Challberg and Kelly 1989; Stillman 1989). One of the best characterized systems still used today is the simian virus 40 (SV40) system (Fanning

and Knippers 1992; Herendeen and Kelly 1996; Bullock 1997). These studies demonstrated that only three proteins are needed for DNA synthesis in the SV40 system: the large-T antigen (the viral initiator protein), RPA (the single-stranded DNA-binding protein), and pol α. The simplicity of this system affords investigation of the interactions and functions of replication factors that would otherwise be elusive. The ability of this viral protein, Large T-antigen (T-ag) to recruit cellular replication factors and initiate replication at origins is well characterized and thus provides a framework for understanding the higher complexity of eukaryotic systems.

T-ag is a 708 amino acid phosphopeptide that forms a double hexamer with each ring consisting of six identical subunits (Mastrangelo, Hough et al. 1989). T-ag is composed of four independently functional domains: an N-terminal J-domain (a.a. 1-102), an origin binding domain (OBD) (a.a. 131-259), a helicase domain (a.a. 260-627), and a C-terminal host-range (HR) domain (a.a. 628-708) (Figure 6). The combined efforts of at least 20 eukaryotic proteins are substituted by the action of the multifunctional viral protein T-ag. T-ag functions to recognize the replication origin (Dean, Dodson et al. 1987), bind to and melt the origin DNA (Borowiec and Hurwitz 1988), recruit pol α (Dornreiter, Hoss et al. 1990), and then to unwind the replication fork (Stahl, Droge et al. 1986).

Figure 6. **Schematic cartoon of T-ag domains.** The amino acid numbers are indicated at the bottom. The functional domains are represented by open boxes and are labeled accordingly. The linkers between domains are represented by thin lines. The Zinc-finger subdomain is indicated by a hatched box. The C-terminal domain from residue 628 to 708, which contains the host-range fragment (residues 682-708), is labeled HR. The HR domain is thought to be unstructured. Figure adapted from (Gai, Li et al. 2004).

In this manner, the functions executed by T-ag circumvent the need for the activity of origin recognition by ORC, helicase loading by Cdc6 and Cdt1, activation of the pre-RC by CDKs and DDKs, and polymerase recruitment by Mcm10 and Cdc45. Thus, the combined efforts of these proteins are equivalent to the functions of a single protein that can initiate multiple rounds of replication during S phase. The simplicity of the SV40 model system suggests that the machinery for replication initiation has grown more complex as the eukaryotic host has evolved to become more complex. To accomplish this feat, additional proteins are needed to regulate the assembly and activation of the pre-RC.

Scope of this Work

The work presented in this dissertation describes the studies performed on the replication initiation protein Mcm10 from *Xenopus laevis* (the African clawed frog) – xMcm10. Chapter 2 describes the biochemical characterization of xMcm10 and the elucidation of the domains contained in this protein. Chapter 3 describes the structure determination of a critical DNA binding domain of xMcm10 (xMcm10-ID) by X-ray crystallography. Chapter 3 details the biochemical characterization of this domain's DNA binding activity by fluorescence anisotropy and NMR (Nuclear Magnetic Resonance) spectroscopy. Chapter 4 describes the determination of a DNA-bound crystal structure of xMcm10-ID and the investigation of the interplay between this domain, ssDNA, and the catalytic subunit of p180. Finally, chapter 5 describes current and future developments with other domains of xMcm10 such as the purification and crystal trials with a dual-domain construct of xMcm10. This chapter also contains future directions and a closing discussion.

CHAPTER II

DOMAIN ARCHITECTURE AND BIOCHEMICAL CHARACTERIZATION OF VERTEBRATE MCM10*

Abstract

Mcm10 plays a key role in initiation and elongation of eukaryotic chromosomal DNA replication. As a first step to better understand the structure and function of vertebrate Mcm10, we have determined the structural architecture of *Xenopus laevis* Mcm10 (xMcm10) and characterized each domain biochemically. Limited proteolytic digestion of the full-length protein revealed amino-terminal (NTD), internal (ID), and carboxy-terminal (CTD) structured domains. Analytical ultracentrifugation revealed that xMcm10 self-associates and that the NTD forms homodimeric assemblies. DNA binding activity of xMcm10 was mapped to the ID and CTD, each of which binds to single- (ss) and double-stranded (ds) DNA with low micromolar affinity. The structural integrity of xMcm10-ID and CTD is dependent on the presence of bound zinc, which was experimentally verified by atomic absorption spectroscopy and proteolysis protection assays. The ID and CTD also bind independently to the amino-terminal 323 residues of the p180 subunit of DNA polymerase α-primase (pol α). We propose that the modularity of the protein architecture, with discrete domains for dimerization and for binding to DNA and pol α, provides an effective means for coordinating the biochemical activities of Mcm10 within the replisome.

* The work presented in this chapter was published in 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) *J Biol Chem.* **283**, 3338-3348.

Introduction

Eukaryotic DNA replication is carried out by large multiprotein machines that coordinate DNA unwinding and synthesis at the replication fork. Initiation of replication involves ordered assembly of the replisome and local denaturation of duplex DNA at the origin, followed by replisome activation. Screens for mutants defective in minichromosome maintenance (Mcm) and DNA replication in yeast identified a number of factors essential for replication (Nasmyth and Nurse 1981; Maine, Sinha et al. 1984; Solomon, Wright et al. 1992; Merchant, Kawasaki et al. 1997). Pre-replicative complexes (pre-RCs) composed of the origin recognition complex (ORC), Cdc6, Cdt1, and the hexameric Mcm2-7 helicase are assembled in G_1 (reviewed in Blow and Dutta 2005) and converted into active replication forks at the onset of S phase. Mcm10 loads onto chromatin after pre-RC assembly (Wohlschlegel, Dhar et al. 2002; Ricke and Bielinsky 2004) and stimulates phosphorylation of Mcm2-7 by Dbf4-Cdc7 kinase (Lee, Seo et al. 2003). Once Mcm10 is present, Cdc45 and GINS are loaded onto chromatin (Walter and Newport 2000; Wohlschlegel, Dhar et al. 2002; Takayama, Kamimura et al. 2003) and form a Cdc45/Mcm2-7/GINS helicase complex (Pacek and Walter 2004; Gambus, Jones et al. 2006; Moyer, Lewis et al. 2006; Pacek, Tutter et al. 2006). Cyclin- and Dbf4-dependent kinases, together with Sld2, Sld3, and Dpb11 in budding yeast (Tanaka, Umemori et al. 2007; Zegerman and Diffley 2007), stimulate origin unwinding, which is signified by recruitment of RPA to single-stranded DNA (Tanaka and Nasmyth 1998; Zou and Stillman 2000). Mcm10, Cdc45, and RPA facilitate subsequent loading of DNA polymerase α-primase (pol α) onto chromatin (Mimura and Takisawa 1998; Walter and Newport 2000; Ricke and Bielinsky 2004; Yang, Gregan et al. 2005). The

association of PCNA, RFC, and replicative DNA polymerases δ and ε with the origin completes the replisome (reviewed in Garg and Burgers 2005).

A number of interactions have been observed between Mcm10 and proteins found in the pre-RC and at the replication fork. Mcm10 is a component of active replication complexes in *Xenopus* and budding yeast (Gambus, Jones et al. 2006; Pacek, Tutter et al. 2006) and is associated with chromatin throughout S-phase (Ricke and Bielinsky 2004). Mcm10 interacts genetically with Mcm2-7, DNA pol δ and ε,. ORC, and Dpb11 (Merchant, Kawasaki et al. 1997; Homesley, Lei et al. 2000; Izumi, Yanagi et al. 2000; Kawasaki, Hiraga et al. 2000). *In vitro*, interactions of Mcm10 with initiation factors ORC, Mcm2-7, Cdc45, and Cdc7/Dbf4 have been observed by co-immunoprecipitation from cell extracts (Homesley, Lei et al. 2000; Kawasaki, Hiraga et al. 2000; Christensen and Tye 2003; Lee, Seo et al. 2003). Importantly, Cdc45 and RPA cannot load onto chromatin in Mcm10-depleted *Xenopus* egg extracts, preventing DNA unwinding (Wohlschlegel, Dhar et al. 2002). Thus, the essential role of Mcm10 in initiation links the pre-RC with origin unwinding.

Several lines of evidence suggest that Mcm10 migrates with the elongating replication fork through association with DNA polymerases and DNA. *Schizosaccharomyces pombe* Mcm10 (spMcm10) affects chromatin binding and sub-nuclear distribution of pol α (Gregan, Lindner et al. 2003; Yang, Gregan et al. 2005), and *Saccharomyces cerevisiae* Mcm10 (scMcm10) has been shown to interact with and stabilize the catalytic subunit of pol α *in vivo* (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006). *In vitro*, spMcm10 interacts with and stimulates the activity of the catalytic (polymerase) subunit of pol α (Fien, Cho et al. 2004), and has been shown to

contain primase activity (Fien and Hurwitz 2006). Additionally, an interaction between diubiquitinated scMcm10 and PCNA is essential for replication in budding yeast (Das-Bradoo, Ricke et al. 2006). Finally, spMcm10 binds to single (ss)- and double-stranded (ds) DNA *in vitro*, and DNA binding activity is localized in the N-terminal 300 residues of the protein (Fien, Cho et al. 2004). The interactions between Mcm10, DNA, and pol α have led to the suggestion that Mcm10 helps to recruit pol α to the replisome and may regulate its activity. Studies in *Xenopus* extracts have demonstrated that when an elongating fork stalls, Mcm10 and DNA polymerases α, δ, and ε are uncoupled from the Cdc45/Mcm2-7/GINS helicase (Pacek, Tutter et al. 2006).

Sequence alignments of Mcm10 from divergent eukaryotes show stretches of consecutive residues that are phylogenetically conserved (Figure 7A), suggesting that these regions may be important to the structure and function of the protein. Mcm10 from metazoa contains ~100-300 residues not present in the yeast proteins, and conservation from yeast to human is limited to ~200-amino acids in the middle of the protein. Consistent with Mcm10's DNA binding activity, the conserved central region contains an invariant CCCH zinc-binding motif (Homesley, Lei et al. 2000; Izumi, Yanagi et al. 2000; Cook, Kung et al. 2003) and a putative oligonucleotide/oligosaccharide binding (OB)-fold (Ricke and Bielinsky 2006).

The lack of sequence similarity outside of the central region raises a question of whether the function of Mcm10 is conserved from yeast to metazoa. In the present study, we report the first structure-function analysis of vertebrate Mcm10 using the *Xenopus laevis* protein (xMcm10). Limited proteolytic digestion of xMcm10 revealed the protein to be composed of at least three structural domains—an amino-terminal domain (NTD)

that forms homodimers in solution, and highly conserved internal (ID) and carboxy-terminal domains (CTD) that bind to ssDNA, dsDNA, and to the p180 subunit of pol α. Our results confirm and extend previous work from yeast, and suggest that vertebrate Mcm10 contains a CTD not present in the yeast orthologs.

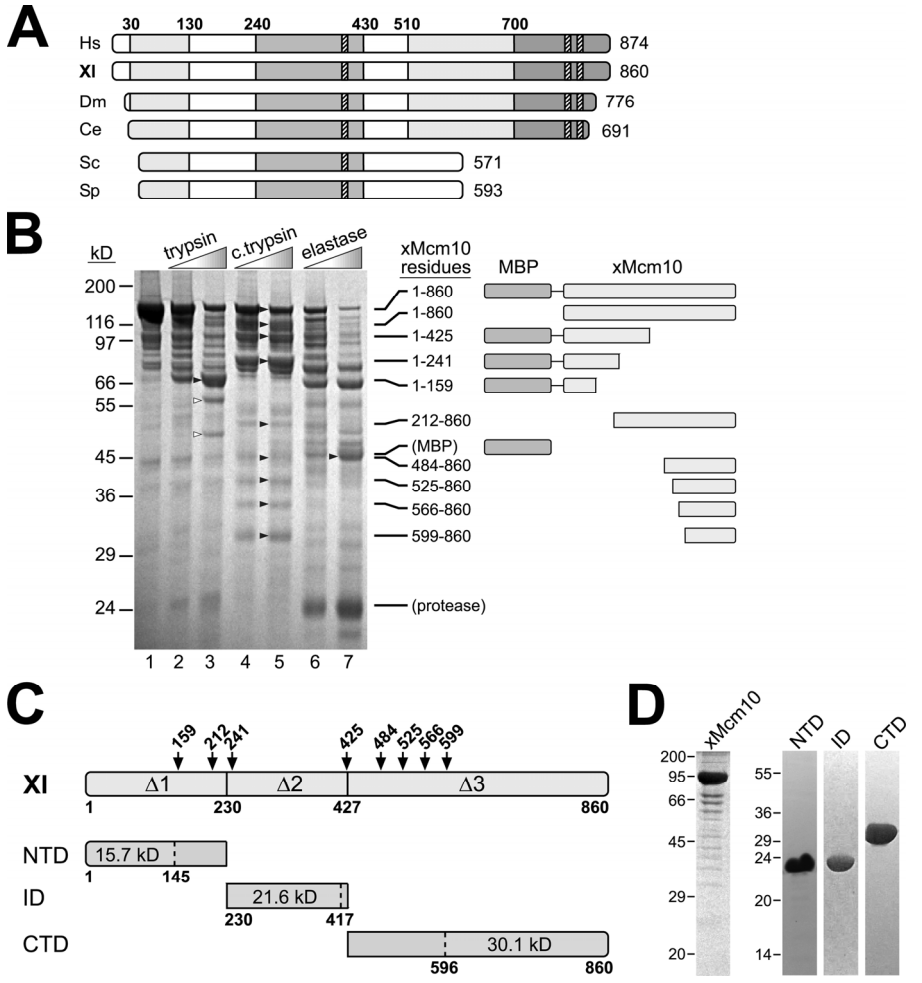


Figure 7. Domain architecture of Mcm10. (**A**), schematic alignment of Mcm10 sequences from *Homo sapiens* (Hs), *X. laevis* (Xl), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce)*, S. cerevisiae* (Sc), and *S. pombe* (Sp). Light and dark gray bars indicate moderate and high sequence conservation, respectively, and hatched boxes represent invariant cysteine/histidine clusters likely involved in zinc coordination. (**B**), limited proteolytic digestion of xMcm10. 50 pmol MBP-xMcm10 (lane 1) was subjected to proteolysis by trypsin (25 and 100 ng, lanes 2 and 3), chymotrypsin (c.trypsin) (100 and 200 ng, lanes 4–5), and elastase (10 and 25 ng, lanes 6 and 7) and visualized by SDS-PAGE. Major proteolytic fragments marked with black arrowheads were unambiguously identified by MALDI-TOF and TOF/TOF tandem mass spectrometry (MS) and are shown schematically to the right. Bands marked with white arrowheads contained several co-migrating Mcm10 fragments. See Appendix 1 for the full peptide coverage map used to identify fragment endpoints.

Figure 7, continued. (**C**), three truncation fragments (Δ 1, Δ2, Δ 3) of xMcm10 were purified and subjected to limited proteolysis to reveal stable domains NTD, ID, and CTD. Proteolytically sensitive sites identified in panel B are highlighted with arrows on top of the full-length protein schematic. Molecular masses and N-terminal sequences shown for each proteolytic fragment were identified by MS and Edman degradation. (**D**), Coomassie-stained SDS-PAGE of purified full-length xMcm10, NTD, ID, and CTD used in this study.

Results

xMcm10 Contains Three Structural Domains

In the current study, experiments to characterize the domain architecture of vertebrate Mcm10 were carried out using the *Xenopus laevis* ortholog because of previous investigations of the function of the protein using *Xenopus* egg extracts (Walter, Sun et al. 1998; Wohlschlegel, Dhar et al. 2002). Homology exists in three distinct regions of the protein (Figure 7A, Supplementary Figure A1). The internal region (aa 240-430) is highly conserved among all known Mcm10 orthologs, with an overall similarity of 21.3% (39.0% for non-yeast Mcm10). Likewise, the carboxy-terminus contains a region of high (aa ~700-860) and moderate (aa 510-700) similarity among higher eukaryotes. However, this region is not present in the yeast proteins (23.3% similarity for metazoan as compared to 3.6% for all eukaryotes). Moderate sequence similarity also exists at the amino-terminus (10% similarity for aa 1-130 in non-yeast sequences). This sequence analysis immediately suggested the presence of at least three domains tethered by disordered linkers. Consistent with this, no secondary structure was predicted in regions 130-230 and 575-624 (Supplementary Figure A1), and region 130-230 was predicted to be largely disordered.

In order to experimentally determine the domain organization of Mcm10, the full-length protein was overexpressed in *E. coli* with a cleavable N-terminal maltose binding

protein (MBP) tag and a C-terminal His_6 tag. The purified MBP-xMcm10-His_6 protein was subjected to limited proteolytic digestion by trypsin, chymotrypsin, and elastase, and the major proteolytic fragments identified by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) and MALDI-TOF/TOF tandem MS (Figure 7B). Peptide masses were mapped to the xMcm10 amino acid sequence to define domains (Supplementary Figure A2). In most cases the endpoint regions were defined by peptide ions that were present in the full-length protein but absent in the fragment under study, and in some cases the endpoint was confirmed with tandem MS on unique peptide(s) that were generated by chymotrypsin cleavage on one side (from limited proteolysis) and trypsin cleavage on the other (from in-gel digestion). Peptides analyzed in this way revealed proteolytic-resistant domains separated by cleavage sites at amino acids 159, 241, 425, 484, 525, 566, and 599 (Figure 7B and Supplementary Figure A2).

Using the proteolytically sensitive regions as a guide, three deletion constructs encompassing the entire protein were designed in order to define the domain boundaries more accurately: $xMcm10^{1\text{-}230}$ (Δ1), $xMcm10^{230\text{-}427}$ (Δ2), and $xMcm10^{427\text{-}860}$ (Δ3). Each of these proteins were expressed in bacteria, purified, and subjected to limited proteolysis by trypsin (Supplementary Figure A3). Precise endpoints of tryptic fragments were identified by Edman degradation and MALDI mass spectrometry (Figure 7C). Chymotrypsin, elastase, and endoproteinase-Glu-C digestion was also performed (data not shown). Despite the unique specificities of each protease tested, the resulting cleavage patterns were similar for each Mcm10 deletion mutant. Proteolysis of each deletion mutant revealed the presence of smaller fragments that were resistant to digestion and that were consistent with the cleavage pattern of the full-length protein

(Figure 7B) and with regions of sequence conservation (Figure 7A). Cleavage of the C-terminal ends of Δ1 and Δ2 yielded xMcm10$^{1-145}$ and xMcm10$^{230-417}$, respectively. For Δ3, ~170 residues were cleaved from the N-terminus, yielding xMcm10$^{596-860}$. The resistance of xMcm10$^{1-145}$, xMcm10$^{230-417}$, and xMcm10$^{596-860}$ to further degradation indicates the presence of stable tertiary folds that sterically preclude protease access to their cleavage sites. To prepare for further characterization, regions 1-145 (NTD), 230-417 (ID), and 596-860 (CTD) were sub-cloned, overexpressed, and purified (Figure 7D). The anomalous electrophoretic mobility of the NTD can be rationalized on the basis of the predicted pI (4.2) and elongated shape of the protein (see below). The NTD, ID, and CTD were relatively stable to further proteolytic digestion, and circular dichroism spectra confirmed the presence of secondary structure in each domain (data not shown).

Dimerization of xMcm10-NTD

Purified scMcm10 and spMcm10 have been reported to oligomerize in solution (Cook, Kung et al. 2003; Lee, Seo et al. 2003; Fien and Hurwitz 2006), and human Mcm10 was recently reported to form hexameric assemblies (Okorokov, Waugh et al. 2007). Prior to a rigorous analysis of xMcm10 oligomerization, we first investigated the hydrodynamic properties of the full-length, NTD, ID, and CTD proteins by gel filtration chromatography (Supplementary Figure A4). The elution volumes of full-length and NTD proteins were considerably less than expected for globular, monomeric proteins. Similarly, the CTD showed a modest decrease in retention volume as compared to that of a 30-kD protein standard. The elution profile of the ID, on the other hand, corresponded exactly to that of a 22-kD protein, indicating that this domain does not self-associate.

These results raised the question of whether xMcm10 oligomerizes in solution or whether the shape of the protein significantly deviates from a globular fold.

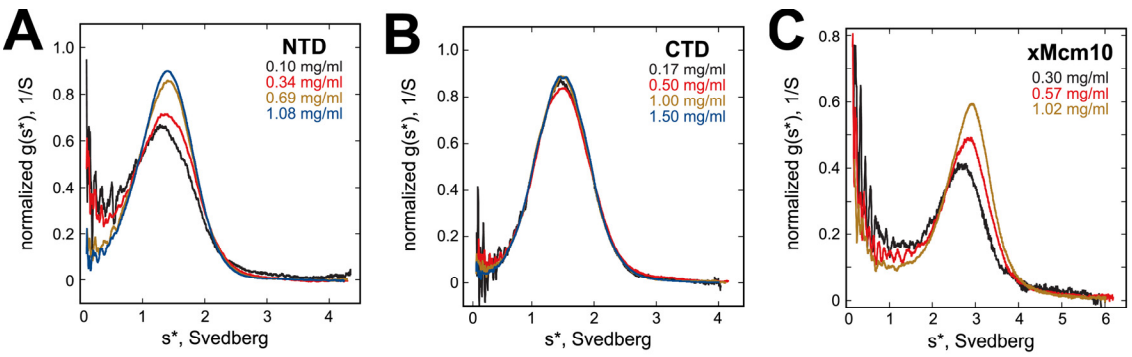


Figure 8. Self-association of xMcm10. Shown are overlays of normalized g(s*) plots from sedimentation velocity experiments at different concentrations of xMcm10-NTD (**A**), CTD (**B**), and full-length enzyme (**C**). NTD and CTD were prepared in 20mMTris, pH7.5, 100mMNaCl, 3.5mM β-mercaptoethanol, and 5% glycerol; and full-length enzyme was prepared in 20 mM Tris, pH 7.5, 500mMNaCl, 1mMdithiothreitol, 5%glycerol. Conditions: rotor speed, 55,000 rpm; temperature, 20 °C; interference optics.

The oligomeric states of the NTD, CTD and full-length proteins were determined using sedimentation velocity experiments (Figure 8). Figure 8A shows an overlay of the normalized g(s*) sedimentation coefficient distributions for four concentrations of the NTD. The distributions shift to the right with increasing concentration, indicating reversible self-association. The best fit to the data was obtained using a monomer-dimer equilibrium model. The sedimentation coefficient for the monomer could not be accurately determined due to the fact that the protein is predominantly dimeric over the concentration range tested. Thus, the sedimentation coefficient ratio s(dimer) / s(monomer) was fixed at 1.45, which is the value predicted for a monomer-dimer system (Garcia de la Torre and Bloomfield 1981). The best fit parameters are $s_{20,w}$ (monomer) = 1.22 S, $s_{20,w}$ (dimer) = 1.77 S, a dissociation constant of K_d = 3.1 μM and an rms error of 0.0048 mg/ml. The corrected sedimentation coefficients of the monomer and dimer can

be used to calculate frictional ratios, f/f_0, of 1.8 and 2.0, respectively, indicating that the NTD is highly asymmetric.

The normalized g(s*) profiles for the CTD superimpose over the concentration range tested (0.17-1.5 mg/ml), indicating that the system does not undergo reversible association under these conditions. The molecular weight obtained from a global fit of the data to single species model is 31.0 kD, which agrees closely with the predicted monomeric value of 30.1 kD. The frictional ratio (f/f_0) of 1.89 indicates that CTD is also quite asymmetric, consistent with its gel filtration behavior.

Figure 8C shows the normalized g(s*) distributions for the full-length enzyme. Like NTD, the distributions shift to the right with increasing concentration, indicating mass-action association. In this case, the presence of lower- and higher-S contaminants precludes further analysis of these data. However, the limiting sedimentation coefficient of ~2.6 S at low concentration indicates that xMcm10 is predominantly monomeric at low concentrations with $f/f_0 \sim 2.2$. Assuming an alternative model where the s=2.6 S species is a dimer yields an unreasonably high $f/f_0 \sim 3.5$.

Zinc-Dependent Stability of xMcm10-ID and CTD

Sequence alignments show clusters of highly invariant cysteine and histidine residues in both the ID and CTD (Figure 9A), suggesting that these domains contain zinc binding motifs. Strong evidence has been provided for the presence of a zinc motif in scMcm10 internal region (Cook, Kung et al. 2003), although zinc binding by the CTD has not yet been reported. To verify the presence and determine the stoichiometry of Zn^{2+} in xMcm10 domains, we analyzed each of the domains by graphite furnace atomic

absorption (GFAA) spectroscopy. Molar ratios of Zn^{2+}/xMcm10 for the NTD, ID, and CTD were determined to be 0.16, 1.3 ± 0.3, and 1.8 ± 0.5, respectively (Table 1). As a positive control, 3-methyladenine DNA glycosylase I (TAG), which has been shown previously to contain 1 Zn^{2+}/molecule (Kwon, Cao et al. 2003; Metz, Hollis et al. 2007), was analyzed by GFAA and returned a value of 0.98 Zn^{2+}/TAG. We therefore conclude that the NTD, CTD, and ID contain 0, 1, and 2 Zn^{2+} ions, respectively. In support of the GFAA data, X-ray fluorescence emission spectra of xMcm10-ID single crystals, which were grown in the absence of Zn^{2+} in the crystallization buffer, revealed a strong peak at 9.6 keV corresponding to the Zn^{2+} absorption edge (data not shown).

Table 1. Molar equivalents of Zn^{2+} in xMcm10 domains

Protein	Zn^{2+} / xMcm10
xMcm10-NTD	0.16
xMcm10-ID	1.3 ± 0.3
xMcm10-CTD	1.8 ± 0.5
TAG (control) [a]	0.98

[a] (Kwon, Cao et al. 2003; Metz, Hollis et al. 2007)

The importance of bound zinc on the tertiary folding of the ID and CTD was investigated by limited proteolysis protection assays. The ID and CTD were subjected to proteolysis by elastase in the presence and absence of ethylenediamine tetraacetic acid (EDTA), a known Zn^{2+} chelator. Both domains were more readily degraded in the presence of EDTA (Figure 9B), suggesting that in the absence of bound Zn^{2+}, the ID and CTD were at least partially unfolded and thus more susceptible to protease cleavage. Similarly, when the ID and CTD were incubated at room temperature for 10 days in the presence or absence of EDTA, spontaneous degradation was increased in the presence of EDTA (Supplementary Figure A5). These results suggest that the zinc motifs in

xMcm10-ID and –CTD play a key role in maintaining the overall structural integrity of these domains.

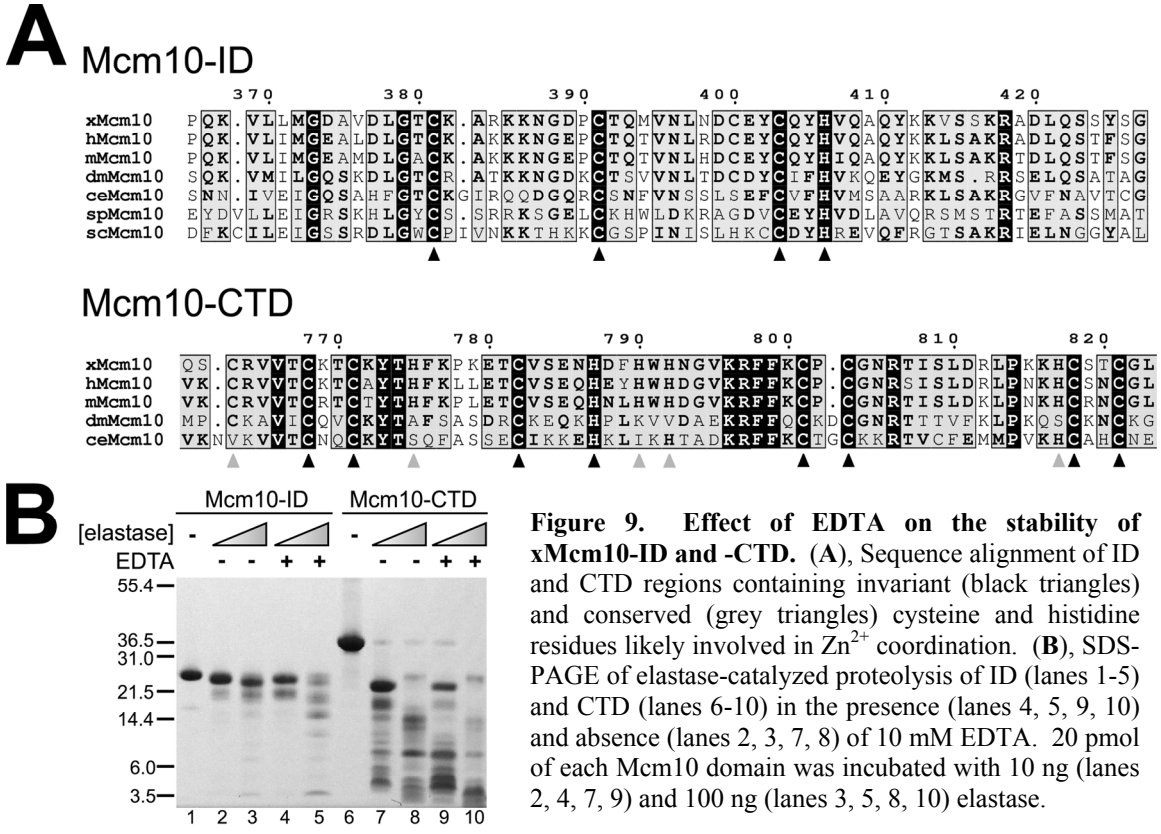


Figure 9. Effect of EDTA on the stability of xMcm10-ID and -CTD. (**A**), Sequence alignment of ID and CTD regions containing invariant (black triangles) and conserved (grey triangles) cysteine and histidine residues likely involved in Zn^{2+} coordination. (**B**), SDS-PAGE of elastase-catalyzed proteolysis of ID (lanes 1-5) and CTD (lanes 6-10) in the presence (lanes 4, 5, 9, 10) and absence (lanes 2, 3, 7, 8) of 10 mM EDTA. 20 pmol of each Mcm10 domain was incubated with 10 ng (lanes 2, 4, 7, 9) and 100 ng (lanes 3, 5, 8, 10) elastase.

xMcm10-ID and CTD are DNA Binding Domains

In order to quantitatively characterize the DNA binding activity of purified xMcm10, the change in fluorescence anisotropy was monitored as the protein was added to a fluorescein-labeled 25mer oligonucleotide (Figure 10). Binding isotherms for MBP-xMcm10-His_6 show that the full-length *Xenopus* protein bound to both ssDNA and dsDNA with the same affinity (K_d ~ 0.1μM) (Figure 10A, Table 2). To determine if Mcm10 might bind to the replication fork at the ss/dsDNA junction, a forked substrate containing both ssDNA and dsDNA regions was also tested and did not show a difference

in binding affinity (K_d = 0.08 ± 0.06 μM) compared to ssDNA and dsDNA (Table 2). Interestingly, in the presence of EDTA, binding of xMcm10 to dsDNA was abolished, whereas the affinity for ssDNA remained unchanged (Figure 10A, Table 2). The overall anisotropy change for ssDNA binding was different between EDTA and non-EDTA titrations, indicating that a change in the tumbling rate of the complex occurred, likely as a result of EDTA-induced local unfolding of the zinc motifs (Figure 9). These results establish that zinc-dependent structural integrity of xMcm10 is important for the dsDNA binding activity.

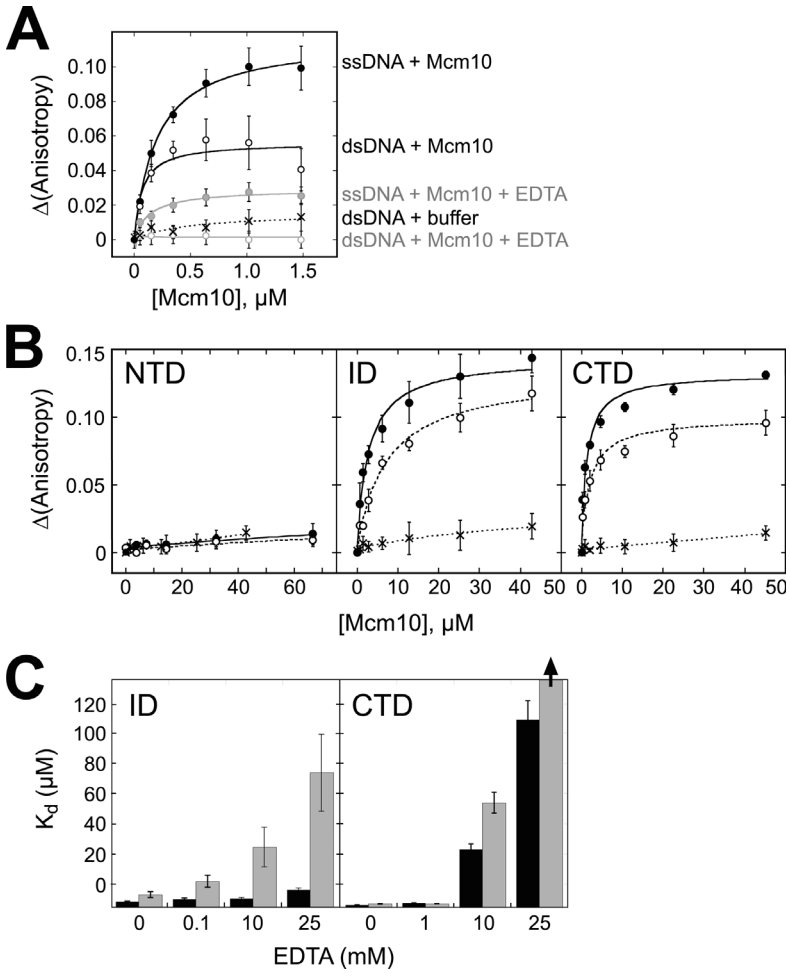


Figure 10. DNA binding of xMcm10. Binding was monitored as a change in fluorescence anisotropy as full-length (**A**) and isolated domain (**B,C**) proteins were titrated into a solution containing fluorescently labeled DNA.

Figure 10, continued. Error bars represent the standard deviation from the average values from three independent measurements. (**A**), Binding isotherms for full-length MBP-xMcm10-His_6 binding to ssDNA (filled symbols) and dsDNA (open symbols) in the absence (black) and presence (grey) of EDTA. A control in which buffer without protein was added to the DNA is shown as black Xs. (**B**), Binding curves for each xMcm10 domain against ssDNA (closed circles) and dsDNA (open circles), and for buffer-only controls (Xs). (**C**), The dissociation constants (K_d) for xMcm10-ID and –CTD binding to ssDNA (black bars) and dsDNA (grey bars) derived from the anisotropy data are plotted as a function of EDTA concentration. The K_d for xMcm10-ID/dsDNA binding in 25 mM EDTA is $\geq$ 300 μM, the limit of detection for this assay.

Binding of DNA to the NTD, ID, and CTD was then measured in order to determine the DNA binding domain of xMcm10. No anisotropy change was observed in the presence of the NTD, indicating that this domain does not interact with DNA (Figure 10B). Unexpectedly, both the ID and CTD showed robust binding to both ssDNA and dsDNA (Figure 10B). The affinity of each domain for DNA was roughly the same and was an order of magnitude less than that of the full-length protein (Table 2). Unlike full-length xMcm10, the affinity of each domain for ssDNA was ~2-fold greater than for dsDNA. In order to test the effect of the Zn^{2+} motifs, binding experiments for each domain were again carried out in the presence of EDTA (Figure 10C). Both xMcm10-ID and –CTD exhibited a dramatic decrease in dsDNA binding affinity as a function of increasing EDTA concentration, whereas the ssDNA affinity was only moderately affected under the same conditions (Figure 10C). Interestingly, EDTA had a greater affect on ssDNA binding to the CTD than the ID, suggesting that ssDNA is able to bind to the ID in the absence of a folded zinc motif.

Table 2. Dissociation constants for DNA binding [a]

	ssDNA [c]	dsDNA [c]	Fork [d]	Bubble [e]
xMcm10 [b]	0.12 ± 0.02	0.09 ± 0.03	0.08 ± 0.06	n.d.
xMcm10 [b] + 10 mM EDTA	0.14 ± 0.04	≥ 300	n.d.	n.d.
xMcm10-NTD	≥ 300	≥ 300	n.d.	n.d.
xMcm10-ID	3.39 ± 0.49	7.83 ± 1.44	3.09 ± 0.99	5.21 ± 1.86
xMcm10-CTD	1.41 ± 0.24	2.21 ± 0.20	2.67 ± 0.34	4.77 ± 2.57

[a] K_d (μM) for xMcm10 binding to deoxyoligonucleotides determined using fluorescence anisotropy. N.d., not determined

[b] Binding data for full-length xMcm10 were measured using MBP-xMcm10-His_6

[c] 25mer single- (ss) and double stranded (ds) DNA substrates

[d] Forked DNA = $(dsDNA)_{25}$-2x$(ssDNA)_{25}$ for full-length and $(dsDNA)_{10}$-2x$(ssDNA)_{15}$ for domains

[e] Bubble DNA = $(dsDNA)_{10}$-2x$(ssDNA)_{15}$-$(dsDNA)_{10}$

xMcm10 Binding to DNA Polymerase α-Primase is Localized to the ID and CTD

We investigated whether vertebrate Mcm10 can undergo direct, physical interactions with pol α, and if so, to map these interactions with the xMcm10 domains. Since purified recombinant human pol α has been shown to substitute functionally for the *Xenopus laevis* protein in *in vitro Xenopus* replication assays (Michael, Ott et al. 2000), human pol α was chosen for these experiments (Figure 11A). The first experiment examined the ability of the purified four-subunit human pol α-primase complex immobilized on beads to capture His-tagged xMcm10 domains from solution. After incubation with purified xMcm10-NTD, ID, or CTD and extensive washing, xMcm10

domains remaining bound to the beads were detected by denaturing gel electrophoresis and anti-His Western blot. Figure 11B shows the results of the pol α-Mcm10 affinity capture, in which both the ID and CTD, but not the NTD, bound to the polymerase complex. The experiment was repeated using only the purified catalytic pol α-p180 subunit in the absence of p48, p58, and p68. Again the NTD was not detected in the bound fraction, and both the ID and CTD bound to p180 (Figure 11C). This result demonstrates that the p180 subunit is sufficient to bind xMcm10-ID and CTD.

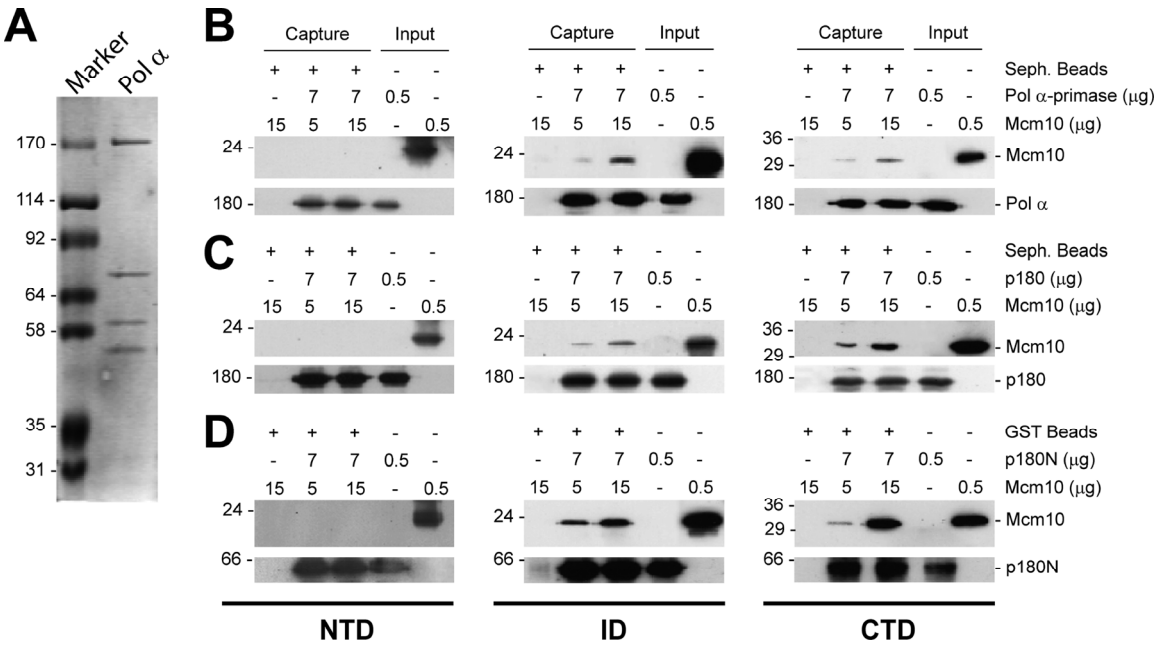


Figure 11. Binding of xMcm10 to the p180 subunit of DNA pol α. (**A**), Coomassie blue stained SDS-PAGE gel of purified pol α. (**B,C,D**), Affinity capture experiments between xMcm10-NTD (left panels), -ID (center panels), or –CTD (right panels) and pol α (**B**), p180 (**C**), or p180N (**D**). Amounts of protein added to each binding reaction are shown above the western blots. (**B**), The intact pol α complex was mixed with xMcm10 domains NTD, ID or CTD as indicated and immunoprecipitated on Sepharose beads coupled to SJK132-20 antibodies against the p180 subunit as indicated. Bound xMcm10 domains were detected by western blotting with an α-His antibody. (**C**), The purified catalytic p180 subunit of pol α was mixed with xMcm10 domains and immunoprecipitated as in *B*. Bound xMcm10 domains were detected by western blotting with an α-His antibody. (**D**), GST fused to the N-terminal 323 residues of p180 (p180N) was adsorbed on glutathione beads and mixed with xMcm10 domains as indicated. Bound xMcm10 was detected by western blotting with an α-His antibody.

We next sought to map the specific Mcm10-binding region of p180. The p180 subunit has a modular organization with a ~300 residue N-terminal region dispensable for polymerase activity, an extended core region containing the conserved polymerase motifs, and a C-terminal region that complexes with the other subunits (Mizuno, Yamagishi et al. 1999). Only the N-terminal region of p180 binds to SV40 T antigen, an interaction essential for viral DNA replication (Dornreiter, Copeland et al. 1993). Based on this information, an N-terminal construct encompassing p180 residues 1-323 (p180N) was tested. GST-tagged p180N immobilized on glutathione-sepharose was able to capture both the ID and CTD, but not the NTD, consistent with the pol α-primase and p180 pull-downs (Figure 11D). Thus, p180N is sufficient for Mcm10 interaction. These results also show that as for binding DNA, the ID and CTD function in a coordinated manner.

xMcm10 Does Not Contain Primase Activity

Based on the recent report that spMcm10 contains primase activity (Fien and Hurwitz 2006), we examined the ability of full-length xMcm10 to synthesize an oligoribonucleotide in the presence of a DNA template. Purified xMcm10 that contained no MBP-tag (Figure 7D) was incubated with dT_{50} template and [α-^{32}P]ATP, and product RNA visualized by denaturing PAGE. No radiolabeled products were apparent when compared to a no-enzyme control reaction (Figure 12). Under identical conditions, pol α-primase showed robust, concentration dependent formation of oligoribonucleotides ~12 nucleotides in length. This result indicates that a purified preparation of xMcm10 is not capable of priming DNA.

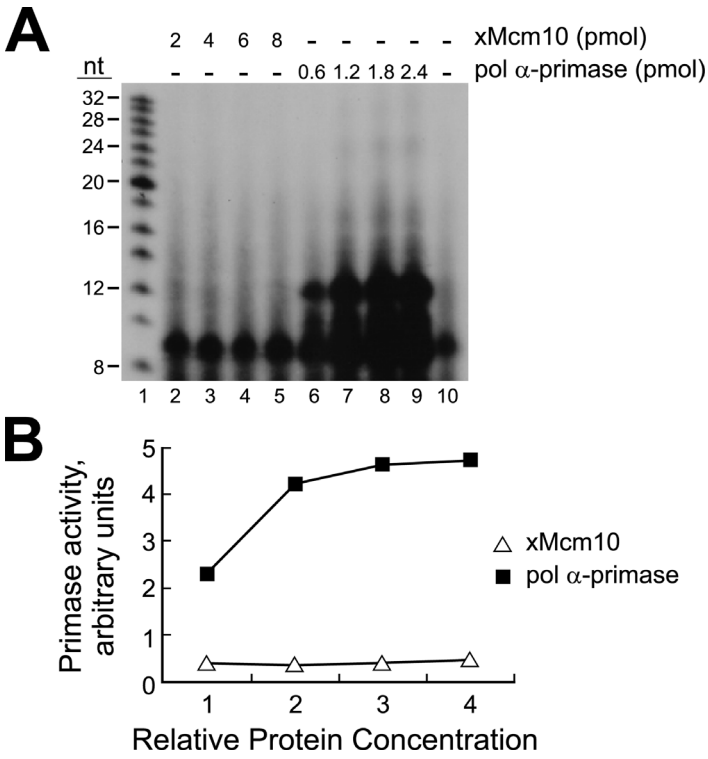


Figure 12. xMcm10 does not contain primase activity. (**A**), Oligoribonucleotide synthesis was assayed in reaction mixtures containing dT_{50} template, [α-^{32}P]ATP, and increasing amounts of xMcm10 (lanes 2-5) or pol α-primase (lanes 6-9). Lane 10, negative control lacking xMcm10 and pol α-primase. Radiolabeled products were analyzed by electrophoresis on a 25% denaturing polyacrylamide gel containing 7 M urea. (**B**), Quantitation of the autoradiogram shown in **A**. Primase activity is expressed in arbitrary units, with the reaction containing no protein set to zero. Relative protein concentration corresponds to 0.2, 0.4, 0.6, and 0.8 μM xMcm10 and 0.06, 0.12, 0.18, 0.24 μM pol α-primase.

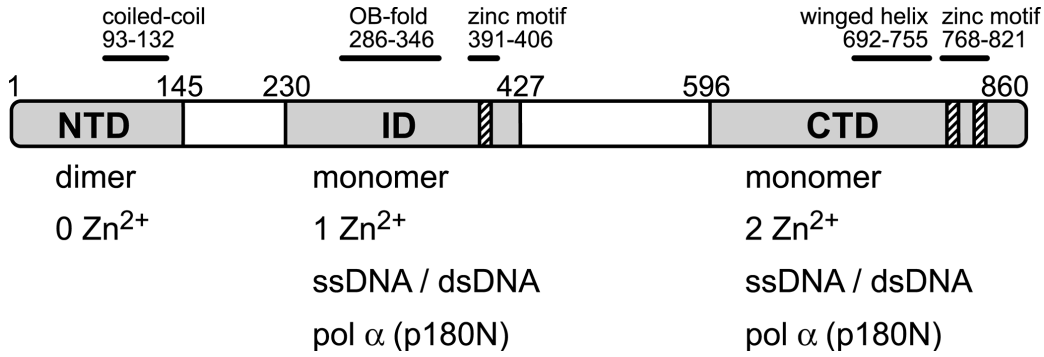


Figure 13. Vertebrate Mcm10. The schematic summarizes the domain organization and functional regions of xMcm10 identified in this study. The NTD, ID, and CTD are shaded grey, and conserved cysteine/histidine clusters predicted to chelate Zn^{2+} are shown as cross-hatched strips. Predicted structural motifs are shown as black bars above the protein. Listed below each domain are the oligomerization states, number of zinc ions bound, and binding partners.

Discussion

Modular architecture of Mcm10

The present work provides new insight into Mcm10's role in initiation and elongation complexes by carrying out the first structure-function analysis of the protein. We have determined using limited proteolysis that purified preparations of xMcm10 contain at least three structural domains located from residues 1-145 (NTD), 230-417 (ID), and 596-860 (CTD) (Figure 13). The extreme proteolytic sensitivity of regions 146-230 and 418-596 suggests that these are exposed flexible linkers connecting each independently folded globular domain. It is likely that these flexible regions become more structured or protected from proteolytic cleavage when Mcm10 is part of the larger multiprotein replisome assembly. Nevertheless, the present work suggests that Mcm10 is at least able to adopt multiple conformations in which each globular domain can move relative to the other two. Such a flexible protein architecture would be necessary for the multiple protein and DNA transactions at an inherently dynamic replication fork. Indeed, many replication proteins have evolved modular architectures with distinct domains that are able to act independently or cooperatively to perform a common task (reviewed in Stauffer and Chazin 2004; Fanning, Klimovich et al. 2006). For example, separate structural domains often provide multiple binding sites that increase the affinity for one ligand, or that enable the protein to contact multiple ligands in a concerted or sequential fashion (Arunkumar, Stauffer et al. 2003).

Structural features of Mcm10-ID and –CTD

Motifs predicted within the ID and CTD provide a rationale for their interactions with DNA and pol α (Figure 13). The protein structure prediction Protein Homology/analogy Recognition Engine (PHYRE) program (Kelley, MacCallum et al. 2000) and manual inspection of the xMcm10-CTD primary sequence identified two putative Zn^{2+}-binding motifs (aa 766-789, 799-821) and a 3-helical bundle from the winged helix superfamily (aa 692-755) (Supplementary Figure A1). These motifs were not identified in yeast Mcm10 proteins. Previously identified motifs in the conserved ID were also found by this method, including an OB-fold (aa 286-346) and zinc motif (391-406) (Homesley, Lei et al. 2000; Izumi, Yanagi et al. 2000; Ricke and Bielinsky 2006). Consistent with the ability of the ID and CTD to bind both DNA and pol α, OB-folds, winged helix bundles, and zinc motifs have each been shown to mediate protein-protein interactions in addition to their role in nucleic acid recognition (Leon and Roth 2000; Mer, Bochkarev et al. 2000; Stauffer and Chazin 2004).

The zinc-binding motifs are essential to the structure and function of Mcm10. Mutations in the putative CCCH-type zinc finger within the conserved ID have been shown to disrupt the association of scMcm10 with chromatin (Homesley, Lei et al. 2000), to cause growth defects in yeast, and to disrupt the NMR chemical shift dispersion of purified scMcm10 (Cook, Kung et al. 2003). Our atomic absorption data show conclusively that one molar equivalent of zinc is present in the ID, and reveal two additional zinc atoms bound to the CTD (Table 1). The effect of Zn^{2+} chelation on Mcm10 DNA binding activity and protein stability (Figures 8B, 9A, and 9C; Table 2)

helps to explain the dissociation of Mcm10 from chromatin in the *S. cerevisiae mcm10-43* (C320Y in the ID) mutant (Solomon, Wright et al. 1992; Homesley, Lei et al. 2000).

The arrangement of the invariant Cys/His clusters in the xMcm10-CTD into a **C**X_2**C**X_{10}**C**X_4**H**-(X_{13})-**C**X**C**X_{14}**C**X_2**C** consensus sequence (Figure 9A) raises several possibilities for the precise role of the CTD zinc motifs. On one hand, the sequences of each CCCH or CCCC cluster do not deviate significantly from the classical DNA sequence-specific **C**X_2**C**X_{12}**H**X_3**H** zinc finger (Klug and Schwabe 1995). However, there was no difference in binding affinities between either the ID or CTD tested against three different oligonucleotide sequences (data not shown), suggesting that Mcm10 does not recognize DNA in a sequence-specific manner. On the other hand, the two tandem cysteine-rich clusters in the CTD are remarkably similar in sequence to LIM domains and RING-finger motifs, which provide protein-binding interfaces important for a variety of cellular functions (reviewed in Borden 2000; Kadrmas and Beckerle 2004). It is noteworthy that the CTD zinc motif is immediately adjacent in the primary sequence to a putative winged helical (WH) bundle, which was predicted based on its similarity to that of the SCF ubiquitin ligase (Murzin, Brenner et al. 1995). The globular assembly formed from the RING protein Rbx1 and the WH of Cul1 in the SCF complex is an interaction integral to the cullin-RING ubiquitin ligase machinery (Zheng, Schulman et al. 2002; Petroski and Deshaies 2005). Thus, the zinc motif in xMcm10-CTD might stabilize the protein fold through a WH-RING interaction.

Structural and functional differences between vertebrate and yeast Mcm10

The lack of sequence conservation within the C-terminal region helps to reconcile differences in DNA binding activities of spMcm10 and xMcm10. The DNA binding affinity for spMcm10 N-terminal (1-303) and C-terminal (295-593) fragments, which are truncated between the putative OB-fold and zinc finger of the ID, was the same as that of the full-length protein (Fien, Cho et al. 2004). Full-length xMcm10, on the other hand, bound to DNA with 10-fold greater affinity than xMcm10-ID or –CTD alone (Table 2). Additionally, spMcm10 exhibited a 20-fold preference for ssDNA over dsDNA (Fien, Cho et al. 2004), whereas xMcm10 bound to ssDNA and dsDNA with the same affinity. Although the domain structure of yeast Mcm10 is unknown, these results are consistent with a second DNA binding domain in vertebrate xMcm10-CTD that is not present in the yeast proteins.

The sequence divergence and different DNA binding activities between vertebrate and yeast Mcm10 suggest that these proteins have evolved subtly different functions. An additional DNA binding domain may have evolved in response to the greater complexity of the genome and the lack of specific nucleotide sequences at origins of replication. Alternatively, the additional DNA and pol α binding domain and the lack of detectable primase activity in xMcm10 suggest that vertebrate Mcm10 evolved a means to recruit pol α-primase in lieu of itself priming DNA. Structural studies will be required to determine whether the ID and CTD are classical DNA binding domains, or if they form versatile structural scaffolds commonly observed in replication proteins (Shamoo, Friedman et al. 1995; Mizuno, Yamagishi et al. 1999; Lee, Chang et al. 2000; Bochkareva, Korolev et al. 2002).

Perspectives on Mcm10's role at the replication fork

Structural arrangement of Mcm10 domains together with their macromolecular interactions provides insight into Mcm10 function. Our results are consistent with the notion that Mcm10 recruits pol α to origins of replication (Fien, Cho et al. 2004; Ricke and Bielinsky 2004; Ricke and Bielinsky 2006). With each of two separate domains encompassing DNA and pol α binding activities, Mcm10 might mediate a hand-off mechanism between pol α and DNA. Domain rearrangement to facilitate a handoff between replication proteins and DNA has been proposed for SV40 T antigen-mediated RPA loading onto DNA (Jiang, Klimovich et al. 2006).

Evidence is provided here for NTD-mediated dimerization of vertebrate Mcm10 (Figure 8). Analytical ultracentrifugation clearly showed dimerization of the NTD with a K_d of ~3.1 μM. The full-length enzyme is predominantly monomeric at low concentration but also self-associates, and by analogy to NTD it is likely also a monomer-dimer system. We observed that the NTD of mammalian and yeast Mcm10 contains a predicted coiled-coil (Supplementary Figure A1), a highly asymmetric motif that would explain protein dimerization and the anomalously short gel filtration retention times of Mcm10 constructs containing the NTD. Indeed, frictional ratios calculated from the sedimentation data are indicative of a highly asymmetric protein. These data are consistent with glycerol gradient sedimentation results showing spMcm10 dimerization and suggesting an elongated scMcm10 structure (Lee, Seo et al. 2003), and are intriguing in light of the recent report that human Mcm10 forms a globular homohexameric assembly (Okorokov, Waugh et al. 2007).

NTD-mediated dimerization raises the interesting possibility that Mcm10 interacts with both leading and lagging strand polymerases at a replication fork. Direct physical interactions between Mcm10 and pol α have now been observed in scMcm10, spMcm10, and xMcm10 (Fien, Cho et al. 2004; Ricke and Bielinsky 2006), and genetic studies raise the possibility that Mcm10 also interacts with replicative polymerases δ and ε. The coiled-coil interaction would orient both subunits of the Mcm10 dimer in the same direction and consequently provide the polarity needed for the individual subunits to associate with co-directional leading and lagging strands.

The fact that xMcm10 did not preferentially bind to forked DNA substrates (Table 2) suggests that Mcm10 does not reside directly at the fork, but rather some distance behind the unwinding DNA. On the other hand, interactions between Mcm10 and Mcm2-7 subunits have been observed by yeast two-hybrid (Izumi, Yanagi et al. 2000). Our data suggest that Mcm10 travels with pol α by association with the N-terminal end of p180. This region is dispensable for polymerase activity of p180 (Mizuno, Yamagishi et al. 1999), suggesting that Mcm10 is capable of interacting with pol α during DNA synthesis. The p68 subunit of pol α has been reported to interact with SV40 T antigen, tethering pol α to the viral replication fork (Collins, Russo et al. 1993; Ott, Rehfuess et al. 2002), but p68 did not interact with xMcm10 (data not shown). In addition, we were unable to detect a direct interaction between xCdc45 and pol α, or between xMcm10 and xCdc45 (data not shown).

In summary, the structural studies begun here provide a framework for future studies to elucidate the spatial arrangement of vertebrate Mcm10 and its binding partners and to develop a model for the action of these proteins within the replisome.

Experimental Procedures

Cloning, Expression and Purification of xMcm10

The cDNAs encoding full-length xMcm10 (FL, 1-860) and deletion fragments 1-145, 1-230, 230-427, 427-860, and 596-860 were PCR amplified from a previously described plasmid encoding a GST-xMcm10 fusion (Wohlschlegel, Dhar et al. 2002). The FL-xMcm10 PCR product was ligated into a modified pMAL-c2x vector (New England Biolabs) to generate an MBP-xMcm10-His_6 fusion protein, and xMcm10 fragments were ligated into a modified pET-32a plasmid (Novagen) to generate N-terminal thioredoxin (Trx)-His_6-fusion proteins. Protein was overexpressed in *E. coli* BL21(DE3) cells in Luria-Bertani medium supplemented with 100 μg/ml ampicillin, 5 μM $ZnSO_4$, and 0.5 mM IPTG. Proteins were overexpressed at 22° C for 4 hr (FL) or at 16° C for 16 hr (fragments). The cells were resuspended in 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 10% glycerol and lysed under pressure (25,000 psi) using an EmulsiFlex-C3 homogenizer (Avestin, Inc.). FL-xMcm10 was purified by tandem nickel-NTA and amylose affinity chromatography, cleavage of the MBP-tag, and SP-sepharose cation exchange. Protein was concentrated and stored in FL-buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM dithiothreitol (DTT) and 5% glycerol). xMcm10 fragments were purified by nickel-NTA affinity chromatography, followed by cleavage of the Trx-His_6 tag. The cleaved proteins were further purified by cation exchange (fragments 230-427, 427-860, 596-860) or anion exchange (1-145 and 1-230) chromatography, followed by gel filtration on a Superdex™ 200 preparative column (GE Healthcare) that had been equilibrated with S-200 buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 5% glycerol,

4mM β-mercaptoethanol [β-ME]). Structural integrity of fragment proteins was verified by circular dichroism spectroscopy.

Limited Proteolysis and Fragment Identification

Proteolysis experiments were carried out in S-200 buffer, in which 5-20 μM xMcm10 was incubated with 1-200 ng protease (trypsin, α-chymotrypsin, elastase, or endoproteinase-Glu-C) in a 10-μl reaction at 37° C for 30 min. Proteolysis protection reactions contained 10 mM EDTA. Proteases were inactivated by adding 10 μl SDS–PAGE sample buffer (63 mM Tris-HCl pH 6.8, 700 mM β-ME, 2% w/v SDS, 0.03% w/v bromophenol blue, and 10% glycerol) and heating 5 min at 95° C. Proteolytic fragments were separated by SDS-PAGE and visualized by Coomassie blue staining.

Proteolytic fragments from MBP-xMcm10-His_6 were excised from the SDS-PAGE gel and subjected to in-gel digestion with Trypsin Gold (Promega) using standard procedures (Anumanthan, Halder et al. 2006). The resulting peptides were analyzed by matrix-assisted laser desorption/ionization, time-of-flight mass spectrometry (MALDI-TOF MS) and TOF/TOF tandem MS using a Voyager 4700 (Applied Biosciences, Framingham MA). Peptide ion masses (M+H) were accurate to within 20 ppm after internal calibration using either trypsin autolytic peptides or xMcm10-derived peptides confirmed by TOF/TOF MS.

Molecular masses of xMcm10 domains resulting from proteolysis of deletion mutants Δ1, Δ2, and Δ3 were obtained by MALDI-TOF mass-spectrometry of the proteolysis reactions prior to SDS-PAGE. Reactions were concentrated in 0.1% trifluoroacetic acid, mixed with 3 μl saturated sinapinic acid in 60:40 (vol/vol)

acetonitrile:1% trifluoroacetic acid/dH_2O, and 1 µl was deposited onto a gold 100-well plate. Mass spectra were acquired on a Perceptive Biosystems Voyager Elite TOF spectrometer equipped with a laser desorption ionization source and an extended-path ion reflector. Protein standards from Sigma (MSCAL1-1KT) were used for mass calibration. For N-terminal sequencing of xMcm10 domains, intact proteolytic fragment proteins were transferred from SDS gel to PVDF membrane, stained with Ponceau-S, extracted from the membrane and subjected to Edman degradation chemistry using an Applied Biosystems Model 492HT Protein/Peptide Sequencer equipped with an on-line PTH-amino acid analyzer.

Zinc Quantitation

Quantitative analysis of zinc bound to xMcm10 was performed using graphite furnace atomic absorption (GFAA) spectroscopy. Analyses were performed using a Perkin Elmer HGA SIMAA 6000 graphite furnace equipped with an Aanalyst 800 GFAA/FLAA spectrophotometer. xMcm10 domains were quantified by absorbance spectroscopy at 280 nm using extinction coefficients of 0.092 (NTD), 1.09 (ID), and 0.524 (CTD) $ml \cdot mg^{-1} \cdot cm^{-1}$.

Gel Filtration Chromatography and Analytical Ultracentrifugation

Size exclusion chromatography of FL-xMcm10 was performed on a Superose 6 column (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 5% glycerol, and 1 mM DTT. xMcm10 domains were eluted from an analytical Superdex^{TM} 200 column (GE Healthcare) equilibrated with S-200 buffer. 50 µl solutions of either

xMcm10 (~1-2 mg/ml) or molecular weight standards were eluted at 0.5 ml/min. The standard curve was generated from thyroglobulin (670 kD), aldolase (158 kD), albumin (67 kD), chicken ovalbumin (44 kD), equine myoglobin (17kD), and Vitamin B_{12} (1.4 kD).

Sedimentation velocity analysis was conducted at 20°C and 55,000 RPM using interference optics with a Beckman-Coulter XL-I analytical ultracentrifuge. Double sector synthetic boundary cells equipped with sapphire windows were used to match the sample and reference menisci. FL-xMcm10 was prepared in FL-buffer, and NTD and CTD were prepared in S-200 buffer. The data were initially analyzed using the program DCDT+ which computes the apparent sedimentation coefficient distribution function g(s*) using the time-derivative method (Stafford 1992; Philo 2000). For CTD, the molecular weight and sedimentation coefficient of the main component was obtained by global fitting of the data sets collected at multiple concentrations to a hybrid discrete-continuous model with Sedphat (Schuck 2003). For NTD, the data were fit to a monomer-dimer equilibrium model using the programs Sedanal (Stafford and Sherwood 2004) and Sedphat. Molecular masses, partial specific volumes and solvent densities were calculated using the SEDNTERP program (Laue 1992).

Fluorescence Anisotropy

DNA binding was measured by following an increase in fluorescence anisotropy as protein (MBP-xMcm10-His_6, NTD, ID, or CTD) was added to oligonucleotide, d(TGACTACTACATGGTTGCCTACCAT), containing a 6-carboxyfluorescein moiety at the 3′-end, either alone (ssDNA) or annealed to an excess of the complementary strand

(dsDNA). Forked DNA substrate tested against full-length Mcm10 was generated from two 50mer deoxyoligonucleotides, in which dC_{25} was added to the 3′-end of the sequence above and to the 5′-end of the complementary sequence. For Mcm10-ID and –CTD, forked and bubble DNA substrates were generated from the sequences d(**GGTAGGCACG**AACCATGTAGTAGTA) / d(AACCATGTAGTAGTA**CGTGCCTACC**) and d(**GGTAGGCACG**AACCATGTAGTAGTA**GGCAATCAGC**) / d(GCTGATTGCCAACCATGTAGTAGTA**CGTGCCTACC**), respectively, in which the bold-face denotes duplex regions. Protein was added over the concentration range of 0.05-50 μM to a solution containing 25 nM DNA in S-200 buffer. For EDTA titrations, the buffer was supplemented with 0.1, 1, 10, and 25 mM EDTA. Polarized fluorescence intensities using excitation and emission wavelengths of 495 and 515 nm, respectively, were measured for 30 s (1/s) and averaged. Anisotropy (r) was calculated using the equation $r = (I_{par}-I_{perp})/(I_{par}+2I_{perp})$, where I_{par} and I_{perp} are the observed fluorescence intensities recorded through polarizers oriented parallel and perpendicular, respectively, to the direction of vertically polarized light. Dissociation constants (K_d) were derived by fitting a simple two-state binding model to data from three experiments using Kaleidagraph 3.6 (Synergy Software, PA).

Mcm10-pol α Binding Assay

Recombinant DNA polymerase α-primase (pol-prim) was purified by immunoaffinity chromatography from extracts of Hi-5 insect cells co-infected with four recombinant baculoviruses as previously described (Voitenleitner, Fanning et al. 1997).

The p180 subunit was prepared identically except only one recombinant baculovirus was used for infection. P180N (aa 1-323) was amplified by PCR on a cDNA template pBR322-p180 and cloned into the BamHI/EcoRI sites of a pGEX-2T expression vector (GE Healthcare). GST fusion proteins were expressed and purified by glutathione-agarose affinity chromatography as described previously (Smith and Johnson 1988).

For the binding experiments, a total of 7 μg of purified pol-prim or p180 was incubated with SJK132-20 antibodies covalently coupled to Sepharose-4B beads (GE Healthcare), or 7 μg of purified p180N was incubated with glutathione-agarose beads (Sigma-Aldrich) in binding buffer (30 mM HEPES-KOH pH 7.8, 10 mM KCl, 7 mM $MgCl_2$) containing 2% nonfat dry milk for 1 hr at 4º with end-over-end rotation. Reactions contained either 5 or 15 μg Trx-His_6-xMcm10-domain proteins. The beads were washed once with binding buffer, three times with wash buffer (30 mM HEPES-KOH pH 7.8, 75 mM KCl, 7 mM $MgCl_2$ 0.25% inositol, 0.1% NP-40) and once with binding buffer (rotated for 10 min during each wash). The beads were resuspended in 30 μl of 2x SDS-PAGE loading buffer and heated at 100° C for 5 min. Half of each sample was analyzed by 10% SDS-PAGE and immunoblotting with monoclonal antibody 2CT25, specific for the p180 subunit of pol-prim, rabbit anti-GST (Invitrogen) for p180N, and H-15 anti-His (Santa Cruz Biotechnology) for xMcm10 domains. Trx-only control experiments were performed to confirm that pol α, p180, and p180N did not interact with the Trx affinity tag.

DNA Primase Assay

Oligoribonucleotide synthesis activity was measured as previously described for spMcm10 (Fien and Hurwitz 2006). Briefly, 2-8 pmol purified xMcm10 or 0.6-2.4 pmol purified pol-prim were incubated at 37° C for 40 min with 1.0 μM dT_{50}, 25 μCi [α-^{32}P]ATP and 0.1 mM ATP in a 10 μM reaction containing 40 mM Tris-HCl (pH 7.4), 10 mM magnesium acetate, 1 mM DTT, and 100 μg/ml BSA. Reactions were treated with 1 U calf intestine phosphatase at 37° C for 40 min. After addition of 3 μl sequencing gel running buffer (98% formamide, 10 mM EDTA pH 8.0, 0.1% xylene cyanol, 0.1% bromophenol blue), samples were heated to 98° C for 5 min and separated on a 25% polyacrylamide/7 M urea gel. RNA was visualized by autoradiography.

Footnotes

The authors wish to thank Tim Bowles, Jami O'Quin and Laura Mizoue for technical assistance, Hassane McHaourab, Albert Beth, and Charles Cobb for access to fluorometers, and Walter Chazin for critical reading of the manuscript. We thank the Vanderbilt Academic Venture Capital Fund for support of the Proteomics Laboratory. Peptide sequencing was provided by the Protein Analysis Lab at Wake Forest University, and atomic zinc analysis was performed by Galbraith Laboratories. This work was funded by the National Institutes of Health (GM080570 to B.F.E. and GM52948 to E.F.) and the Vanderbilt Discovery Grant Program. Additional support for facilities came from the Vanderbilt Center in Molecular Toxicology (P30 ES000267). E.M.W. was supported from the Molecular Biophysics Training Grant (T32 GM08320).

CHAPTER III

STRUCTURAL BASIS FOR DNA BINDING BY REPLICATION INITIATOR MCM10[*]

Abstract

Mcm10 is an essential eukaryotic DNA replication protein required for assembly and progression of the replication fork. Mcm10 has been shown to physically interact with single (ss) and double-stranded (ds) DNA, DNA polymerase α, and PCNA. The crystal structure of the conserved internal domain of *Xenopus laevis* Mcm10 (Mcm10-ID) presented here reveals a novel architecture that helps to explain Mcm10's ability to bind various DNA substrates. Mcm10-ID is composed of an OB-fold followed in tandem by a variant and highly basic zinc finger motif, which together form a unique DNA binding surface. NMR chemical shift perturbation and mutational studies of Mcm10-ID DNA binding *in vitro* reveal that the protein uses this contact surface to engage both ssDNA and dsDNA. Corresponding mutations in *Saccharomyces cerevisiae* Mcm10 result in increased sensitivity to replication stress, demonstrating the functional importance of DNA binding by this region of Mcm10 to replication.

[*] The work presented in this chapter was published in Warren, E. M., Vaithiyalingam, S.R., 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): 1892-1901.

Introduction

Eukaryotic DNA replication is carried out by large multiprotein machines that coordinate DNA unwinding and synthesis at the replication fork. Assembly of the replisome is highly regulated, and proceeds in stages through a series of protein complexes that recognize and denature origin DNA (Figure 14) (reviewed in Bell and Dutta 2002). Pre-replicative complexes (pre-RCs) composed of the origin recognition complex (ORC), Cdc6, Cdt1, and the minichromosome maintenance (Mcm) 2-7 helicase are assembled in G_1 (reviewed in Blow and Dutta 2005). Initiation of replication begins at the G_1/S transition and involves conversion of pre-RCs into the functional replisome. Mcm10 loads onto chromatin after pre-RC assembly and is required for Cdc45 and replication protein A (RPA) recruitment (Wohlschlegel, Dhar et al. 2002; Ricke and Bielinsky 2004). Cdc45 and GINS (Takayama, Kamimura et al. 2003) associate with Mcm2-7 to form a helicase complex (Pacek and Walter 2004; Gambus, Jones et al. 2006; Moyer, Lewis et al. 2006; Pacek, Tutter et al. 2006). Phosphorylation of Mcm2-7 and several other replication factors by cyclin- and Dbf4-dependent kinases (CDK, DDK) stimulate origin unwinding, which is signaled by recruitment of RPA to the origin (Lei, Kawasaki et al. 1997; Tanaka and Nasmyth 1998; Zou and Stillman 2000; Tanaka, Umemori et al. 2007; Zegerman and Diffley 2007). Mcm10, Cdc45, and RPA facilitate subsequent loading of DNA polymerase α-primase (pol α) (Mimura and Takisawa 1998; Walter and Newport 2000; Ricke and Bielinsky 2004; Yang, Gregan et al. 2005). Replicative DNA polymerases δ and ε and associated PCNA are recruited to form the intact replisome (reviewed in Garg and Burgers 2005).

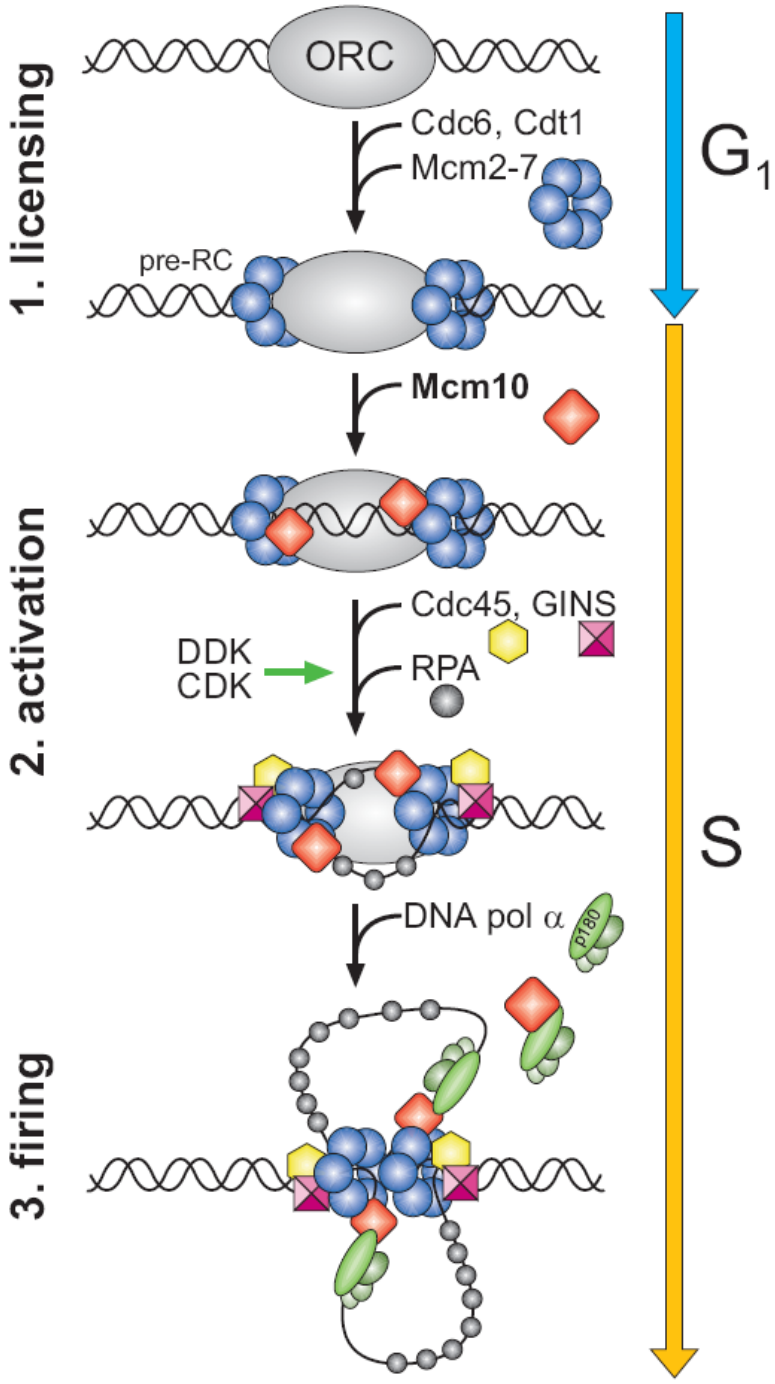


Figure 14. Initiation of eukaryotic DNA replication. The schematic shows some of the key steps necessary for DNA unwinding and replication fork assembly at a eukaryotic origin of replication. Details are described in the text.

Mcm10 is essential for replisome assembly and fork progression. Originally identified by independent yeast screens (Nasmyth and Nurse 1981; Dumas, Lussky et al. 1982; Maine, Sinha et al. 1984; Merchant, Kawasaki et al. 1997), Mcm10 is associated with chromatin throughout S-phase and is a component of active replication complexes in *Xenopus* and budding yeast (Ricke and Bielinsky 2004; Gambus, Jones et al. 2006; Pacek, Tutter et al. 2006). A number of genetic interactions have been observed between Mcm10 and proteins found in the pre-RC and at the replication fork, including Mcm2-7, DNA pol δ and ε, ORC, and Dpb11 (Merchant, Kawasaki et al. 1997; Homesley, Lei et al. 2000; Izumi, Yanagi et al. 2000; Kawasaki, Hiraga et al. 2000). *In vitro*, physical interactions between Mcm10 and initiation factors ORC, Mcm2-7, Cdc45, And-1, and Cdc7/Dbf4 have been observed (Homesley, Lei et al. 2000; Kawasaki, Hiraga et al. 2000; Christensen and Tye 2003; Lee, Seo et al. 2003; Zhu, Ukomadu et al. 2007), and Mcm10 has been shown to stimulate DDK phosphorylation of Mcm2-7 (Lee, Seo et al. 2003). Interestingly, Mcm10 has been shown to be diubiquitinated during G1 and S phase in budding yeast (Das-Bradoo, Ricke et al. 2006). Diubiquitination of Mcm10 is a prerequisite to interact with PCNA. Importantly, the binding between Mcm10 and PCNA serves an essential function during DNA replication. Thus, Mcm10 plays key roles in both initiation and elongation.

A number of reports have demonstrated an interaction between Mcm10 and pol α. *In vivo*, Mcm10 interacts with and stabilizes the catalytic (polymerase) subunit of pol α in human and *Saccharomyces cerevisiae* (sc) cells (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007), and *Schizosaccharomyces pombe* Mcm10 (spMcm10) affects the association of pol α with chromatin (Gregan, Lindner et

al. 2003; Yang, Gregan et al. 2005). *In vitro*, spMcm10 stimulates the polymerase activity of pol α (Fien, Cho et al. 2004) and has been shown to contain primase activity (Fien and Hurwitz 2006), although *Xenopus laevis* Mcm10 (xMcm10) does not synthesize RNA primers under identical conditions (Robertson, Warren et al. 2008). The interaction between xMcm10 and pol α has been mapped to the N-terminal 323 residues of the catalytic p180 subunit (Robertson, Warren et al. 2008). The Mcm10-pol α interaction has led to the suggestion that Mcm10 helps to recruit the polymerase to the replisome and may regulate its activity.

In addition to its interactions with the replisome, Mcm10 is a DNA binding protein. Mcm10 from fission yeast, frogs, and humans have been shown to bind to both single (ss)- and double-stranded (ds) DNA (Fien, Cho et al. 2004; Okorokov, Waugh et al. 2007; Robertson, Warren et al. 2008). Alignment of known Mcm10 protein sequences from yeast to human shows a high degree of conservation within a 200-amino acid internal region (Figure 15A). We have previously identified this conserved internal domain (Mcm10-ID) to be one of two domains that bind to both DNA and pol α *in vitro* (Robertson, Warren et al. 2008). In addition to a conserved Cys_3His-type zinc finger (Izumi, Yanagi et al. 2000; Cook, Kung et al. 2003), Mcm10-ID has been predicted to contain an oligonucleotide/oligosaccharide (OB)-fold (Ricke and Bielinsky 2006), both of which are classic DNA binding motifs. Recently, human Mcm10 was reported to form ring-shaped hexameric assemblies reminiscent of DNA helicases (Okorokov, Waugh et al. 2007).

It is unclear how Mcm10's interactions with protein and DNA contribute to its association with chromatin during initiation and elongation. Remarkably, the mutations

discovered from yeast genetic screens are all located within the Mcm10-ID (Nasmyth and Nurse 1981; Dumas, Lussky et al. 1982; Maine, Sinha et al. 1984; Solomon, Wright et al. 1992; Grallert and Nurse 1997; Merchant, Kawasaki et al. 1997). Additionally, several mutations within the ID have been shown to disrupt scMcm10 association with PCNA and pol α (Das-Bradoo, Ricke et al. 2006; Ricke and Bielinsky 2006). In *S. cerevisiae*, the primary interaction site for pol α is confined to a hydrophobic stretch, the heat-shock protein (Hsp)10-like motif (Ricke and Bielinsky 2006), which lies adjacent to Mcm10's PIP box (Das-Bradoo, Ricke et al. 2006) within Mcm10-ID. It is worthwhile to note that Mcm10 can stabilize pol α through its Hsp10-like domain even if the putative DNA binding OB-fold and zinc finger motifs are structurally disrupted (Ricke and Bielinsky 2006). Collectively, these mutations demonstrate the importance of the ID in Mcm10 function and motivate structural analysis of this domain.

Presented here is a high resolution structure-function analysis of the conserved internal domain of xMcm10. The crystal structure reveals that the central domain of Mcm10 is composed of a unique arrangement of two classic DNA binding motifs. NMR chemical shift perturbation and mutational analyses demonstrate a common binding surface for ssDNA and dsDNA, and provide a model for how Mcm10 engages DNA. Mutation of residues lining the DNA binding surface in scMcm10 increases the sensitivity of yeast that have been subjected to hydroxyurea-induced replication stress. These data establish that DNA binding by the core domain of Mcm10 is critical for maintenance of the replication fork. Additionally, mapping scMcm10 mutations known to disrupt PCNA, pol α, and DNA interactions onto the crystal structure provides

important insight into how xMcm10-ID may coordinate protein and DNA binding within the replisome.

Results

We have previously identified the domain architecture of xMcm10 (Robertson, Warren et al. 2008). Limited proteolytic digestion of the full-length protein produced N-terminal (NTD, aa 1-145), internal (ID, 230-417), and C-terminal (CTD, 596-860) domains that correspond to the sequence conservation among Mcm10 proteins (Figure 15A). The NTD encompasses an oligomerization function, while the ID and CTD both bind to ssDNA, dsDNA, and pol α (Robertson, Warren et al. 2008). Mcm10-ID is the only region of the 860-residue protein that shows significant homology across all species from vertebrates to yeast (Figure 15A). The amino acid sequence of xMcm10-ID is 73% and 24% identical to human and scMcm10, respectively. Extensive homology suggests that Mcm10-ID contains an essential function and prompted a rigorous structure-function analysis of this core domain.

The unique structure of the conserved domain of Mcm10

The crystal structure of Mcm10-ID from *Xenopus laevis* was determined to 2.3 Å resolution (Figure 15B). Experimental phases were obtained from a multiwavelength anomalous dispersion (MAD) experiment using a single gold derivative crystal (Table 3). The atomic model, which consists of three Mcm10-ID molecules in the asymmetric unit, was built into 3.0 Å Au-MAD electron density and refined against 2.3 Å native diffraction data (Table 3) to a crystallographic residual of 0.202 (R_{free} = 0.247). The

accuracy of the structure is reflected in part by 91.5% and 8.3% of the 515 total residues residing within the favored and allowed regions of the Ramachandran plot, respectively.

Mcm10-ID forms a globular domain consisting of an OB-fold (β1-β5.2) flanked by an α-helical/random coil region (αA-αB) at the amino-terminus and a zinc finger motif (βC-αE) at the carboxy-terminal end (Figure 15C,D). The α-helical region is packed against the back side of the OB-fold to form an essentially flat molecular surface. The zinc finger protrudes to one side of the structure and makes extensive electrostatic and van der Waals contacts with the OB-fold L23 loop and the α-helical/coil region (Figure 15C). The three crystallographically independent Mcm10-ID molecules in the asymmetric unit superimpose with an r.m.s. deviation of 0.7 Å for all atoms with the zinc finger in the same relative position with respect to the OB-fold in each protomer. Thus, there is no evidence to suggest free movement between the OB-fold and zinc finger, despite the cluster of invariant glycine residues Gly339, Gly373, and Gly379 at the OB-fold/zinc finger junction (Figure 15D). In summary, Mcm10-ID forms a single structural domain with the OB-fold and zinc finger motifs in an orientation that makes them both accessible to DNA.

Table 3. Data collection, phasing and refinement statistics for xMcm10-ID

	Native	$K_2Au(CN)_2$		
Data collection[a]				
Space group	$P2_1$		$P2_1$	
Cell dimensions				
a, *b*, *c* (Å)	54.6, 94.4, 69.8		54.5, 94.4, 69.8	
α, β, γ (°)	90, 112.8, 90		90, 112.6, 90	
		Peak	*Inflection*	*Remote*
Wavelength	1.000	1.0388	1.0370	1.0311
Resolution (Å)	50-2.3 (2.38-2.3)	50-2.8 (2.9-2.8)	50-2.8 (2.9-2.8)	50-2.8 (2.9-2.8)
R_{sym} or R_{merge}	0.074 (0.425)	0.034 (0.098)	0.034 (0.131)	0.038 (0.129)
$I / \sigma I$	27.7 (2.3)	25.7 (7.9)	19.3 (4.4)	24.0 (6.5)
Completeness (%)	99.6 (97.0)	99.6 (100.0)	79.2 (80.7)	99.7 (99.8)
Redundancy	6.7 (3.9)	3.7 (3.6)	2.3 (2.1)	3.7 (3.6)
Refinement				
Resolution (Å)	50-2.3			
No. reflections	27,448			
R_{work} / R_{free}	0.202 / 0.247			
No. atoms				
Protein	4128			
Ligand/ion	3			
Water	149			
B-factors				
Protein	56.3			
Ligand/ion	47.2			
Water	54.3			
R.m.s deviations				
Bond lengths (Å)	0.014			
Bond angles (°)	1.572			

(a) Values in parentheses refer to the highest-resolution shell.

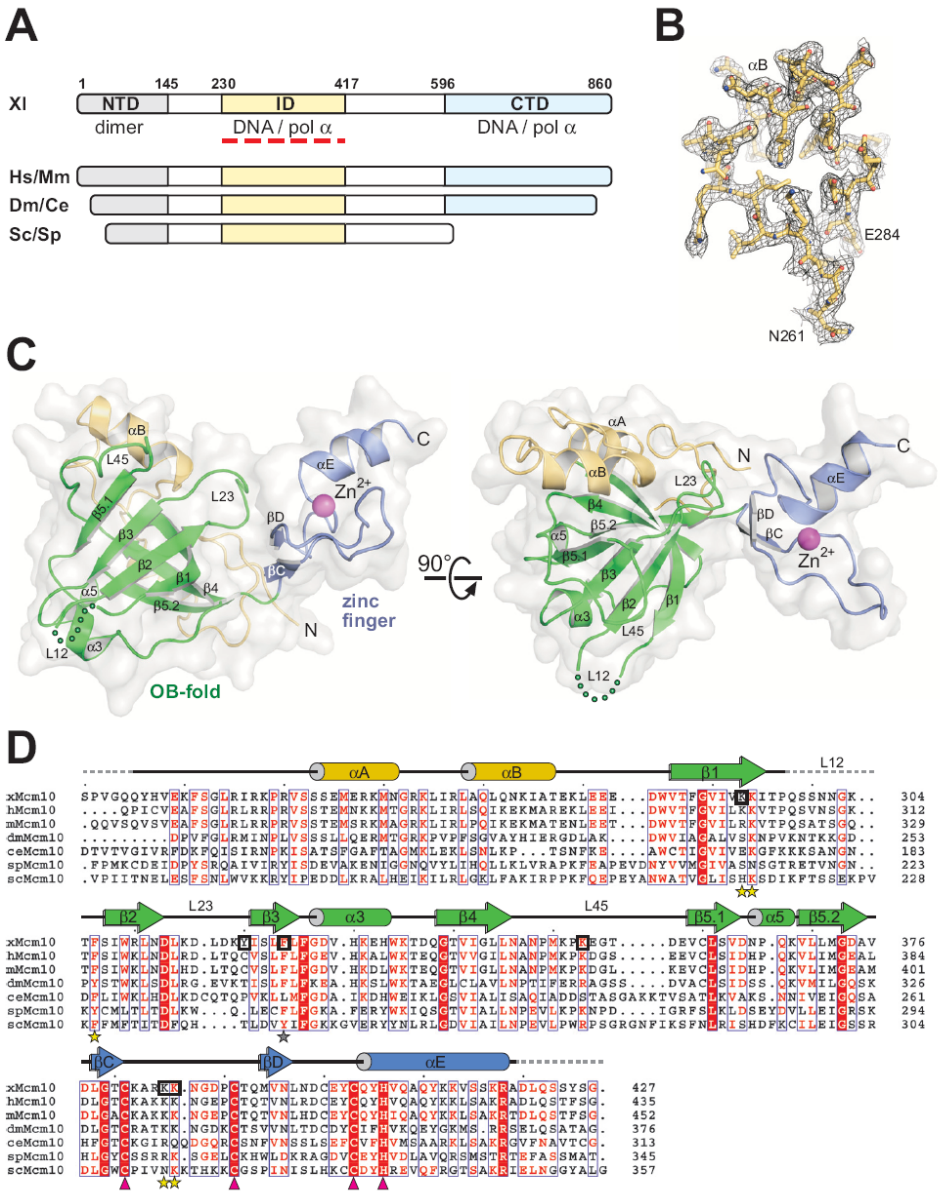


Figure 15. Structure of the conserved central domain of Mcm10. (**A**), Mcm10 domain architecture. The three structural domains identified from *Xenopus laevis* (Xl) Mcm10 are shown as colored bars and aligned with homologous regions of Mcm10 from *Homo sapiens* (Hs), *Mus musculis* (Mm), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce), *Saccharomyces cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp). (**B**), A representative section of the refined crystallographic model superimposed onto composite omit electron density contoured at 1σ. Stereoviews of electron density maps can be found as Supplementary Figure B1. (**C**), The crystal structure of *Xenopus laevis* Mcm10-ID (residues 230-427) is shown as a ribbon diagram with a white molecular surface. Two orthogonal views show the relative orientation of the OB-fold (green), zinc finger (blue ribbon, magenta Zn^{2+} sphere), and N-terminal α-helical/coil region (gold). (**D**), Sequence alignment of Mcm10-ID with schematic secondary structural elements colored as in panel C. xMcm10 residues identified by mutagenesis to interact with DNA *in vitro* are boxed. scMcm10 mutations that result in increased sensitivity of yeast to hydroxyurea are marked with yellow stars, and the grey star denotes a lethal mutation. Conserved Zn^{2+}-coordinating residues are marked with magenta triangles.

The combination of OB-fold and zinc finger domains in a single structural domain is intriguing. The OB-fold is a canonical ssDNA/RNA binding motif that typically binds nucleic acids in the concave cleft formed by strands β2 and β3. The loops between β1/β2 (L12), β4/β5.1 (L45) and flanking helix α3 (L3α and Lα4) stabilize the protein-ssDNA interaction (reviewed in Theobald, Mitton-Fry et al. 2003). Zinc fingers, on the other hand, are dsDNA recognition motifs in which the α-helix inserts itself into the major groove surface of the DNA in order to maximize sequence-specific hydrogen bonding interactions (Pavletich and Pabo 1991; Krishna, Majumdar et al. 2003). While the OB-fold and zinc finger are considered to be classic ssDNA and sequence-specific dsDNA binding motifs, respectively, some OB-folds also bind to dsDNA (Pascal, O'Brien et al. 2004; Nair, Nandakumar et al. 2007), and variations of the zinc finger allow for sequence-independent DNA binding (Finerty and Bass 1999).

The two DNA binding motifs in Mcm10 form a unique molecular surface based on several key structural features. Firstly, the spatial arrangement of OB-fold and zinc finger in Mcm10-ID is different from other DNA processing proteins that contain both structural motifs (Figure 16A). Whereas zinc motifs of RPA70, T4 gp32, NAD^+-dependent DNA ligase, and archaeal MCM helicase reside within or adjacent to the L12 loop and have been suggested to play a structural, rather than ligand-binding role (Shamoo, Friedman et al. 1995; Bochkarev, Pfuetzner et al. 1997; Lee, Chang et al. 2000; Fletcher, Bishop et al. 2003), the zinc finger of Mcm10 is on the opposite (L23) face of the OB-fold. Secondly, Mcm10's variant Cys_3His-type zinc finger contains an extended loop between βC and βD that is oriented perpendicular to the α-helix as a result of Zn^{2+} coordination at the N-terminal end of the helix (Figure 16B). Consequently, the zinc loop

is positioned immediately adjacent to the putative DNA binding cleft of the OB-fold (Figure 15C).

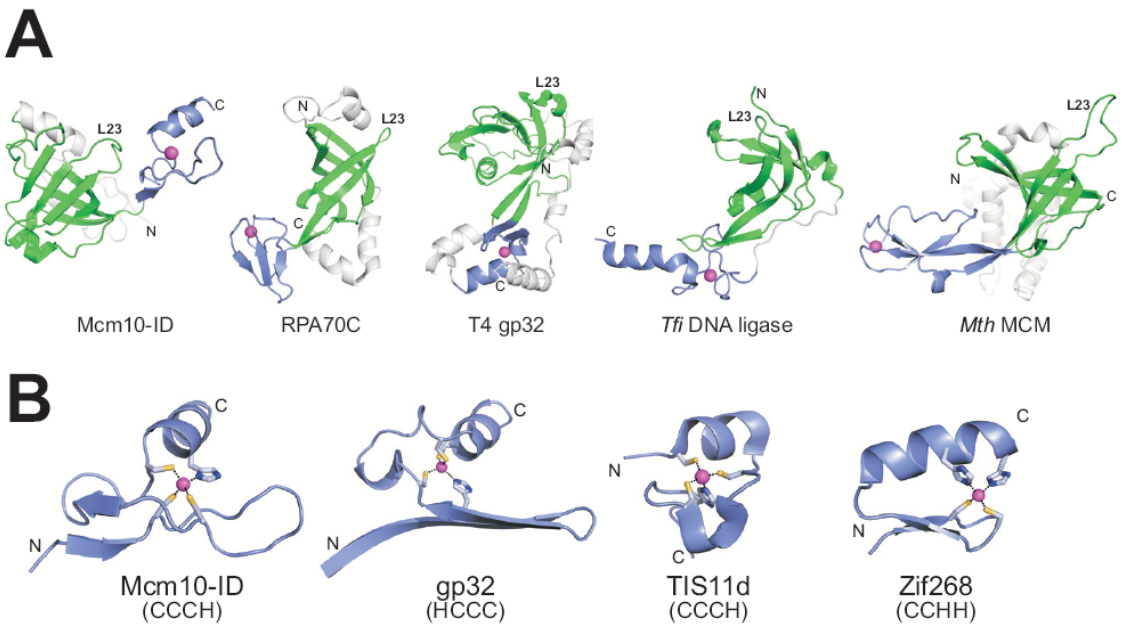


Figure 16. Comparison of OB-fold and zinc finger motifs in DNA binding proteins. (**A**), Structures of five proteins that contain an OB-fold (green) and zinc motif (blue) in the same domain. The structures are oriented with respect to the OB-fold β1 strand, with the L23 loops labeled to illustrate the relative positions of the zinc-binding motifs. (**B**), Variant zinc fingers from Mcm10-ID (**C**X_9**C**X_{11}**C**X_2**H**), T7 gp32 (**H**X_{12}**C**X_9**C**X_2**C**), and TS11d (**C**X_8**C**X_5**C**X_3**H**), and the archetypical zinc finger from Zif268 (**C**X_2**C**X_{12}**H**X_3**H**) are oriented with the long axis of the beta sheet running horizontally across the plane of the page. These structures suggest that the zinc finger fold depends more on the number of intervening residues between cysteines and histidines than on the order of the coordinating residues themselves. PDB IDs: Mcm10, 2q0w (this work); RPA70C, 1jmc (Bochkarev, Pfuetzner et al. 1997); gp32, 1gpc (Shamoo, Friedman et al. 1995); DNA ligase, 1v9p (Lee, Chang et al. 2000); MCM, 1ltl (Fletcher, Bishop et al. 2003); TIS11d, 1rgo, (Hudson, Martinez-Yamout et al. 2004); Zif268, 1zaa, (Pavletich and Pabo 1991).

A novel DNA binding platform

The differences in the structure of the zinc finger and its proximity to the OB-fold raised the possibility that Mcm10-ID engages DNA in a manner distinct from other OB-fold/zinc finger proteins. Indeed, in contrast to RPA's marked preference for ssDNA, xMcm10-ID binds ssDNA with only 2-3-fold higher affinity than dsDNA (Lao, Lee et al.

1999; Robertson, Warren et al. 2008). In order to expand our understanding of nucleic acid binding by Mcm10-ID, ssDNA and dsDNA binding was investigated using heteronuclear NMR spectroscopy. The strategy involved monitoring perturbations in NMR chemical shifts as DNA was titrated into ^{15}N-enriched Mcm10-ID. The chemical shift is an exquisitely sensitive parameter that responds to both binding and conformational changes. This approach has been used extensively to map the location of ssDNA and dsDNA binding sites on protein structures (e.g., Bhattacharya, Botuyan et al. 2002).

The first step in the NMR analysis was determination of the sequence-specific resonance assignments (Supplementary Figure B2A). These were obtained using standard multi-dimensional triple resonance experiments performed on ^{13}C,^{15}N-enriched Mcm10-ID. The addition of ssDNA to Mcm10-ID resulted in a shift of a significant number but not all of the peaks in the 2D ^{15}N-^{1}H HSQC spectrum (Supplementary Figure B2B), which is consistent with a combination of effects from DNA binding and small conformational changes in the protein. Over the course of the titration, the peaks shifted continuously (fast exchange) with changes saturated at a 1:3 protein:DNA ratio (data not shown). These observations are consistent with the 3 μM binding affinity measured by fluorescence anisotropy (Robertson, Warren et al. 2008). The corresponding titration with dsDNA resulted in perturbations of the chemical shifts of largely the same residues as for ssDNA (Figure 17A). The similarity of the chemical shift perturbations from titration with ss- and dsDNA was surprising, as it was anticipated that dsDNA would bind to the zinc finger and ssDNA to the OB-fold. These observations strongly imply that the binding sites for ssDNA and dsDNA are similar.

In order to determine the DNA binding site of Mcm10-ID, the residues exhibiting the most significant NMR chemical shift perturbations were mapped onto the crystal structure (Figure 17B). The largest shifts were observed for residues lining the β-barrel of the OB-fold and the extended loop of the zinc finger. Very few perturbations were observed on the opposite face of the protein, demonstrating that DNA primarily contacts the common OB-fold/zinc loop surface. This suggested that a model could be generated for the DNA bound state based on combination of structural homology and the NMR data.

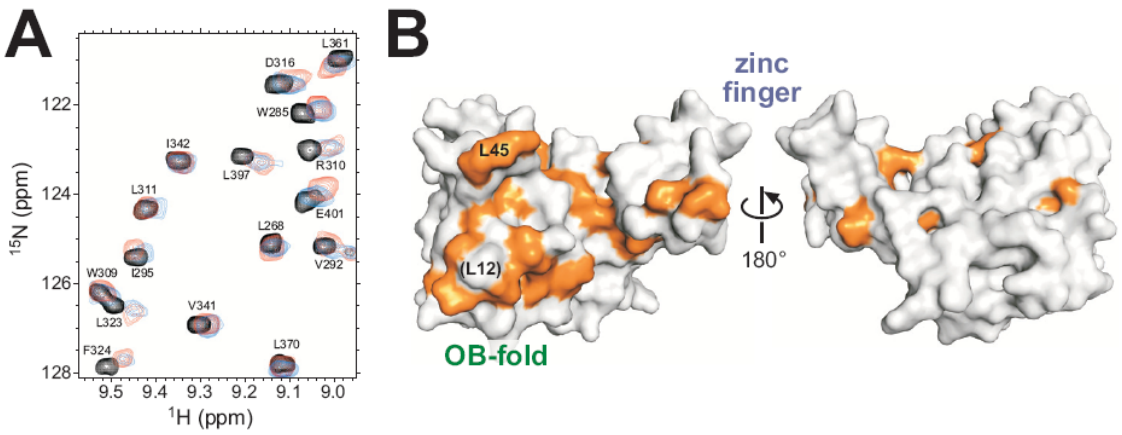
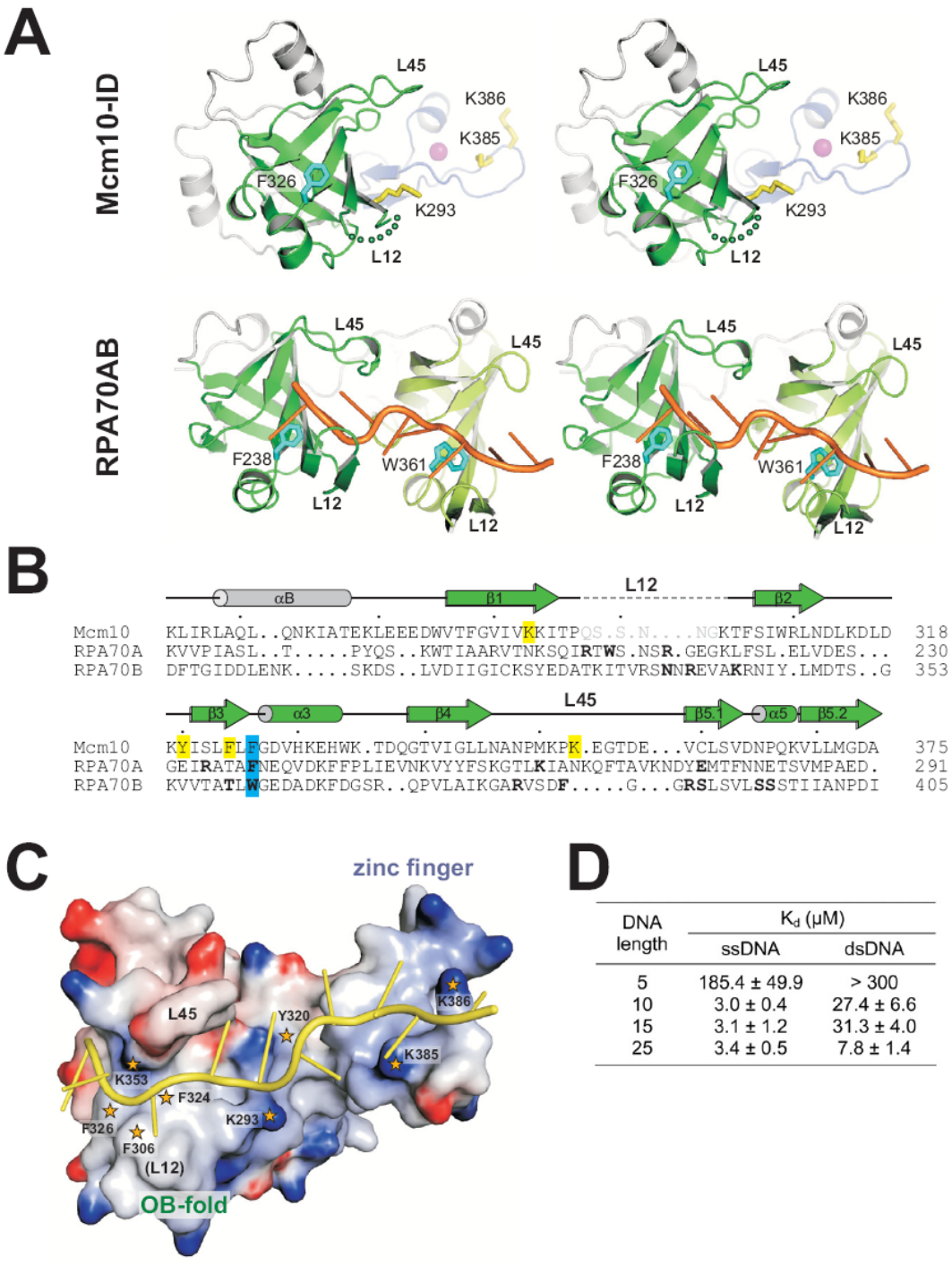


Figure 17. Mapping the Mcm10 DNA binding site. (**A**), NMR chemical shift perturbations in response to ssDNA and dsDNA binding to Mcm10. An overlay of a region of ^{15}N-^{1}H HSQC spectra of ^{15}N-enriched Mcm10 in the absence (black) and presence of ssDNA (red) and dsDNA (blue) is shown. The spectra of the complexes were acquired at protein:DNA ratios of 1:0.4 (ssDNA) and 1:0.8 (dsDNA). Peak assignments are labeled. (**B**), Surface representation of Mcm10-ID with all assigned residues showing significant chemical shift perturbation (>0.057 ppm) colored orange. Two orientations rotated 180° with respect to one another show that perturbations occur almost exclusively on one surface of the protein.

To this end, a comparative structure search using the DALI server (Holm and Sander 1993) was performed. This analysis revealed that the OB-fold in Mcm10-ID is most similar to those of the high affinity ssDNA-binding domains from the 70-kD

subunit of human RPA (RPA70AB). Consequently, the crystal structure of the RPA70AB/ssDNA complex was used as a basis to model ssDNA binding to Mcm10 (Figure 18). Structures are not yet available for complexes of OB-fold proteins with dsDNA, so modeling was restricted to the ssDNA complex. RPA70AB is composed of tandem OB-folds oriented with their ssDNA binding surfaces side-by-side, which allows ssDNA to traverse both binding pockets in a linear fashion (Figure 18A). Superposition of the Mcm10-ID and RPA70A OB-folds places the zinc finger in the same location as the RPA70B OB-fold (Figure 18A). Thus, the OB-fold and extended zinc loop in Mcm10-ID forms a molecular surface analogous to the DNA binding platform of the two RPA70AB OB-folds.

Using the NMR chemical shift perturbation data as restraints, DNA from the RPA70AB/ssDNA complex was docked onto the Mcm10-ID structure and energy minimized. The resulting model shows that a minimum of ~8-10 nucleotides are needed to span the OB-fold/zinc loop surface (Figure 18C and B4). This correlates well with the length dependence of DNA binding to Mcm10-ID determined by fluorescence anisotropy (Supplementary Figure B3), which showed that 10-nucleotide oligomers were the shortest DNA that supported high affinity binding for both ss- and ds-DNA (Figure 18D).



DNA length	K_d (μM)	
	ssDNA	dsDNA
5	185.4 ± 49.9	> 300
10	3.0 ± 0.4	27.4 ± 6.6
15	3.1 ± 1.2	31.3 ± 4.0
25	3.4 ± 0.5	7.8 ± 1.4

Figure 18. A model for DNA binding to Mcm10. (**A**), Stereodiagram of crystal structures of Mcm10-ID and an RPA70AB/ssDNA complex. OB-folds are colored green, zinc finger blue with magenta Zn^{2+}, and ssDNA orange. Residues important for DNA binding to Mcm10 and RPA are rendered as yellow and cyan sticks, respectively. (**B**), Structure-based sequence alignment of OB-folds from Mcm10, RPA70A, and RPA70B. Residues shown by mutagenesis to affect *in vitro* DNA binding in Mcm10-ID are highlighted yellow, and conserved aromatic residues contacting ssDNA in RPA are highlighted cyan. DNA-binding residues identified from the RPA crystal structure are in bold-face, and disordered residues in the Mcm10-ID crystal structure are grey.

Figure 18, continued. (**C**), Structural model of DNA bound to Mcm10. The ssDNA (yellow sticks) was docked onto the protein and refined using restraints from NMR chemical shift perturbation data. Mcm10-ID is shown as an electrostatic potential surface (blue, positive; red, negative). Residues implicated in DNA binding are highlighted with orange stars, and the positions of L12 and L45 loops known to contact ssDNA in other OB-folds are labeled. (**D**), Length dependence of DNA binding to Mcm10-ID. The dissociation constants (K_d) for Mcm10-ID and four different lengths of ss- and dsDNA were determined *in vitro* using fluorescence anisotropy (Supplementary Figure B3).

The model of the Mcm10-ID/ssDNA structure was used as a guide for mutational studies of Mcm10's DNA binding activities. The assessment of mutants was based on DNA binding affinities for 25mer ssDNA and dsDNA measured by the fluorescence anisotropy assay (Table 4). A significant difference between Mcm10's putative DNA binding surface and that of RPA70AB is the cluster of basic residues (Lys293, Lys385, Lys386) on the zinc finger loop and in the cleft formed between the two motifs (Figure 18). Mutations in this electropositive region had a marked effect on Mcm10 binding to DNA. Most strikingly, a Lys385Glu/Lys386Glu double mutant on the extended zinc loop exhibited a 10- and 5-fold reduction in ssDNA and dsDNA binding affinity, respectively (Table 4). This interaction is likely electrostatic, given that substitution of Lys385 and Lys386 to alanines did not affect DNA binding. Lysine → glutamate substitutions along the zinc finger helix (Lys412Glu/Lys413Glu and Lys417Glu/Arg418Glu), on the other hand, had no affect on DNA binding (Table 4). These data are consistent with the observation that dsDNA binding to Mcm10 is sequence independent (E.W. and B.F.E., unpublished data), and thus is not likely mediated by a zinc finger recognition helix binding in the major groove. Rather, an interaction between DNA and the flexible zinc loop helps to explain the observation that chelation of Zn^{2+} by EDTA (Robertson, Warren et al. 2008) and 1,10-phenanthroline (E.W. and B.F.E., unpublished data) does not affect ssDNA binding. In contrast to the

zinc loop, substitution of Lys293 with alanine reduced the affinity for ssDNA by 5-fold with respect to wild-type Mcm10-ID, and had only a modest effect on dsDNA binding (Table 4). Tyr320, also located in the OB-fold/zinc cleft, had a marginal but significant effect on DNA binding (Table 4).

On the OB-fold side, a cluster of three phenylalanine side chains (Phe306, Phe324, Phe326) and Lys353 are poised to interact with ssDNA in our model. Indeed, Phe324Ala on strand β3 and Lys353 in the L45 loop had a modest effect on DNA binding (Table 4). Substitution of any residue within the L12 loop, including Phe306, resulted in insoluble protein, which precluded analysis of L12 in our DNA binding assay. In RPA70AB, both OB-folds clamp the ssDNA between loops L12 and L45, and aromatic residues Phe238 (RPA70A) and Trp361 (RPA70B) at the C-terminus of β3 form DNA base stacking interactions (Bochkarev, Pfuetzner et al. 1997). Phe326 in xMcm10 is invariant among Mcm10 orthologs and superimposes with RPA Phe238 and Trp361 (Figure 18A,B). Surprisingly, substitution of Phe326 with alanine did not affect DNA binding to Mcm10-ID (Table 4). However, a single Phe238Ala mutation in RPA70A was also reported as not having a measurable effect on ssDNA binding, despite the observation of a direct contact to ssDNA in the crystal structure (Walther, Gomes et al. 1999). Thus, it is not possible to draw specific conclusions from the mutational data alone. The data do, however, reflect the redundancy in protein-DNA contacts along the extended binding site in Mcm10-ID. That is, the extended contacts with DNA preclude any one mutation from having a large effect on binding due to the significant portion of the DNA that can engage the protein outside of the area affected by the substitution. Taken together, the NMR and mutational data demonstrate that DNA spans the

hydrophobic cleft of the OB-fold and the extended, positively-charged loop of the variant zinc finger (Figure 18C and B4).

Table 4. DNA binding activity of Mcm10-ID mutants

Mutant	ssDNA		dsDNA	
	K_d (μM)	Relative binding	K_d (μM)	Relative binding
WT	3.4 ± 0.5	(1.00)	7.8 ± 1.4	(1.00)
K293A	17.8 ± 4.0	0.19	15.0 ± 1.2	0.52
Y320A	4.8 ± 0.5	0.71	15.5 ± 1.8	0.51
F324A	5.5 ± 1.9	0.61	19.5 ± 3.3	0.40
F326A	3.1 ± 0.7	1.08	5.5 ± 1.6	1.42
K353A	6.6 ± 1.3	0.51	55.8 ± 14.0	0.14
K385A / K386A	1.5 ± 0.6	2.25	9.6 ± 1.3	0.82
K386E	27.3 ± 2.1	0.12	34.0 ± 5.5	0.23
K385E / K386E	40.2 ± 8.6	0.08	45.2 ± 5.0	0.17
K412E / K413E	1.7 ± 0.3	1.94	7.6 ± 0.5	1.03
K417E / R418E	2.4 ± 0.4	1.45	4.3 ± 0.1	1.84

Functional relevance of DNA binding to Mcm10

In order to establish that the residues affecting DNA binding of xMcm10 *in vitro* have a role during DNA replication *in vivo*, we introduced the corresponding mutations into the endogenous *MCM10* locus of *S. cerevisiae* and tested for sensitivity to hydroxyurea (HU). HU inhibits replication by decreasing the cellular deoxyribonucleoside triphosphate (dNTP) pool. Mid-log phase cultures were incubated in 0.2 M HU for 60, 120 or 180 min before they were diluted and plated in the absence of HU to determine the rate of recovery. Figure 19A shows that all mutants were expressed at levels similar to wild-type Mcm10. Under our test conditions, wild-type cells doubled in number, and mutations that showed no effect on DNA binding *in vitro* (Asn313Ala/Lys314Ala) behaved in the same manner. Those mutants exhibiting a

modest decrease in DNA binding (His215Ala/Lys216Ala, corresponding to xMcm10 Lys293Ala) displayed a 2-fold decline in survival (Figure 19B). Viability was more strongly compromised in Phe230Ala/Phe231Ala mutants. These two phenylalanines correspond to Phe306 in xMcm10, which is implicated in DNA binding by our structural modeling but we were unable to test this directly because the protein could not be purified. Most strikingly, when Asn313 and Lys314 (corresponding to Lys385/386 in xMcm10) were changed to glutamic acid instead of alanine, survival was drastically reduced by more than 7-fold to about 30% even after a very short exposure to HU (Figure 19B). Taken together, these results extend our *in vitro* DNA binding studies and suggest that the corresponding residues in scMcm10 have an important role in DNA replication. Because neither Phe230/231 (located in the OB-fold) nor Asn313/Lys314 (located in the zinc finger loop) lie within the binding sites for pol α or PCNA (Das-Bradoo, Ricke et al. 2006; Ricke and Bielinsky 2006), it is highly likely that the HU sensitivity we detected is directly due to a defect of scMcm10 in DNA binding.

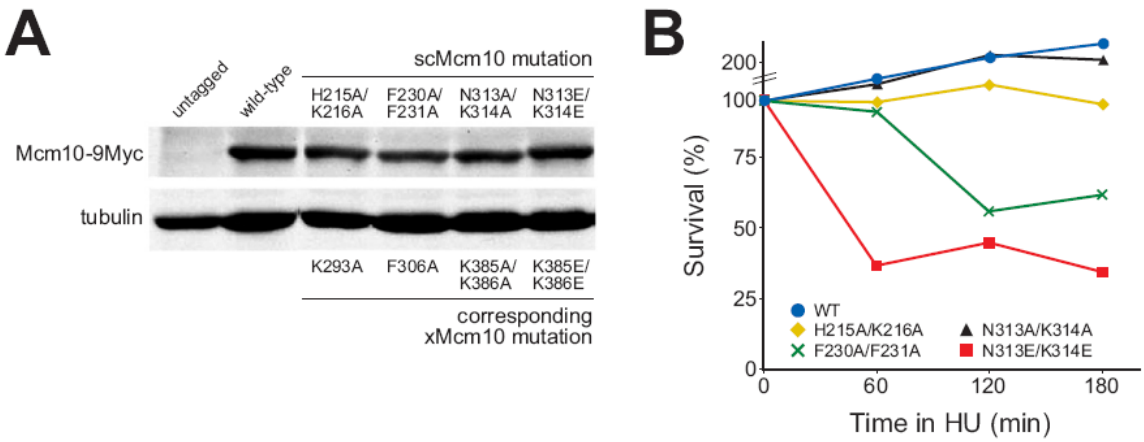


Figure 19. Mutations in the OB-fold and zinc finger domain of scMcm10 affect cell viability in hydroxyurea. (**A**), Total protein extracts prepared from mid-logarithmic phase cycling cells were analyzed by western blot with anti-Myc and anti-α-tubulin antibodies. (**B**), Survival of wild type, *mcm10-H215A/K216A, mcm10-F230A/F231A, mcm10-N313A/K314A,* and *mcm10-N313E/K314E* cells after treatment with 0.2 M hydroxyurea for 60, 120 or 180 min is shown in one representative experiment.

Discussion

Mcm10-ID is the highly conserved region of the protein that binds both DNA and pol α. The crystal structure of *Xenopus* Mcm10-ID reveals a unique arrangement of OB-fold and zinc finger motifs. NMR chemical shift perturbation and mutational analysis show that DNA spans both the hydrophobic β barrel of the OB-fold and the electrostatic, extended loop of the zinc finger. This model is further substantiated by the finding that substitutions of conserved key residues within the OB-fold and zinc finger in scMcm10 decrease cell viability in the face of replication stress.

At this time we are unable to reconcile the high-resolution structure of the core domain of Mcm10 with the proposed hexameric structure of the full-length human protein (Okorokov, Waugh et al. 2007). Efforts to dock our high-resolution crystal structure into the large volume within the hexameric EM reconstruction did not result in a clear solution, and thus is it not possible to draw specific conclusions regarding the

orientation of the DNA binding surface with respect to the hexameric ring. Moreover, the novel OB-fold/zinc finger configuration in Mcm10-ID bears no structural or functional resemblance to the OB-fold/zinc finger domain in the ring-shaped archaeal MCM helicase (Figure 16A) (Fletcher, Bishop et al. 2003). In addition, burying the ID within a hexameric assembly would likely occlude one or more sites of protein or DNA interactions described above. The functional implications of the Mcm10-ID structure, which likely contacts DNA and/or pol α on both leading and lagging strands, are more consistent with previous reports of an Mcm10 dimer (Cook, Kung et al. 2003; Lee, Seo et al. 2003; Fien and Hurwitz 2006; Robertson, Warren et al. 2008). Nevertheless, the Mcm10 structures serve as an important launching pad for further work to investigate the function of this essential eukaryotic replication factor.

The unique DNA binding platform observed here raises interesting questions regarding Mcm10 association with chromatin. The point at which origin DNA is initially denatured is undefined, but is likely to occur after Mcm2-7 and up to RPA loading onto chromatin. Mcm10 associates with chromatin after pre-RC formation and before Cdc45 and RPA (Wohlschlegel, Dhar et al. 2002; Gregan, Lindner et al. 2003; Ricke and Bielinsky 2004; Sawyer, Cheng et al. 2004). Given that the Mcm10 possesses two DNA binding domains and can bind to both ss- and dsDNA (Robertson, Warren et al. 2008), it is likely that the protein is anchored to DNA throughout replisome assembly. The lack of specificity for a particular DNA structure and the common ss/ds-DNA binding site within the ID imply that Mcm10 functions as a molecular scaffold to stabilize the replisome on DNA. As such, the effect of DNA binding residues on HU sensitivity in yeast strongly

suggests that Mcm10's DNA binding function is critical for fork integrity during DNA synthesis.

The structure of Mcm10-ID enables localization of residues identified previously as having an effect on DNA replication and cell viability (Supplementary Figure B5). The *cdc23-1E2* Cys239Tyr (Grallert and Nurse 1997) and *cdc23-M30* Leu287Pro (Liang and Forsburg 2001) mutations map to xMcm10 Leu323 and Leu369, respectively, which point into the core of the OB-fold β-barrel and are thus likely to disrupt the protein fold. Similarly, *cdc23-M36* Asp232Gly (Nasmyth and Nurse 1981) relates to an invariant aspartate (xMcm10 Asp313) on the interior of the L23 loop, and thus likely alters the conformation of L23. This loop might mediate Mcm10-protein interactions given its surface exposed location immediately outside of the DNA binding region (Supplementary Figure B5). Similarly, *cdc23-M36* Val265Ile and *mcm10-1* Pro269Leu mutations (Nasmyth and Nurse 1981; Maine, Sinha et al. 1984) map to solvent exposed positions on L45 and thus likely mediate intermolecular interactions. Indeed, these residues are adjacent to the putative pol α binding surface (described below), and extensive NMR chemical shift perturbation was observed in the L45 loop upon addition of DNA (Figure 17B).

The relative positions of residues that mediate protein-protein and protein-DNA interactions provide further insight into Mcm10's role at the replication fork. Both OB-fold and zinc finger motifs have been shown to mediate protein-protein interactions in addition to their DNA binding function (Matthews and Sunde 2002; Ball, Ehrhardt et al. 2007). We note that the putative PCNA interacting protein (PIP) box and Hsp10-like regions predicted from the scMcm10 sequence (Ricke and Bielinsky 2006) coincide with

the OB-fold β3 and β4 strands, respectively (Supplementary Figure B5). scMcm10 Tyr245 was found to be important for an interaction between diubiquinated scMcm10 and PCNA (Das-Bradoo, Ricke et al. 2006). This residue (xMcm10 Phe324) is located on the concave, DNA-binding face of β3 (Figure 18C) and had a modest effect on DNA binding (Table 4). We therefore speculate that it might contribute to DNA binding in unmodified Mcm10, but alters its interaction with DNA upon Mcm10 ubiquitination, which has been suggested to trigger a conformational change (Das-Bradoo, Ricke et al. 2006). Further work to elucidate the site of ubiquitination and its structural consequences will be required to understand the mechanistic basis for how Mcm10 modulates its interactions with DNA and PCNA.

Mapping functionally important residues onto this novel DNA binding platform supports the proposal that Mcm10 associates with pol α during initiation and elongation. scMcm10 Asn268 (xMcm10 Asn346) is important for Mcm10 stabilization of pol α and maps to the C-terminal end of the β4 strand (Supplementary Figure B5). We have previously shown that Mcm10-ID binds to the N-terminus of the catalytic p180 subunit of pol α *in vitro* (Robertson, Warren et al. 2008). The fact that Asn346 is surface exposed is consistent with a role for this residue in binding pol α. In addition, Asn346 is clearly located outside of the DNA binding interface, raising the possibility that Mcm10-ID can bind DNA and pol α simultaneously and consistent with the proposal that Mcm10 acts to recruit pol α to the origin (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006). Furthermore, evidence is mounting to suggest that Mcm10 likely associates with pol α during elongation (Ricke and Bielinsky 2004; Yang, Gregan et al. 2005; Pacek, Tutter et al. 2006; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007). It is intriguing

to speculate that Mcm10's ability to bind ssDNA, dsDNA, and pol α allows it to interact with both a ssDNA template and a duplex DNA product of the polymerase reaction. Structures of Mcm10 in complex with its binding partners will be critical to understand the physical basis for function of this modular, multi-functional protein.

Experimental Procedures

Mcm10 Purification

The coding sequence for amino acids 230-427 of *Xenopus laevis* Mcm10 was ligated into a modified pET-32a expression vector (Novagen) to produce an N-terminal thioredoxin-His_6 fusion protein. *E. coli* BL21(DE3) cells transformed with the Mcm10-ID/pET-32a plasmid were grown at 37°C in LB medium containing 100 μg/ml ampicillin, and protein was overexpressed for 16 hrs at 16° C upon addition of 0.5 mM IPTG. For NMR experiments, protein was uniformly enriched with ^{13}C and ^{15}N by propagating cells in M9 minimal medium supplemented with 2 mg/ml $^{13}C_6$-glucose and/or 1 mg/ml $^{15}NH_4Cl$ (Cambridge Isotope Laboratories). The cells were harvested in 50 mM Tris pH 7.5, 500mM NaCl, and 10% glycerol, and lysed under pressure (25,000 psi) using an Avestin EmulsiFlex C3 homogenizer. Mcm10-ID was purified by Ni^{2+}-NTA chromatography (Qiagen), followed by cleavage of the affinity tag. The cleaved protein was further purified by ssDNA-cellulose (Sigma), followed by gel filtration using a Superdex™ 200 preparative column (GE Healthcare) equilibrated in 20mM Tris pH 7.5, 150mM NaCl, and 5% glycerol.

X-ray Crystallography

Crystals were grown by hanging drop vapor diffusion by mixing 2 μl each of protein solution and a reservoir solution containing 100mM Tris pH 8.0, 100mM KSCN, and 40% PEG 4000. Rod-shaped crystals appeared overnight and grew to approximately 50 x 50 x 200 μm^3 after 2-3 days. Crystals were soaked 5 min in mother liquor containing 10% (v/v) glycerol and flash frozen in liquid nitrogen. X-ray diffraction data (Table 3) were collected at beamline 22-ID at the Advanced Photon Source (Argonne, IL) and processed with HKL2000 (Otwinowski and Minor 1997). Mcm10-ID crystallized in space group $P2_1$ with three molecules in the asymmetric unit.

Experimental X-ray phases were obtained from a multiwavelength anomalous dispersion (MAD) experiment using a single crystal that was soaked for 5 h at 25°C in mother liquor supplemented with 1 mM $Kau(CN)_2$. Diffraction data (Table 3) were collected at 110 K at energies corresponding to the peak (1.0388 Å), inflection (1.0370 Å) and high-energy remote (1.0311 Å) settings for the gold L_{III} absorption edge. The positions of 10 gold atoms in the asymmetric unit were located by automated Patterson searching using SHELXD (Uson and Sheldrick 1999), and initial phases to 3 Å were refined by solvent flattening using the SOLOMON option within autoSHARP (Vonrhein, Blanc et al. 2006). The model containing all three proteins in the asymmetric unit was built manually into the experimentally phased electron density using XtalView/Xfit (McRee 1999). Electron density for residues 230-234 (N-terminus), 420-427 (C-terminus), and 298-304 (loop L12) were unobserved.

The model was refined against the native X-ray data (50-2.3 Å) with a maximum likelihood target for experimental phases, as implemented in REFMAC 5.2 (Murshudov, A.A.Vagin et al. 1997). Improvements to the model were made by manual inspection of σ_A-weighted $2mF_o$-DF_c and mF_o-DF_c electron density maps, and they were judged successful by a decrease in R_{free} during refinement. Translation/libration/screw-rotation (TLS) refinement in REFMAC was used to model anisotropic motion of each protein domain (three in total). Individual anisotropic B-factors were derived from the refined TLS parameters and held fixed during subsequent rounds of refinement, which resulted in a decrease in both R and R_{free} and a noticeable improvement in the electron density maps (Supplementary Figure B1).

Analysis of the final structure using PROCHECK (Laskowski, MacArthur et al. 1993) showed 91.5% and 8.3% of the total of 515 residues to be within the favored and allowed regions of the Ramachandran plot, respectively. Only one residue, located at the extreme N-terminus, resided in the disallowed region. The coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under accession number 3EBE.

NMR Spectroscopy

Gradient enhanced ^{15}N-^{1}H HSQC NMR spectra were recorded at 25 °C using a Bruker DRX 800 NMR spectrometer equipped with single axis z-gradient cryoprobe. All spectra were acquired with 1024 complex points over a sweep width of 15 ppm in the ^{1}H dimension and 128 complex points over 37 ppm in the ^{15}N dimension. The center of the ^{15}N spectral width was set to 117.5 ppm, and the ^{1}H carrier was placed on the water signal

at 4.7 ppm from the respective base spectrometer frequencies. All spectra were processed and analyzed using Topspin v1.3 (Bruker, Billerica, MA) and Sparky v3.1 (University of California, San Francisco, CA). Data were treated with shifted sine-bell functions and zero-filled to twice the number of data points in both dimensions.

Backbone resonance assignments were made using 3D triple resonance experiments: HNCA, HNCACB, CBCA(CO)NH, (H)C(CO)NH-TOCSY, and HNCO. Acquisition parameters are provided in Table B1. All 3D-NMR spectra were collected at 25° C on a Bruker AVANCE 600 MHz spectrometer equipped with a cryoprobe.

Chemical shift perturbation data were collected by titrating unlabelled DNA into 250 μM ^{15}N-Mcm10-ID in 20 mM Tris-d_{11} (pH 7.0), 75 mM NaCl and 5% D_2O. Spectra were recorded at protein/DNA ratios of 1:0, 1:1, 1:3, and 1:5. The 15mer oligonucleotide d(GGCGCATTGTCGCAA) was used for ssDNA titrations. For dsDNA titrations, a similar sequence d(GGCACATTGTCCTCGTTTTCGAGGACAATGTGCC) was annealed into a 15-base pair hairpin by flash cooling from 80° C to 4° C. The observation of chemical shift perturbations in the fast exchange limit (on the NMR time scale) enabled the peaks in the free protein and DNA complexes to be correlated.

Molecular Modeling

The d(C_{10}) ssDNA was modeled onto the binding surface of Mcm10 with AMBER9 (D.A. Case 2006) restrained molecular dynamics simulated annealing (rMD-SA) techniques, guided by 25 loose distance restraints derived from the NMR chemical shift perturbation and ssDNA binding assay data. A starting conformation of the complex was created by superposition of the Mcm10 OB-fold with that of RPA70A from

the RPA70/ssDNA complex (PDB ID 1jmc) using the sequence-structure analysis functions of UCSF Chimera software (Pettersen, Goddard et al. 2004; Meng, Pettersen et al. 2006). The 8-mer RPA70 ssDNA coordinates were extended by 2 residues at the 3' end followed by energy minimization in AMBER. One hundred diverse ssDNA conformations were captured from 200ps of high-temperature molecular dynamics (1000K). Each of these conformations were then cooled to low temperatures over 20ps of simulation time while imposing the set of 25 sparse distance constraints to flexibly dock the ssDNA onto the Mcm10 surface. Mcm10 and the 3 residues at the 5′-end of the ssDNA (the DNA residues most closely associated with the RPA70A OB-fold) were held fixed during the heating and cooling process. Each of these 100 intermediate models was subjected to a final 20ps rMD-SA protocol at 600K with just the Mcm10 residues held fixed. More than 50% of the resulting 100 conformers after rMD-SA satisfied the large majority of the input restraints. In order to select a representative model for Figure 18, each member of the ensemble was inspected manually to search for local regions of poor agreement with the experimental data or incorrect chemical structure. This selection arrived independently at a conformer with one of the lowest molecular mechanics energies, indicating it to be one of the most favored by the AMBER force field as well as being in good agreement with the available experimental evidence.

Mutagenesis and in vitro DNA Binding Assays

xMcm10 mutants were prepared using a Quik-Change Kit (Stratagene) and purified similarly to wild-type protein, except that the ssDNA-cellulose affinity step was replaced with an SP-sepharose (GE Healthcare) ion exchange step. DNA binding to

Mcm10 mutants was measured by following an increase in fluorescence anisotropy as protein was added to a fluorescently-labeled oligonucleotide d(ATGGTAGGCAACCATGTAGTAGTCA) containing a 6-carboxyfluorescein moiety at the 3'-end, either alone (ssDNA) or annealed to a 1.2-fold molar excess of the complementary strand (dsDNA). For DNA length dependence measurements, 5-, 10-, and 15mer oligonucleotides were derived from the 5'-end of the 25mer sequence above. Protein was added over the concentration range of 0.1-50 μM to a solution containing 25 nM DNA in 20 mM Tris pH 7.5, 100mM NaCl, and 5% glycerol. Polarized fluorescence intensities using excitation and emission wavelengths of 495 and 515 nm, respectively, were measured for 30 s (1/s) and averaged. Anisotropy (r) was calculated using the equation $r = (I_{par}-I_{perp})/(I_{par}+2I_{perp})$, where I_{par} and I_{perp} are the observed fluorescence intensities recorded through polarizers oriented parallel and perpendicular to the direction of vertically polarized light. Dissociation constants (K_d) were derived by fitting a simple two-state binding model to data from three experiments using Kaleidagraph 3.5 (Synergy Software, PA).

Hydroxyurea Survival Assay

scMcm10 mutant yeast strains were constructed as previously described (Das-Bradoo, Ricke et al. 2006). pRS406-MCM10-9MYC (Aby491), pRS406-MCM10-9MYC-H215A/K216A (Aby492), pRS406-MCM10-9MYC-F230A/F231A (Aby496), pRS406-MCM10-9MYC-N313E/K314E (Aby503), and pRS406-MCM10-9MYC-N313A/K314A (Aby525) were integrated at the endogenous *MCM10* locus of Aby014 (W303). Total protein extracts were prepared from mid-logarithmic phase cycling yeast

cultures (OD_{600} = 0.6) as described previously (Ricke and Bielinsky 2006). Proteins were transferred to a nitrocellulose membrane and probed by western blot with anti-Myc (9E11, LabVision Neomarkers) for Myc-tagged scMcm10 and anti-α-tubulin (MMS-407R, Covance). For the hydroxyurea survival assay, cells were grown to mid-logarithmic phase (OD_{600} = 0.6). All mutants tested had doubling times comparable to wild type (data not shown). A 100 μl aliquot of cells was removed from each culture before adding 0.2 M hydroxyurea. 100 μl aliquots were removed at timed intervals, diluted, and colony-forming units were scored for viability on YPD plates as described previously (Allen, Zhou et al. 1994). Percentage survival was determined relative to cells that were not exposed to hydroxyurea at the beginning of the experiment.

Footnotes

The authors are indebted to Audrey Metz for her assistance in crystal preparation, and to Johannes Walter for stimulating discussions. We thank the staff at the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source (Argonne, IL), and Hassane Mchaourab and Albert Beth for access to fluorimeters. This work was funded by the National Institutes of Health (R01 GM080570 to B.F.E.; R01 GM065484 to W.J.C.; R01 GM074917 to A.-K.B.). Additional support for facilities came from the Vanderbilt Center in Molecular Toxicology (P30 ES000267) and the Vanderbilt Discovery Grant Program. E.M.W. was supported by the Molecular Biophysics Training Grant (T32 GM08320).

CHAPTER IV

CHARACTERIZATION OF PHYSICAL INTERACTIONS BETWEEN MCM10, DNA POLYMERASE α, AND DNA*

Abstract

Mcm10 is an essential eukaryotic protein required for the initiation and elongation phases of chromosomal replication. Specifically, Mcm10 is required for the association of several replication proteins, including DNA polymerase α (pol α), with chromatin. We showed previously that the internal (ID) and carboxy-terminal (CTD) domains of Mcm10 physically interact with both single-stranded (ss)DNA and the catalytic p180 subunit of pol α. However, the mechanism by which Mcm10 interacts with pol α on and off DNA is unclear. As a first step towards understanding the structural details for these critical intermolecular interactions, X-ray crystallography and NMR spectroscopy were used to map the binary interfaces between Mcm10-ID, ssDNA, and p180. The crystal structure of an Mcm10-ID/ssDNA complex confirmed and extended our previous evidence that ssDNA binds within the OB-fold cleft of Mcm10-ID. We show using NMR chemical shift perturbation and fluorescence spectroscopy that p180 also binds to the OB-fold, and that ssDNA and p180 compete for binding to this motif. In addition, we map a minimal Mcm10 binding site on p180 to a small region within the p180 amino-terminal domain (residues 286-310). These findings, together with data for DNA and p180 binding to an Mcm10 construct that contains both the ID and CTD, provide the first mechanistic insight

* The work presented in this chapter was submitted as Warren, E. M., Huang, H., Fanning, E., Chazin, W. J., and Eichman, B. F., (2009) *J Biol Chem.* In press.

into how Mcm10 might use a hand-off mechanism to load and stabilize pol α within the replication fork.

Introduction

In order to maintain their genomic integrity, cells must ensure complete and accurate DNA replication once per cell cycle. Consequently, DNA replication is a highly regulated and orchestrated series of molecular events. Multiprotein complexes assembled at origins of replication lead to assembly of additional proteins that unwind chromosomal DNA and synthesize nascent strands. The first event is the formation of a pre-replicative complex (pre-RC), which is composed of the origin recognition complex (ORC), Cdc6, Cdt1, and Mcm2-7 (reviewed in Arias and Walter 2007). Initiation of replication at the onset of S-phase involves the activity of cyclin- and Dbf4-dependent kinases, CDK and DDK, concurrent with recruitment of key factors to the origin. Among these, Mcm10 (Solomon, Wright et al. 1992; Merchant, Kawasaki et al. 1997) is recruited in early S-phase and is required for loading of Cdc45 (Wohlschlegel, Dhar et al. 2002). Mcm2-7, Cdc45, and GINS form the replicative helicase (Pacek and Walter 2004; Gambus, Jones et al. 2006; Moyer, Lewis et al. 2006; Pacek, Tutter et al. 2006). Origin unwinding is followed by loading of RPA, And-1/Ctf4, and pol α onto ssDNA (Tanaka and Nasmyth 1998; Walter and Newport 2000; Zou and Stillman 2000; Zhu, Ukomadu et al. 2007). In addition, recruitment of Sld2, Sld3, and Dpb11/TopBP1 are essential for replication initiation (Tanaka, Umemori et al. 2007; Zegerman and Diffley 2007), and association of topoisomerase I, proliferating cellular nuclear antigen (PCNA), replication factor C

(RFC), and the replicative DNA polymerases δ and ε completes the replisome (reviewed in Garg and Burgers 2005).

Mcm10 is exclusive to eukaryotes and is essential to both initiation and elongation phases of chromosomal DNA replication (Ricke and Bielinsky 2004; Gambus, Jones et al. 2006; Pacek, Tutter et al. 2006). Mutations in Mcm10 in yeast result in stalled replication, cell cycle arrest, and cell death (Nasmyth and Nurse 1981; Maine, Sinha et al. 1984; Solomon, Wright et al. 1992; Merchant, Kawasaki et al. 1997; Kawasaki, Hiraga et al. 2000). These defects can be explained by the number of genetic and physical interactions between Mcm10 and many essential replication proteins, including ORC, Mcm2-7, and PCNA (Merchant, Kawasaki et al. 1997; Homesley, Lei et al. 2000; Izumi, Yanagi et al. 2000; Hart, Bryant et al. 2002; Christensen and Tye 2003; Das-Bradoo, Ricke et al. 2006; Zhu, Ukomadu et al. 2007). In addition, Mcm10 has been shown to stimulate the phosphorylation of Mcm2-7 by DDK *in vitro* (Lee, Seo et al. 2003). Thus, Mcm10 is an integral component of the replication machinery.

Importantly, Mcm10 physically interacts with and stabilizes pol α and helps to maintain its association with chromatin (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007). This is a critical interaction during replication because pol α is the only enzyme in eukaryotic cells that is capable of initiating DNA synthesis *de novo*. Indeed, Mcm10 stimulates the polymerase activity of pol α *in vitro* (Fien, Cho et al. 2004), and, interestingly, the fission yeast Mcm10, but not *Xenopus* Mcm10, has been shown to exhibit primase activity (Fien and Hurwitz 2006; Robertson, Warren et al.). Mcm10 is composed of three domains, the N-terminal (NTD), internal (ID), and C-terminal (CTD) domains (Robertson, Warren et al. 2008). The NTD

is presumably an oligomerization domain, while the ID and CTD both interact with DNA and pol α (Robertson, Warren et al. 2008). The CTD is not found in yeast, while the ID is highly conserved among all eukaryotes. The crystal structure of Mcm10-ID showed that this domain is composed of an oligonucleotide/oligosaccharide binding (OB)-fold and a zinc finger motif, which form a unified DNA binding platform (Warren 2008). An Hsp10-like motif important for the interaction with pol α, has been identified in the sequence of *S. cerevisiae* Mcm10-ID (Ricke and Bielinsky 2004; Ricke and Bielinsky 2006).

DNA pol α-primase is composed of four subunits: p180, p68, p58, and p48. The p180 subunit possesses the catalytic DNA polymerase activity, and disruption of this gene is lethal (D'Urso, Grallert et al. 1995; LaRocque, Dougherty et al. 2007). p58 and p48 form the DNA-dependent RNA polymerase (primase) activity (Wang 1991; Arezi and Kuchta 2000), while the p68 subunit has no known catalytic activity but serves a regulatory role (Foiani, Marini et al. 1994; Ott, Rehfuess et al. 2002). Pol α plays an essential role in lagging strand synthesis by first creating short (7-12 nucleotide) RNA primers, followed by DNA extension. At the critical length of ~30 nucleotides, RFC binds to the nascent strand to displace pol α and loads PCNA with pols δ and ε (reviewed in Kunkel and Burgers 2008).

The interaction between Mcm10 and pol α has led to the suggestion that Mcm10 may help recruit the polymerase to the emerging replisome. However, the molecular details of this interaction and the mechanism by which Mcm10 may recruit and stabilize the pol α complex on DNA has not been investigated. Presented here is the high resolution structure of the conserved Mcm10-ID bound to ssDNA together with NMR

chemical shift perturbation competition data for pol α binding in the presence of ssDNA. Collectively, these data demonstrate a shared binding site for DNA and pol α in the OB-fold cleft of Mcm10-ID, with a preference for ssDNA over pol α. In addition, we have mapped the Mcm10-ID binding site on pol α to a 24-residue segment of the N-terminal domain of p180. Based on these results, we propose Mcm10 helps to recruit pol α to origins of replication by a molecular hand-off mechanism.

Results

The crystal structure of Mcm10-ID bound to ssDNA

The highly conserved internal domain of *Xenopus laevis* Mcm10 (Mcm10-ID) has previously been suggested to bind DNA along the surface formed by the concave OB-fold β-barrel and the extended zinc finger loop (Warren 2008). To elucidate the details of the Mcm10-ID/ssDNA interaction at high resolution, we determined the crystal structure of Mcm10-ID in complex with ssDNA using the unliganded Mcm10-ID structure as a molecular replacement search model (Figure 20). Strong electron density corresponding to three consecutive nucleotides of ssDNA was clearly visible inside the OB-fold cleft (Figure 20A), similar to the location of bound ssDNA in OB-fold structures of RPA, Rho, RecG, and RumA (Bochkarev, Pfuetzner et al. 1997; Bogden, Fass et al. 1999; Raghunathan, Kozlov et al. 2000; Singleton, Scaife et al. 2001; Lee, Agarwalla et al. 2005). These proteins typically bind ssDNA between OB-fold L12 and L45 loops, which are often flexible in the absence of DNA (Bhattacharya, Arunkumar et al. 2004). Consistent with other OB-fold/ssDNA complexes, Mcm10-ID's L12 loop (residues 297-305), which was not observed in the unliganded structure, is now visible in the complex

as the flexibility of this loop is quenched by interactions with the DNA (Figure 20B). The atomic model for Mcm10-ID/ssDNA was refined to 2.7 Å to a crystallographic residual of 19.7 % (R_{free} = 23.2%). Data collection and refinement statistics are shown in Table 1.

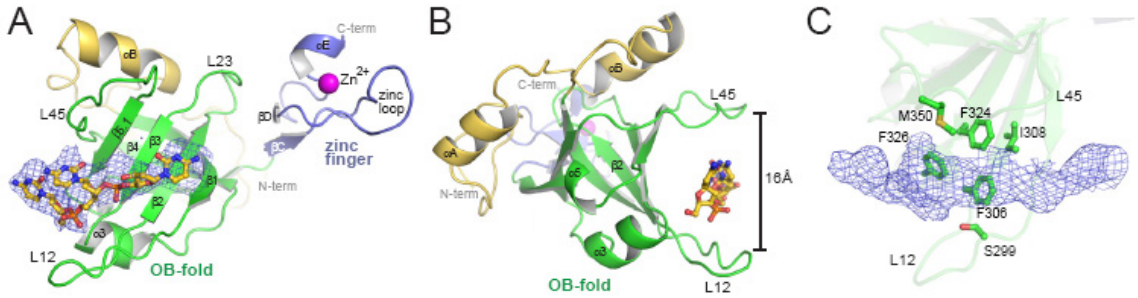


Figure 20. Crystal structure of Mcm10-ID bound to ssDNA. (**A**), The crystal structure of Mcm10-ID bound to ssDNA is shown as a ribbon representation. The various structural motifs are colored to distinguish the N-terminal helical region (gold), the OB-fold (green), the zinc-finger (blue with the zinc colored magenta), and the strands of the OB-fold are labeled as previously (Warren 2008). The DNA molecule is rendered as yellow sticks and fit to annealed omit electron density contoured at 3σ. (**B**), A side view of panel A illustrates the closing of loops L12 and L45 in the presence of ssDNA. (**C**), Close-up of the DNA binding site. The annealed omit electron density shows the location of the ssDNA. Mcm10-ID side chains contacting the DNA are rendered as sticks.

The electron density for the ssDNA molecule traverses β-strands β1-β3 and β5.1, which form the concave surface of the OB-fold cleft (Figure 20A). The polarity of the DNA is such that the 5′-end starts at the β5.1 strand and the 3′-end points toward β1 and the zinc finger, in a similar manner to the RPA70AB/ssDNA complex (Bochkarev, Pfuetzner et al. 1997). Refining the DNA in the opposite orientation had a detrimental impact on the crystallographic residual. The L12 and L45 loops wrap around the DNA, creating a channel ~16Å in diameter (Figure 20B). Polar and hydrophobic side chains from both loops and lining the β-sheet make van der Waals contact with the DNA, including Ser299, Phe306, Ile308, Phe324, Phe326, Met350, and Lys353 (Figure 20C, Supplementary Figure C2A). The high B-factors for the DNA (Table 5) indicate that the DNA is highly mobile and somewhat disordered within this hydrophobic cleft, which

precludes our ability to precisely model the DNA atoms that contact the protein. Nonetheless, the DNA binding surface on the protein and the polarity of the DNA are clearly defined. Electron density was not observed for the DNA around the zinc finger, presumably due to steric occlusion of the zinc finger binding site due to crystal packing. In this crystal form, the L45 loop of a neighboring molecule protrudes into the cleft between the OB-fold and the zinc finger, blocking ssDNA access to the zinc finger (Supplementary Figure C3). Although the entire DNA molecule cannot be identified from the present crystallographic data, this result confirms previous evidence that ssDNA binds directly to the OB-fold cleft (Warren 2008), and is consistent with the orientation of DNA observed in other OB-folds (Bochkarev, Pfuetzner et al. 1997; Bogden, Fass et al. 1999; Raghunathan, Kozlov et al. 2000).

Table 5. Data collection and refinement statistics for Mcm10-ID/ssDNA complex

Data collection	
Space group	$P3_121$
Cell dimensions	
a, *b*, *c* (Å)	95.02, 95.02, 61.16
α, β, γ (°)	90, 90, 120
Resolution (Å)	50.0-2.72 (2.72)
R_{sym}	8.2 (30.5)
$I / \sigma(I)$	11.38 (1.55)
Completeness (%)	90.2 (51.4)
Redundancy	5.3 (2.1)
Refinement	
Resolution (Å)	2.72
No. reflections	8221
R_{work} / R_{free}	0.197 / 0.232
No. atoms	
Protein	1395
DNA	54
Water	15
B-factors	
Protein	80.1
DNA	171.3
Water	69.3
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.245

Values in parentheses are for highest-resolution shell.

The structure of ssDNA-bound Mcm10-ID is nearly identical to the unliganded structure previously published with an rmsd of 0.77Å for all Cα atoms (Warren 2008). Apart from the now ordered L12 DNA-binding loop, the only notable difference between the two structures lies at the zinc finger helix at the extreme C-terminus of the ID (residues 405-416). This helix is well-defined in the unliganded protein and is engaged

in intermolecular protein-protein contacts in each of the three protomers in the asymmetric unit (Supplementary Figure C4) (Warren 2008). In the complex, which crystallizes in a different lattice with one protein/DNA complex per asymmetric unit, the zinc finger helix is disordered past the Zn^{2+}-coordinating His406. This local unfolding is presumably due to the lack of any intermolecular contacts in the present crystal lattice, and suggests that the fold of this helix in the full-length protein may be stabilized through protein contacts outside of the ID.

Mcm10-ID binds to p180$^{189-323}$

Mcm10-ID has previously been shown to bind to the N-terminal 323 residues of the p180 subunit of human pol α-primase (Robertson, Warren et al. 2008). This region is highly conserved but lacks appreciable predicted secondary structure or sequence complexity. In order to map the Mcm10-p180 interaction in detail, p180$^{1-323}$ was subjected to limited proteolysis and the resulting stable fragments were identified by mass spectrometry. Proteolytically sensitive sites were found at residues 145 and 189. Consequently, p180$^{1-145}$ and p180$^{189-323}$ were sub-cloned, purified, and tested for physical interaction with Mcm10-ID by affinity chromatography pull-down assays. The p180$^{1-145}$ protein was not sufficiently stable in solution to test for a putative interaction. However, GST-tagged p180$^{189-323}$ immobilized on glutathione-sepharose was able to capture His-tagged Mcm10-ID from solution (Figure 21A), demonstrating that this region of the p180 subunit is sufficient to bind to Mcm10-ID. The strength of the Mcm10-ID-p180$^{189-323}$ interaction was quantified using a fluorescence anisotropy assay. Titration of unlabeled Mcm10-ID into a solution of fluorescein-labeled p180$^{189-323}$ resulted in a robust increase

in fluorescence anisotropy, whereas addition of either ssDNA or buffer alone had no effect (Figure 21B). Analysis of the titration data by a single-state binding model provided an apparent dissociation constant (K_d) of 12 ± 2 μM for Mcm10-ID binding to $p180^{189-323}$. This value is in good agreement with the K_d (30 ± 1 μM) determined by isothermal titration calorimetry using unlabeled proteins (Figure 21C).

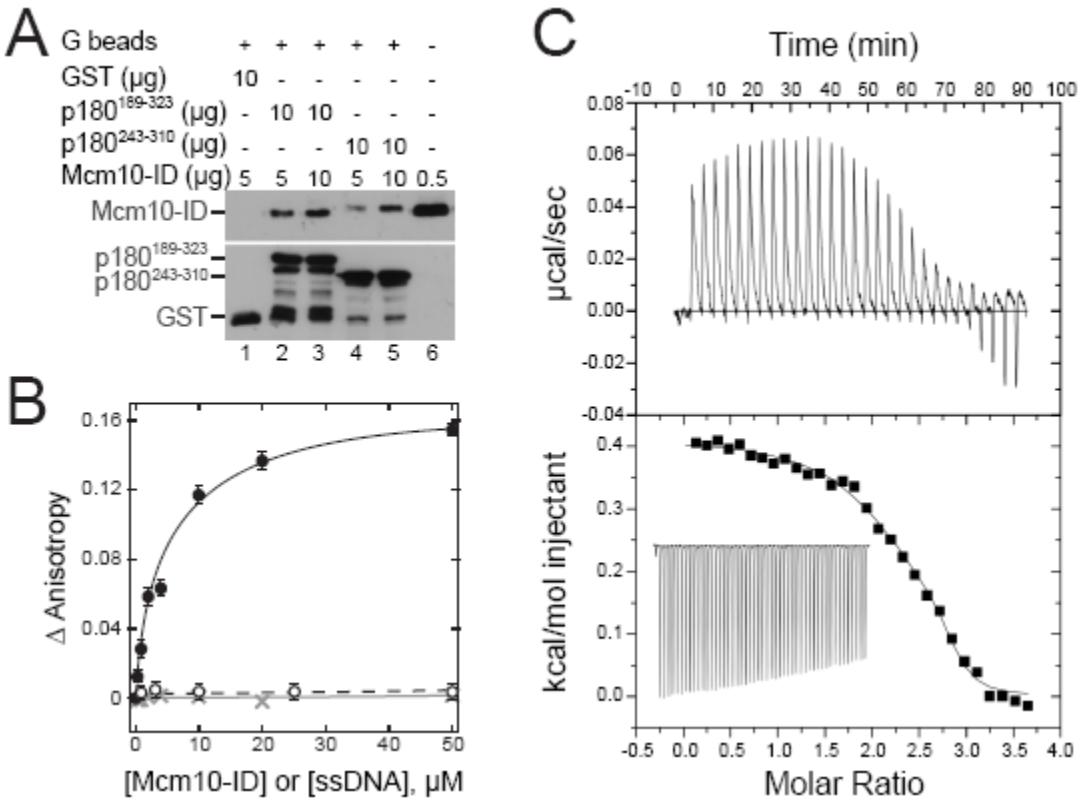


Figure 21. $p180^{189-323}$ physically interacts with Mcm10-ID. (**A**), Affinity capture experiments between His-tagged Mcm10-ID and GST-tagged p180 fragments. p180 fragments 189-323 and 243-310 were adsorbed on glutathione (G) beads and mixed with Mcm10-ID in the amounts indicated. Bound Mcm10-ID was detected by western blotting with α-His antibody (upper blot) while retention of p180 fragments was detected by western blotting with α-GST antibodies (lower blot). (**B**), Protein-protein binding was monitored by following a change in fluorescence anisotropy as Mcm10-ID (closed circles) or 25mer-ssDNA (open circles) was titrated into a solution containing MTS-fluorescein labeled $p180^{189-323}$ (open circles). A control in which only buffer was added to MTS-fluorescein-$p180^{189-323}$ is also shown (crosses).

Figure 21, continued. Error bars represent the standard deviation from the average values from three independent measurements. The dissociation constant (K_d) for Mcm10-ID binding to $p180^{189-323}$ was determined to be 12 ± 2 μM as described in Materials and Methods. (**C**), Isothermal titration calorimetry measurements for $p180^{189-323}$ titrated into Mcm10-ID at 21ºC. *Upper panel*, raw ITC data for sequential injections of $p180^{189-323}$; *lower panel*, integrated heat responses (squares) fit with a single site binding model (continuous line). Best fit parameters for the curve fit were K_d = 30 ± 1 μM, ΔH = 0.4 kcal/mol, and TΔS = 6.5 kcal/mol. The inset shows the heat effects resulting from successive injections of buffer into protein that was subtracted from binding isotherms.

NMR chemical shift perturbation experiments were used to probe the Mcm10-ID and $p180^{189-323}$ interaction in greater detail and map the $p180^{189-323}$ binding site on Mcm10-ID. We previously obtained sequence specific backbone assignments of Mcm10-ID and used chemical shift perturbation to map the DNA binding site (Warren 2008). Here, we monitored ^{1}H and ^{15}N chemical shift perturbations of uniformly ^{15}N-labeled Mcm10-ID upon addition of unlabeled $p180^{189-323}$ (Figure 22A). These experiments revealed a number of significant chemical shift perturbations in the 2D ^{15}N-^{1}H HSQC spectrum that mapped onto the OB-fold cleft, with the strongest perturbations observed for residues in β1, β2, β5.1, L12, and L45 (Figure 22B,C). In fact, the resulting spectrum for the Mcm10-ID/$p180^{189-323}$ complex is remarkably similar to that previously measured for Mcm10-ID/ssDNA complex (Figure C9) (Warren 2008), suggesting that Mcm10-ID utilizes a common binding site for both ssDNA and p180. Moreover, both $p180^{189-323}$ and ssDNA bind to Mcm10-ID in the fast-to-intermediate-exchange regime, with some peaks gradually shifting over the course of the titration while others broaden and disappear. However, by comparing the magnitude of the chemical shift perturbations in response to $p180^{189-323}$ and ssDNA binding, it appears that Mcm10-ID binds more

weakly to p180$^{189-323}$ than to ssDNA (data not shown). This observation is consistent with the 4-fold weaker Mcm10 dissociation constant determined by fluorescence anisotropy (3 μM for ssDNA versus 12 μM for p180$^{189-323}$).

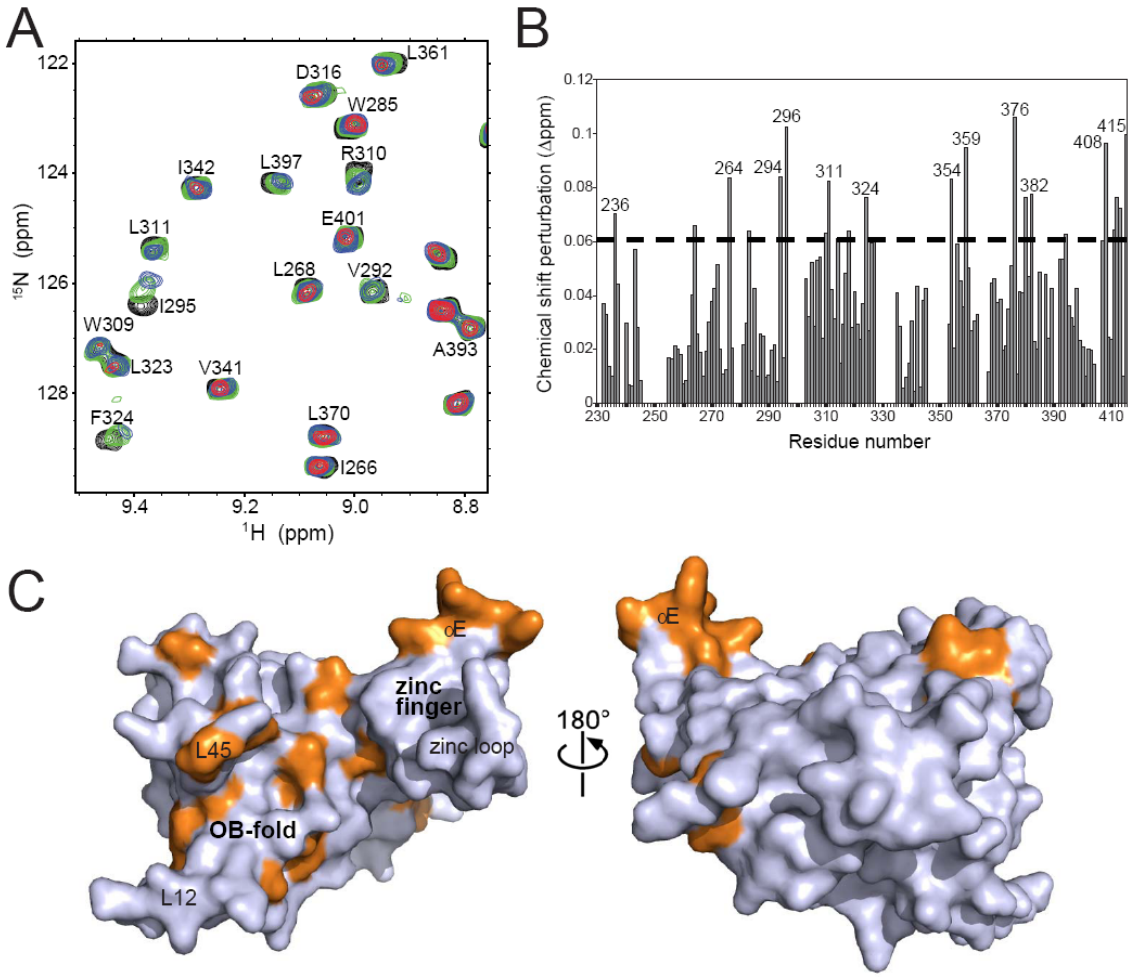


Figure 22. Mapping the p180$^{189-323}$ binding site onto Mcm10-ID. (**A**), An overlay of a region of ^{15}N-^{1}H HSQC spectra of ^{15}N-enriched Mcm10-ID in response to the addition of p180$^{189-323}$, with the peak assignments labeled. The titration was performed at Mcm10-ID:p180$^{189-323}$ ratios of 1:0 (black), 1:0.25 (green), 1:0.67 (blue), and 1:1 (red). (**B**), Quantitation of chemical shift perturbations from addition of equimolar amounts of Mcm10-ID and p180$^{189-323}$ (see Materials and Methods). The dashed line denotes 1 standard deviation above the mean. A shift of zero indicates an unassigned residue. (**C**), Surface representation of Mcm10-ID with the residues exhibiting a significant shift (above the dashed line in panel B) colored orange. The 180º rotation illustrates that the shifts occur almost exclusively on the ssDNA binding face of the protein.

ssDNA and p180$^{189-323}$ compete for the same site on Mcm10-ID

A common binding site for ssDNA and p180 suggests that these two ligands either compete for binding or bind cooperatively to Mcm10. To distinguish between

these two possibilities, competition experiments were performed utilizing NMR chemical shift perturbations to monitor the interaction of $p180^{189-323}$ and ssDNA with Mcm10-ID (Figure 23A, Supplementary Figure C5A). First, ssDNA was titrated into a sample containing ^{15}N-labeled Mcm10-ID and peak shifts were observed as previously reported (Figure 23A, red spectrum) (Warren 2008). Next, unlabeled $p180^{189-323}$ was titrated into the same sample containing ssDNA-saturated Mcm10-ID (Figure 23A, green spectrum). No further chemical shift perturbations were observed with the addition of protein, which suggests that $p180^{189-323}$ neither interacts with an Mcm10-ID/ssDNA complex nor does it displace ssDNA from Mcm10-ID.

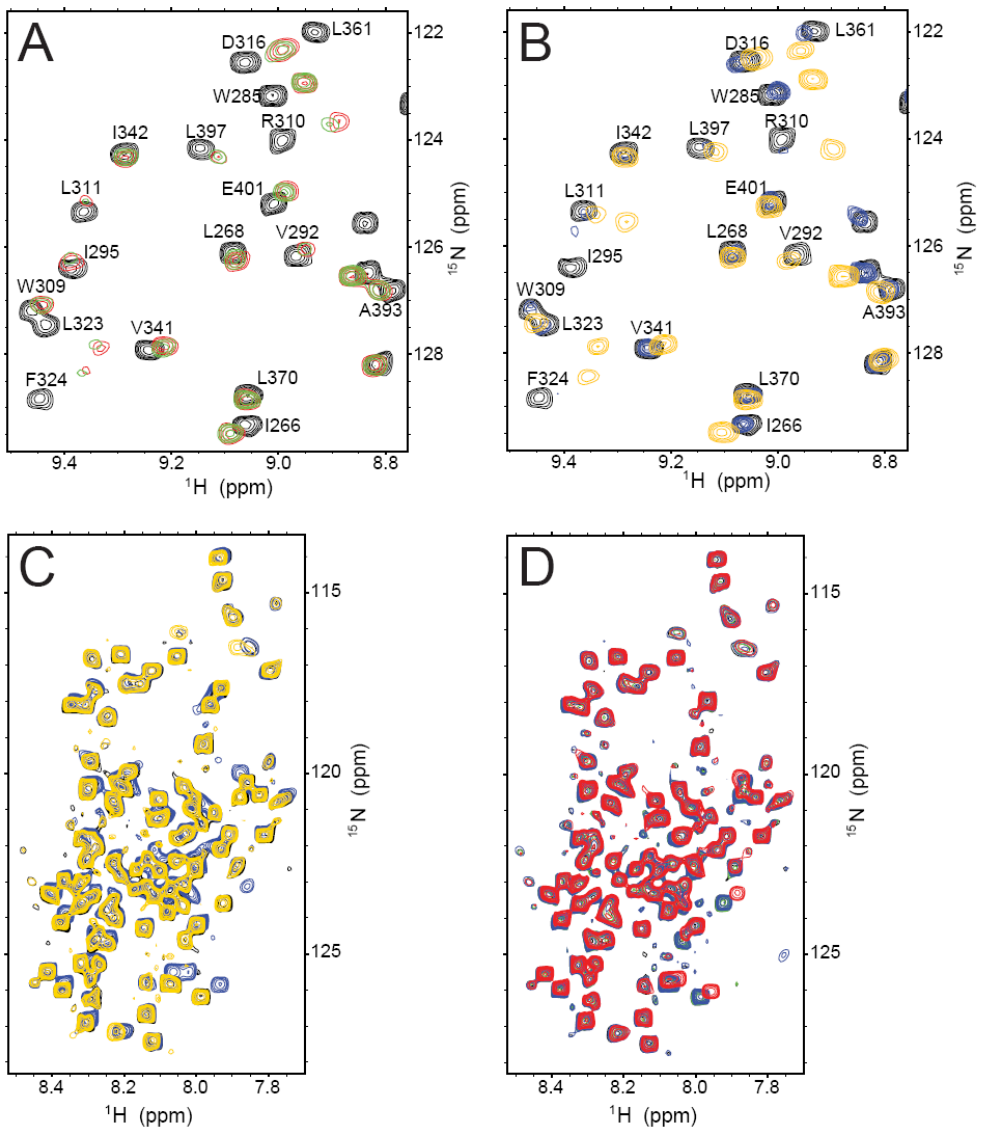


Figure 23. Competition for Mcm10-ID binding by ssDNA and p180$^{189\text{-}323}$. (**A**), NMR chemical shift perturbations in response to ssDNA and p180$^{189\text{-}323}$ binding to ^{15}N-enriched Mcm10-ID. The region of the overlaid spectra is the same as in Figure 22A, with Mcm10-ID alone (black), 1:1 ratio of Mcm10-ID:ssDNA (red), and a 1:1:1 ratio Mcm10-ID:ssDNA:p180$^{189\text{-}323}$ (green). (**B**), The reverse titration with the same overlay region of the ^{15}N-^{1}H HSQC spectra as before. Mcm10-ID alone (black) was mixed in a 1:1 ratio of p180$^{189\text{-}323}$ (blue), then ssDNA was added to this mixture in a 1:1:1 ratio (gold). (**C**), Reciprocal titrations from panels A and B, performed with ^{15}N-enriched p180$^{189\text{-}323}$ (black), 1:1 molar ratio of p180$^{189\text{-}323}$:Mcm10-ID (blue), and a 1:1:1 ratio of p180$^{189\text{-}323}$:Mcm10-ID:ssDNA (gold). (**D**), The reverse titration as in panel C with p180$^{189\text{-}323}$ alone (black), 1:1 molar ratio of p180$^{189\text{-}323}$:ssDNA (green), a 1:1:0.67 molar ratio of p180$^{189\text{-}323}$:ssDNA:Mcm10-ID (blue), and then a 1:1:1 molar ratio of ^{15}N-p180$^{189\text{-}323}$:ssDNA:Mcm10-ID (red).

To test whether ssDNA is able to disrupt a pre-formed Mcm10-ID/p180$^{189\text{-}323}$ complex, the reverse titration was performed in which unlabeled p180$^{189\text{-}323}$ was first

added to a sample containing ^{15}N-labeled Mcm10-ID followed by addition of ssDNA (Figure 23B, Supplementary Figure C5B). As in Figure 22, addition of p180$^{189-323}$ to ^{15}N-labeled Mcm10-ID resulted in significant perturbation in chemical shifts for a discrete set of residues (Figure 23B, blue spectrum). Upon addition of ssDNA, the peaks that were perturbed by p180$^{189-323}$ changed trajectory and shifted to resemble the Mcm10-ID/ssDNA spectrum (Figure 23B, gold spectrum).

To test the ability of ssDNA to displace Mcm10 from p180, we performed a NMR titration in which Mcm10-ID and ssDNA were added in succession to a solution containing ^{15}N-enriched p180$^{189-323}$. When Mcm10-ID was titrated into ^{15}N-p180$^{189-323}$, displacement of a discrete number of chemical shifts was observed, indicative of formation of the Mcm10-ID/p180 complex (Figure 23C, Supplementary Figure C6A, blue spectrum). Addition of ssDNA to the protein complex caused the chemical shifts to revert back to their starting location in the spectrum of ^{15}N-p180$^{189-323}$ alone (Figure 23C, gold). This directly demonstrates that ssDNA is capable of displacing Mcm10-ID from p180$^{189-323}$. A fourth titration was performed in which ssDNA was first added into the ^{15}N-labeled p180$^{189-323}$ sample. In this case, no peak shifts were observed, indicating that p180$^{189-323}$ does not bind ssDNA (Figure 23D, Supplementary Figure C6B, blue spectrum). When Mcm10-ID was titrated into this sample containing free p180$^{189-323}$ and ssDNA, perturbation of p180 chemical shifts that mimicked the ^{15}N-p180$^{189-323}$/Mcm10-ID spectrum were observed (compare blue spectrum in Figure 23C with red spectrum in Figure 23D). However, the magnitude of Mcm10-ID induced ^{15}N-p180$^{189-323}$ peak shifts in the presence of ssDNA were not as large as those in the absence of ssDNA, consistent with a partitioning of Mcm10 between both p180$^{189-323}$ and ssDNA. Taken together,

these data demonstrate that ssDNA and p180$^{189\text{-}323}$ compete for binding to the OB-fold cleft of Mcm10-ID, and that ssDNA is able to displace p180$^{189\text{-}323}$ from Mcm10-ID, consistent with the moderate preference of Mcm10-ID for ssDNA over p180$^{189\text{-}323}$.

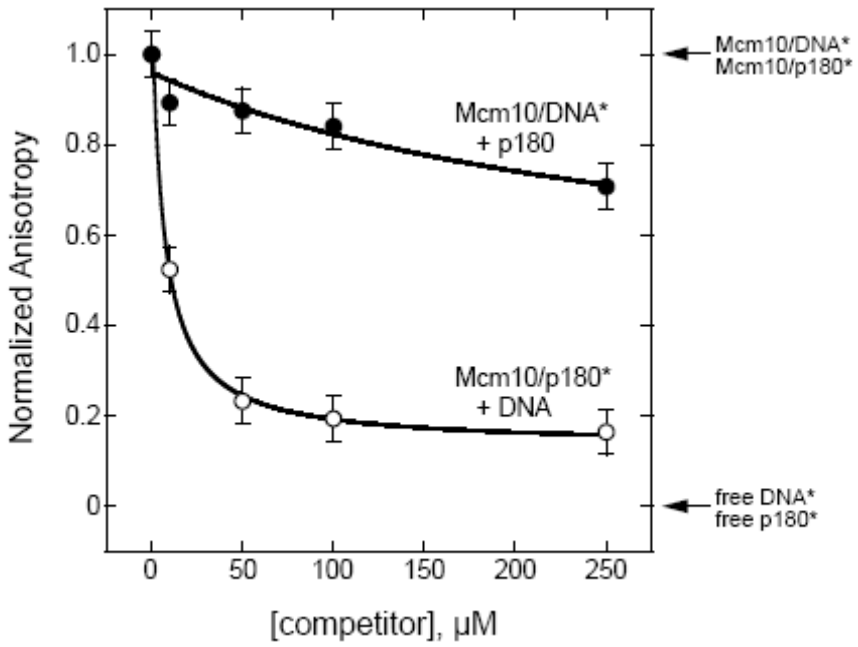


Figure 24. Quantitation of p180$^{189\text{-}323}$ and ssDNA displacement from Mcm10-ID. The decay in fluorescence anisotropy was monitored as unlabeled ssDNA was added to preformed Mcm10-ID/fluorescein-p180$^{189\text{-}323}$ complexes (open circles), and as unlabeled p180$^{189\text{-}323}$ was added to preformed Mcm10-ID/fluorescein-ssDNA complexes (closed circles). Mcm10-ID and fluorescein-labeled molecules (DNA*, p180*) were held at 15 μM and 50 nM, respectively. Anisotropy (*r*) values were normalized to correct for the differences in signal between DNA* and p180*, in which *r*=0 reflects the anisotropy for free DNA* and p180* (prior to addition of Mcm10-ID), and *r*=1 reflects the anisotropy for Mcm10-ID/ssDNA* and Mcm10-ID/p180* before addition of unlabeled competitor.

To quantify the competition of ssDNA and p180$^{189\text{-}323}$ for Mcm10-ID, we examined the concentration dependence on the displacement reaction using the fluorescence anisotropy assay (Figure 24). In separate experiments, Mcm10-ID/fluorescein-ssDNA and Mcm10-ID/fluorescein-p180$^{189\text{-}323}$ complexes were preformed, and the decay in fluorescence anisotropy was plotted as increasing amounts of unlabeled competitor was added. Upon addition of ssDNA to Mcm10-ID/fluorescein-p180$^{189\text{-}323}$, we observed a return in the anisotropy signal back to within 15% of that of

free fluorescein-p180$^{189-323}$, demonstrating a robust displacement of p180$^{189-323}$ from Mcm10 by ssDNA. The calculated K_i for this reaction is 2.8 ± 0.2 μM, roughly equal to the K_d for the Mcm10-ID/ssDNA interaction (3.4 ± 0.5 μM). Conversely, addition of p180$^{189-323}$ to Mcm10/fluorescein-ssDNA complexes reduced the anisotropy signal only ~30% at the highest concentration of p180 tested (Ki>50 μM), indicating that p180$^{189-323}$ only weakly competes for binding. These results are fully consistent with the chemical shift perturbation experiments described above.

The minimal Mcm10-ID binding site maps to p180$^{286-310}$

To map the minimal region of p180$^{189-323}$ needed to interact with Mcm10-ID, the p180 sequence was aligned and examined for conservation and predicted secondary structure. In addition, the data from the ^{15}N-^{1}H HSQC titration of ^{15}N labeled p180$^{189-323}$ with unlabeled Mcm10-ID (Supplementary Figure C6) was carefully examined to determine if insights could be obtained into the location of the Mcm10-ID binding site on p180$^{189-323}$, even in the absence of sequence specific assignments, following the strategy described previously (Mer, Bochkarev et al. 2000). The key to this approach is to monitor the total number of resonances perturbed in the titration, and take advantage of the unique chemical shifts of the glycine backbone and glutamine and asparagine side chain amides. Analysis of the data in this way suggests that the binding sequence should contain approximately 20 residues, including at least one glycine and no asparagine or glutamine residues. In the sequence of p180$^{189-323}$, only two peptides fit these criteria, p180$^{243-256}$ and p180$^{286-310}$. Indeed, glutathione-immobilized GST-p180$^{243-310}$, which spanned both peptides, was able to capture Mcm10-ID from solution in our affinity

chromatography assay (Figure 21A). To assess which peptide contains the Mcm10-ID binding sequence, the individual p180$^{243-256}$ and p180$^{286-310}$ peptides were synthesized and used in chemical shift perturbation experiments with ^{15}N-enriched Mcm10-ID. Of the two peptides, only p180$^{286-310}$ induced significant chemical shift perturbations in Mcm10-ID that resembled those of p180$^{189-323}$ (Figure 25A, Supplementary Figures C7). Binding of p180$^{286-310}$ occurred on the fast-exchange timescale and resulted in a magnitude of chemical shift perturbations similar to those caused by p180$^{189-323}$ (Supplemntary Figure C7). These data show that p180 residues 286-310 bind to the same region of Mcm10-ID as p180$^{189-323}$ and ssDNA (Figure 25A,B), and are consistent with the relative binding affinities for p180$^{189-323}$ (K_d = 12 ± 2 μM) and p180$^{286-310}$ (K_d = 32 ± 2 μM) measured by fluorescence anisotropy (Supplementary Figure C7E).

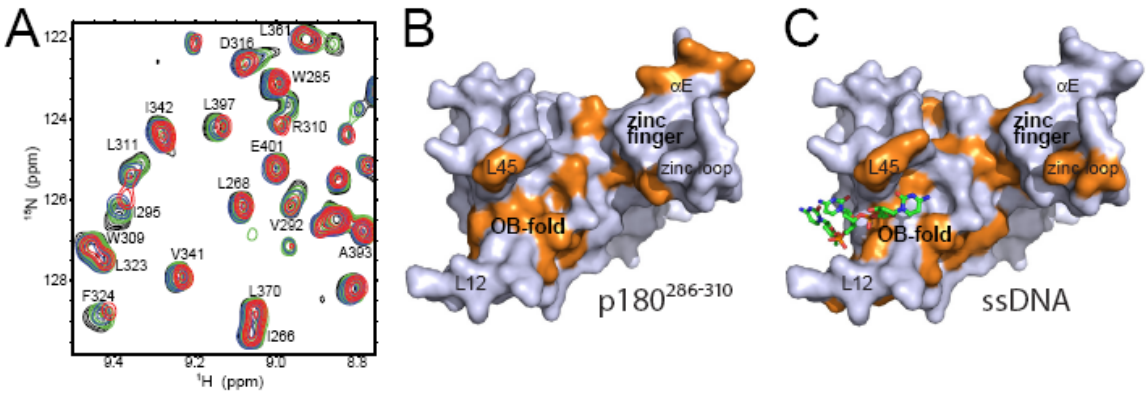


Figure 25. Mapping the p180$^{286-310}$ binding site into Mcm10. (**A**), Overlay of a representative section of the ^{15}N-^{1}H HSQC spectra from ^{15}N-enriched Mcm10-ID performed at Mcm10-ID:p180$^{286-310}$ ratios of 1:0 (black), 1:0.25 (green), 1:0.5 (blue), and 1:1 (red). (**B**), Surface representation of Mcm10-ID showing that residues exhibiting a significant shift in response to p180$^{286-310}$ (orange) predominate on the ssDNA binding face of the protein. (**C**), The Mcm10-ID/ssDNA co-crystal structure, with NMR chemical shift perturbations from titration with ssDNA highlighted orange (data from Warren 2008).

Mcm10-ID+CTD binds ssDNA and p180$^{189-323}$

Having thoroughly characterized the binding of ssDNA and p180 to Mcm10-ID, we asked how studies of the isolated domain relate to the biochemical functions of the intact protein. Mcm10-CTD has previously been shown to bind both DNA and p180$^{1-323}$ (Robertson, Warren et al. 2008). The linking of the ID and CTD are therefore anticipated to result in higher affinity and possibly altered specificity. To this end, a protein deletion construct encompassing Mcm10-ID and -CTD was constructed, purified, and characterized by biochemical approaches.

Binding of Mcm10-ID+CTD to various DNA substrates designed to resemble replication intermediates was measured by fluorescence anisotropy (Table 2). Mcm10-ID+CTD bound all DNAs tested with 10-fold greater affinity than previously determined for ID or CTD alone. This observation is similar to our previous results obtained with an MBP-tagged full-length Mcm10 (Robertson, Warren et al. 2008). Additionally, the ID+CTD protein bound ssDNA with a slightly higher affinity than dsDNA when tested against both 25mer and 45mer oligonucleotides, as observed previously for the isolated domains and the intact protein (Fien, Cho et al. 2004; Robertson, Warren et al. 2008; Warren 2008) (Table 2). Moreover, Mcm10-ID+CTD does not demonstrate a significant preference for ssDNA, dsDNA, or constructs containing ss/dsDNA junctions, including 5′- and 3′-overhangs, fork and bubble substrates (Table 2). This lack of specificity for a particular DNA structure was observed previously for the isolated ID and CTD. Thus, together these two binding modules enhance the strength of the DNA interaction but do not provide additional specificity.

Binding of Mcm10-ID+CTD to p180$^{189-323}$ was investigated by fluorescence anisotropy using fluorescein-tagged p180$^{189-323}$ protein. Mcm10-ID+CTD bound to fluorescein-p180$^{189-323}$ with a K_d of 0.19 ± 0.03 μM (Table 2). The strength of the Mcm10-ID+CTD interaction with p180$^{189-323}$ is ~50-fold greater than that measured for Mcm10-ID alone (K_d = 12 ± 2 μM). Importantly, the affinity observed for this tandem construct brings the strength of binding of p180 to the same level as for ssDNA (Table 2). This has critically important implications for understanding the function of Mcm10 in recruiting p180 (and therefore pol α) to the active replication machinery.

Table 2. Mcm10-ID+CTD binding to DNA and pol α-p180$^{189-323}$

	K_d (μM)[1]	Relative affinity[2]	DNA sequence[3]
ssDNA 25mer	0.22 ± 0.05	1.0	5′-ATGGTAGGCAACCATGTAGTAGTCA*
dsDNA 25mer	0.91 ± 0.10	0.2	5′-ATGGTAGGCAACCATGTAGTAGTCA* TACCATCCGTTGGTACATCATCAGT-5′
ssDNA 45mer	0.12 ± 0.02 (4.30 ± 1.12)	1.8 (0.05)	5′-GGTAGGCACGAACCATGTAGTAGTAGGCAATCAGCATTTGATAGC*
dsDNA 45mer	0.40 ± 0.08 (34.86 ± 8.70)	0.6 (0.01)	5′-GGTAGGCACGAACCATGTAGTAGTAGGCAATCAGCATTTGATAGC* CCATCCGTGCTTGGTACATCATCATCCGTTAGTCGTAAACTATCG-5′
Fork	0.11 ± 0.01	2.0	5′-GGTAGGCACG**AACCATGTAGTAGTA*** CCATCCGTGC**ATGATGATGTACCAA**-5′
Bubble	0.06 ± 0.003	3.7	5′-GGTAGGCACG**AACCATGTAGTAGTA**GGCAATCAGC* CCATCCGTGC**ATGATGATGTACCAA**CCGTTAGTCG-5′
5'-overhang	0.11 ± 0.01	2.0	5′-TGACTACTACATGGT *ACTGATGATGTACCA**ACGGATGGTA**-5′
3'-overhang	0.05 ± 0.004	4.4	5′-ATGGTTGCCTACCAT ***ACTGATGATG**TACCAACGGATGGTA-5′
p180$^{189-323}$	0.19 ± 0.03	1.2	

[1] Values determined by fluorescence anisotropy as described in Experimental Procedures. Values in parentheses recorded at 300 mM NaCl; all others at 150 mM NaCl.
[2] Binding affinities relative to ssDNA 25mer
[3] ssDNA regions of ss/dsDNA hybrids are in boldface

Discussion

Chemical nature of Mcm10 interactions with DNA and pol α

In this study, we show that both ssDNA and the N-terminal region of p180 compete for binding to a relatively hydrophobic surface within the OB-fold cleft of Mcm10. Our previous analysis showed that in addition to the OB-fold, ssDNA binds to the highly basic extended loop on the zinc finger motif (Warren 2008). In the structure of the Mcm10/ssDNA complex, the crystal lattice prevented DNA access to the zinc loop, which precluded direct visualization of the interaction between DNA and the zinc finger. However, additional information regarding the nature of intermolecular Mcm10-ID interactions can be extracted from thermodynamic information derived from ITC measurements. Titration of Mcm10-ID with ssDNA (Supplementary Figure C2B) revealed an enthalpically driven, spontaneous reaction (ΔH = -9.8 kcal/mol, $T\Delta S$ = -3.6 kcal/mol). This is consistent with our previous NMR titration and mutational analyses that showed electrostatic interactions play a large role in ssDNA binding to Mcm10-ID (Warren 2008). Taking the structural and biochemical data together, binding of ssDNA to Mcm10-ID is largely mediated by polar/charged residues located on the L12 and L45 loops (e.g., Ser299 and Lys353), between the OB-fold and the zinc finger (Lys293), and on the zinc loop (Lys385 and Lys386). In contrast, calorimetric titration of Mcm10-ID with $p180^{189\text{-}323}$ (Figure 21C) revealed a large entropic contribution (ΔH = 0.4 kcal/mol; $T\Delta S$ = 6.5 kcal/mol), suggesting that hydrophobic interactions are important to the protein-protein interaction. Indeed, $p180^{189\text{-}323}$ and $p180^{286\text{-}310}$ binding mapped to aliphatic residues within the OB-fold cleft (e.g., Phe324). Interestingly, NMR chemical

shifts corresponding to basic residues on the zinc finger helix (αE), and not the DNA-binding zinc loop, were perturbed by all three p180 constructs tested, including $p180^{243-256}$, which perturbed helix αE almost exclusively (Figure C9). Thus, the hydrophobic interaction at the OB-fold likely provides additional specificity for $p180^{286-310}$.

Pol α and Mcm10 binding domains

The overall domain structure of p180 is known, and the activities of the central polymerase and the C-terminal subunit-assembly domains have been characterized (Mizuno, Yamagishi et al. 1999; Muzi-Falconi, Giannattasio et al. 2003). However, the function of the N-terminal domain is less clear. This region of p180 is dispensable for polymerase activity and is not required for assembly of the pol α-primase complex. The N-terminus of p180 is phosphorylated by Cyclin A-Cdk2 on residues 174, 209, and 219 (Nasheuer, Moore et al. 1991; Schub, Rohaly et al. 2001) and it interacts with several proteins of various functions including Mcl1 (And-1/Ctf4) (Williams and McIntosh 2005), PP2A (Schub, Rohaly et al. 2001), ConA and RCA (Hsi, Copeland et al. 1990), SV40 Large T-antigen (Dornreiter, Copeland et al. 1993; Taneja, Nasheuer et al. 2007), and Mcm10 (Robertson, Warren et al. 2008). Although the importance of these interactions has yet to be determined, our observation that Mcm10 interacts with $p180^{286-310}$, outside of the polymerase domain, is consistent with Mcm10 anchoring the pol α complex onto DNA in such a way as to not interfere with RNA or DNA synthesis.

The finding that Mcm10-ID interacts with both ssDNA and pol α through contacts in the OB-fold domain reflects the adaptability of this motif to bind a range of different biological molecules. This cleft is used by various proteins to engage RNA (Berthet-

Colominas, Seignovert et al. 1998), DNA (Bochkarev, Pfuetzner et al. 1997), oligosaccharides (Stein, Boodhoo et al. 1994), proteins (Bochkareva, Belegu et al. 2001), and even metals and inorganic phosphates (Heikinheimo, Lehtonen et al. 1996; Hall, Gourley et al. 1999). For example, RPA, a eukaryotic recruiting and scaffolding protein critical to DNA replication, has been shown to bind both oligonucleotides and peptides through its six OB-fold domains (Daughdrill, Buchko et al. 2003; Lee, Park et al. 2003; Stauffer and Chazin 2004; Bochkareva, Kaustov et al. 2005; Xu, Vaithiyalingam et al. 2008). Mcm10 appears to exhibit similar behavior by binding to DNA, pol α, DDK and PCNA (Lee, Seo et al. 2003; Das-Bradoo, Ricke et al. 2006; Warren 2008), although the role of the OB-fold in Mcm10 interactions with DDK and PCNA remains to be determined.

A molecular mechanism for Mcm10 hand-off of pol α to DNA

This is the first report of competition between DNA and pol α for binding to Mcm10. Competition for sites provides a ready mechanism for direct coupling of the protein interaction with DNA binding as a means to promote progression of the replication machinery. Although the exact role of Mcm10 in replication initiation has yet to be elucidated, it is reasonable to envision Mcm10 as a macromolecular recruiting and/or scaffolding protein due to the fact that Mcm10 contains two domains that can bind to DNA and pol α (Robertson, Warren et al. 2008). This follows a common strategy for numerous modular proteins involved in DNA processing; there is a significant kinetic advantage to deconstructing protein interactions into two or more weak binding sites (Stauffer and Chazin 2004). The recruitment of pol α to origins of replication by Mcm10

would be a significant step to signal nascent DNA synthesis and contribute to fork stability (Ricke and Bielinsky 2004; Yang, Gregan et al. 2005; Das-Bradoo, Ricke et al. 2006; Pacek, Tutter et al. 2006; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007; Zhu, Ukomadu et al. 2007; Robertson, Warren et al. 2008). Indeed, Mcm10 has been shown to be necessary for pol α loading onto chromatin (Wohlschlegel, Dhar et al. 2002).

The detailed analysis of binding affinities and competition experiments presented here demonstrate that the highly conserved Mcm10-ID transitions between interaction with DNA and pol α, consistent with an Mcm10-mediated hand-off mechanism (Figure 26). The relatively similar affinities of $p180^{189-323}$ and ssDNA for Mcm10-ID+CTD suggest that full-length Mcm10 also partitions between DNA and pol α binding. Two scenarios for hand-off of pol α onto DNA by a single Mcm10 molecule can be envisioned, the first in which the CTD interacts with ssDNA while the ID engages $p180^{286-310}$ (Figure 26A). Equivalently, Mcm10 could bind to the DNA through the ID while the CTD tethers the N-terminal region of p180 (Figure 26B). It is interesting to note that binding of $p180^{189-323}$ to CTD alone was undetectable by our fluorescence assay (data not shown), raising the possibilities that either the CTD binds to $p180^{1-323}$ outside of the 189-323 subdomain, that the ID and CTD bind p180 cooperatively, or that the CTD indirectly stimulates binding of Mcm10-ID to p180. Indeed, Mcm10-ID+CTD binds both ssDNA and $p180^{189-323}$ with 15-fold greater affinity than Mcm10-ID alone, suggesting that protein and DNA binding can be modified by domain interactions within Mcm10. However, anisotropy binding studies carried out with a mixture of Mcm10-ID and Mcm10-CTD did not enhance the binding affinity relative to either domain alone,

and thus far we have been unable to observe a direct interaction between the ID and CTD. Regardless, to recruit pol α, Mcm10 would have to partially release DNA. The slightly weaker affinity of Mcm10-ID for $p180^{189-323}$ than for ssDNA suggests that an external molecular trigger is necessary to shift Mcm10-ID binding from ssDNA to pol α. An alternate interpretation for the enhanced binding with the Mcm10-ID+CTD construct is that a second binding site on the CTD provides an extended interaction surface for DNA, which results in a synergistic effect on binding similar to that observed for the multiple OB-fold binding motifs in RPA (Arunkumar, Stauffer et al. 2003).

Mcm10 oligomerization provides yet a third mechanism for mediating DNA and p180 binding (Figure 26C). Mcm10 has been reported to form dimeric and hexameric assemblies (Lee, Seo et al. 2003; Okorokov, Waugh et al. 2007). We previously showed that Mcm10-NTD, which is predicted to contain a coiled-coil motif, forms a highly asymmetric dimer in solution (Robertson, Warren et al. 2008). Dimerization of Mcm10 through the NTD would expose multiple ID+CTD high-affinity binding platforms for binding to DNA and/or pol α. The higher affinity of the ID+CTD construct for both ssDNA and $p180^{189-323}$ suggests that this is the preferred binding mode over the individual domains. Importantly, the similar affinities of Mcm10-ID+CTD for pol α and DNA provides a physical basis for simultaneous binding of ssDNA and pol α by Mcm10. This condition also raises the possibility that a structural change would be necessary to facilitate Mcm10 release of pol α during a molecular hand-off to DNA. Previous studies suggest that phosphorylation (Izumi, Yatagai et al. 2001) or ubiquitination (Das-Bradoo, Ricke et al. 2006) are likely candidates for altering Mcm10 binding affinities. Additional studies beyond the scope of this paper, including elucidating the structure of full-length

Mcm10 and determining interaction partners, are required to fully understand how Mcm10 mediates critical interactions at the eukaryotic replication fork.

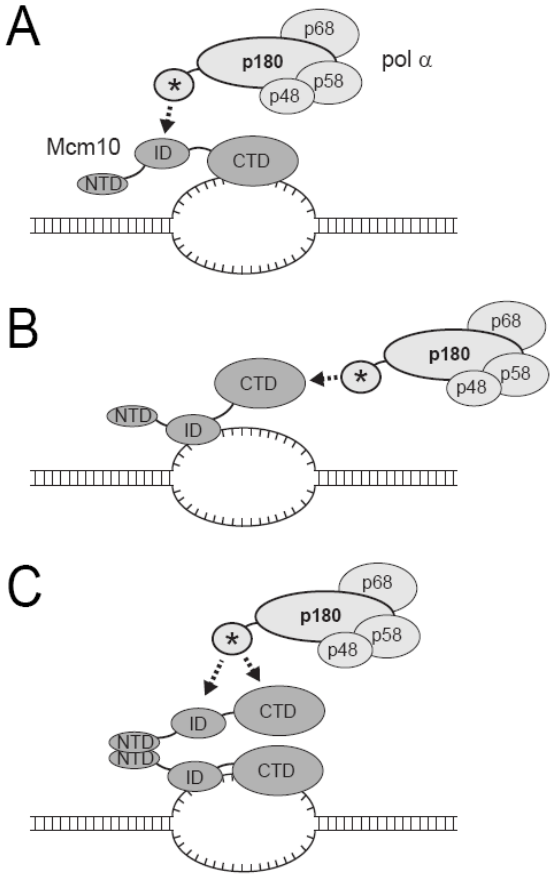


Figure 26. Three possible modes of Mcm10 binding to ssDNA and pol α at a replication fork. Mcm10-ID (**A**) and Mcm10-CTD (**B**) form binary complexes with either DNA or the N-terminal region of p180 (labeled with an asterisk). **C.** The ID+CTD together bind DNA and p180 with higher affinity than either the ID or CTD alone. Oligomerization via the NTD could also allow the ID and CTD from one subunit to contact DNA while a second subunit recruits pol α.

Experimental Procedures

Protein Expression and Purification

Mcm10-ID was prepared as described previously (Warren 2008). Briefly, the protein was overexpressed from a modified pET-32a vector (Novagen) in *E. coli* BL21(DE3) cells for 16 hours at 16°C and isolated using nickel affinity chromatography. After cleavage of the thioredoxin-His_6 tag, Mcm10-ID was purified using ssDNA affinity

and size exclusion chromatography. An Mcm10 construct spanning amino acid residues 230-860 (Mcm10-ID+CTD) was cloned and expressed similarly to Mcm10-ID except protein expression was induced at 21°C for 4 hours. Mcm10-ID+CTD protein was purified by Ni-NTA affinity chromatography (Qiagen), followed by S-sepharose (GE Healthcare) ion exchange chromatography and cleavage of the affinity tag. The cleaved protein was further purified by gel filtration using a Superdex 200 preparative column (GE Healthcare) equilibrated in 20 mM Tris pH 7.5, 150 mM NaCl, 3.5 mM β-mercaptoethanol, and 5% glycerol (Buffer A).

The DNA encoding amino acids 189-323 of human p180 was ligated into a modified pET-27 vector (Novagen) to produce an N-terminal His_6-fusion protein (pBG100, Vanderbilt Center for Structural Biology). *E. coli* BL21(DE3) cells transformed with the $p180^{189-323}$/pBG100 plasmid were grown at 37°C in LB medium containing 100 μg/ml ampicillin, and protein was overexpressed by the addition of 0.5 mM IPTG for 4 hours. For NMR experiments, protein was uniformly enriched with ^{15}N by propagating cells in M9 minimal medium supplemented with 1 mg/ml $^{15}NH_4Cl$ (Cambridge Isotope Laboratories) as the sole source of nitrogen. The cells were harvested in 50 mM Tris pH 7.5, 500 mM NaCl, and 10% glycerol, and lysed under pressure (25,000 psi) using an Avestin EmulsiFlex C3 homogenizer. $p180^{189-323}$ was purified by Ni-NTA affinity chromatography (Qiagen), followed by cleavage of the affinity tag. The cleaved protein was further purified by Q-sepharose (GE Healthcare) ion exchange chromatography, followed by gel filtration using a Superdex 200 preparative column (GE Healthcare) equilibrated in Buffer A. $p180^{243-256}$ and $p180^{286-310}$

peptides used for NMR titrations were synthesized and purified by Genescript Corp. (Piscataway, NJ).

X-ray Crystallography

Crystals were grown by sitting drop vapor diffusion by mixing 2 μl protein/DNA solution containing 300 μM Mcm10-ID and 360 μM dC_9 ssDNA with 2 μl reservoir solution containing 100 mM TAPS pH 9.0 and 17% PEG 3350. Crystals appeared overnight and grew to approximately 50 x 50 x 200 μm^3 after 2-3 days. Crystals were soaked 5 min in mother liquor containing 10% (v/v) butanediol and flash frozen in liquid nitrogen. Preliminary X-ray diffraction data (Table 1) were collected at beamline 21-ID at the Advanced Photon Source (Argonne, IL) and processed with HKL2000 (Otwinowski and Minor 1997). The Mcm10-ID/ssDNA complex crystallized in space group $P3_121$ with one molecule in the asymmetric unit.

X-ray phases were obtained by molecular replacement using unliganded Mcm10-ID (PDB ID 3EBE) as the search model in the program Molrep (Vagin and Teplyakov 1997). A clear rotation/translation solution was verified by the quality of the resulting composite annealed $2F_o$-F_c omit electron density maps generated using CNS (Brunger, Adams et al. 1998) (Supplementary Figure C1). Several iterative rounds of restrained atomic and temperature factor refinement against a maximum likelihood crystallographic target in Phenix (Afonine, Grosse-Kunstleve et al. 2005), together with manual model adjustment and building the L12 loop in Coot (Emsley and Cowtan 2004), resulted in R and R_{free} values of 21.7% and 26.1%, respectively. Strong F_o-F_c difference Fourier density was observed within the OB-fold cleft (residues 292-360). Three nucleotides of ssDNA were fit into this density using Coot and refined, which lowered R and R_{free} by

0.8% and 1.61%, respectively. The polarity of the DNA was established by parallel refinement of both orientations of DNA, which differed from each other by 1% in R_{free}. Translation/libration/screw (TLS) refinement was used to model anisotropic motion of 6 groups, defined by protein residues 235-241, 242-260, 261-300, 301-371, 372-500, and DNA as determined by the TLSMD server (http://skuld.bmsc.washington.edu/~tlsmd/) (Painter and Merritt 2006). Individual anisotropic B-factors were derived from the refined TLS parameters and held fixed during subsequent rounds of refinement, which resulted in a noticeable improvement of the electron density maps (Supplementary Figure C1) and a 1% decrease in both R and R_{free}. No additional electron density was discerned corresponding to the six remaining nucleotides or for residues 230-231 and 416-427 at the extreme N- and C-termini.

Analysis of the final structure using PROCHECK (Laskowski, MacArthur et al. 1993) showed 80.5% and 19.5% of the total of 154 non-glycine and non-proline residues to be within the most favored and allowed regions of the Ramachandran plot, respectively, with no residues in the disallowed region. The coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under accession number 3H15.

NMR Spectroscopy. Gradient enhanced ^{15}N-^{1}H HSQC NMR spectra were recorded at 25 °C using a Bruker DRX 800 NMR spectrometer equipped with single axis z-gradient cryoprobe. All spectra were acquired with 1024 complex points over a sweep width of 15 ppm in the ^{1}H dimension and 128 complex points over 37 ppm in the ^{15}N dimension. The center of the ^{15}N spectral width was set to 117.5 ppm, and the ^{1}H carrier was placed on the water signal at 4.7 ppm from the respective base spectrometer

frequencies. All spectra were processed and analyzed using Topspin v1.3 (Bruker, Billerica, MA) and Sparky v3.1 (University of California, San Francisco, CA). Data were treated with shifted sine-bell functions and zero-filled to twice the number of data points in both dimensions.

Chemical shift perturbation data were collected by titrating 1.8 mM unlabeled $p180^{189-323}$ into 400 μM ^{15}N-Mcm10-ID in 20 mM Tris-d_{11} (pH 7.0), 75 mM NaCl and 5% D_2O. Additionally, unlabeled Mcm10-ID was titrated into 400 μM ^{15}N-$p180^{189-323}$ in the same buffer. Spectra were recorded at Mcm10-ID:$p180^{189-323}$ ratios of 1:0, 1:0.25, 1:0.63, 1:1, and 1:2. The 15mer oligonucleotide d(GGCGCATTGTCGCAA) was used for ssDNA titrations. The observation of chemical shift perturbations in the fast-to-intermediate-exchange regime (on the NMR time scale) enabled most of the peaks in the free protein and complexes to be correlated. In cases where the peaks disappeared, the last titration point where the peak was observed was used for determining chemical shift perturbations. The magnitudes of the average chemical shift (δ) perturbations shown in Figures 22B, 24B, and S8A were calculated from contributions of both ^{15}N and ^{1}H dimensions in the HSQC spectrum by using the equation, $\Delta\delta_{ave} = (((\Delta\delta_{1H})^2 + (\Delta\delta_{15N}/5)^2)/2)^{1/2}$.

Fluorescence Anisotropy

Mcm10-ID and Mcm10-ID+CTD binding to $p180^{189-323}$ was measured by following an increase in fluorescence anisotropy as Mcm10-ID or Mcm10-ID+CTD was added to fluorescein labeled $p180^{189-323}$ or $p180^{286-310}$. Fluorescein-$p180^{189-323}$ was prepared by incubating purified $p180^{189-323}$ with a 20-fold molar excess of MTS-

fluorescein (2-[(5-fluoreceinyl) aminocarbonyl]ethyl methanethiosulfunate, Toronto Research Chemicals) at 25 ºC for 6 hours, followed by purification on a 1-ml Q-Sepharose column (GE Healthcare). Fluorescein isothiocyanate (FITC) conjugated $p180^{286-310}$ was synthesized and purified by Genescript Corp. (Piscataway, NJ). For anisotropy measurements, unlabeled protein was added over the concentration range of 0.1-50 μM to a solution containing 50 nM fluorescein-p180 in 20 mM Tris pH 7.5, 100 mM NaCl, and 5% glycerol. Polarized fluorescence intensities were measured at excitation and emission wavelengths of 495 and 538 nm, respectively. Dissociation constants (K_d) were derived by fitting a single-state binding model to data from three experiments using Kaleidagraph 3.6 (Synergy Software) according to the equation $r = r_o + (r_{max}\text{-}r_o)([E]/(K_d+[E]))$, where r is the fluorescence anisotropy, r_o and r_{max} are the anisotropy values of unbound and fully bound DNA, respectively, and [E] is the total concentration of protein.

Isothermal Titration Calorimetry (ITC)

Proteins were buffer exchanged into 25 mM Tris (pH 7.5) and 100 mM NaCl and concentrated to 50 μM (Mcm10-ID) and 1 mM ($p180^{189-323}$). 1.7 ml Mcm10-ID was placed in the sample cell, into which $p180^{189-323}$ was injected in 6 μl steps during the run. Data were collected at 25 ºC using a MicroCal VP-ITC and analyzed using the accompanying Origin software (Origin Lab, Northampton, MA). Thermodynamic parameters were calculated from fitting the data to the best binding model using Origin according to the Gibbs free energy equation, $\Delta G = \Delta H - T\Delta S = -RTlnK_a$.

Footnotes

The authors gratefully acknowledge Sivaraja Vaithiyalingam for his assistance with NMR experiments, and thank the staff at beamlines 21-ID at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) and the RapiData and X29A staff at the National Synchrotron Light Source (Brookhaven National Laboratory, Upton, NY). Use of the Advanced Photon Source was supported by the U.S. Department of Energy Office of Basic Energy Sciences under Contract No. DE-AC02-06CH11357. Use of LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor for the support of this research program (Grant 085P1000817). Additional thanks to Haijiang Zhang for preliminary Mcm10-p180$^{189\text{-}323}$ binding data. This work was funded by National Institutes of Health Grants R01 GM080570 (to B.F.E.), R01 GM65484 (to W.J.C.) and, and R01 GM52948 (to E.F.). E.M.W. was supported in part by the Molecular Biophysics Training Grant (T32 GM08320). Additional support for crystallography and mass spectrometry facilities were provided by the Vanderbilt Center in Molecular Toxicology (P30 ES000267) and the Vanderbilt-Ingram Cancer Center (P30 CA068485).

CHAPTER V

DISCUSSION AND FUTURE DIRECTIONS

Ever since the structure of DNA was first discovered, scientists have tried to understand the complex mechanism of DNA replication. Many different approaches were used in many different systems and model organisms. As more proteins were discovered, the focus began to shift towards the interactions of these proteins with one another and the layers of regulation that occur via these interactions. Undertsanding has been greatly accelerated by visualizing the interactions that make up the replication machinery at atomic level resolution. As structural biology techniques have advanced in the past decade, renewed interest in the field of DNA replication has arisen.

The focus of recent research has been on initiation proteins such as Mcm2-7, pol α, GINS, Mcm10, Cdc45, and Dpb11/TopBP1, and many of these proteins are multi-domain and multi-subunit factors. Due to the emerging complexity of the eukaryotic protein machinery at origins of replication, a combination of several biochemical, structural, and computational approaches have become necessary. Many of these approaches have been utilized in this thesis, and a few not mentioned previously will be discussed here. As the layers of complexity are slowly peeled away, many questions will be answered that will facilitate our understanding of the intricacies of the process of DNA replication.

The work presented in this thesis contains the most thorough and definitive characterization of the structural architecture of Mcm10 to date. While Mcm10 has been shown to be essential for replication *in vivo* (Nasmyth and Nurse 1981; Dumas, Lussky et al. 1982; Maine, Sinha et al. 1984), its function at origins has remained elusive. Determining the structure of the complete Mcm10 protein or of each domain will facilitate structure-function analysis that will expand our understanding of the biological role of Mcm10. This thesis presents a first step in this direction several data were collected and are presented here, including domain characterization, crystallization of unliganded and ssDNA-bound Mcm10-ID, and determination of the interactions with DNA polymerase α.

Crystallization of xMcm10-ID with dsDNA

While the structure of xMcm10-ID bound to ssDNA (presented in chapter 4) provided insight into the manner in which this domain interacts with DNA, there is also ample data demonstrating that Mcm10-ID binds to dsDNA with an affinity roughly half of the ssDNA couterpart (chapter 2). To understand the Mcm10-dsDNA interaction, co-crystallization trials were performed with xMcm10-ID and dsDNA at a 1:1.2 molar ratio. DNA sequences used in these trials are shown in Table 7. All dsDNA substrates were blunt ended and thus the complementary strands are omitted for simplicity. Crystals with dsDNA 5-10 nucleotides in length were obtained in several conditions and ranged from bundles of tiny needles to polygon blocks. Samples of Mcm10-ID with ds5mer and ds15mer were subjected to crystallization trails at the Hauptman-Woodward High Throughput screening (HWI-HTS) facility in Buffalo, NY (Luft, Collins et al. 2003).

Several conditions yielded crystals with the 5mer DNA, but none with the 15mer (Figure 27). Optimization of these conditions was attempted by varying DNA length, buffers, pH, salts, and precipitates, but could not be optimized further to yield larger crystals.

Table 7. DNAs used for crystallization	
Oligo Name	Oligo Sequence (5' to 3')
ds4mer	ATGG
ds5mer	ATGGT
ds6mer	ATGGTA
ds7mer	ATGGTAG
ds8mer	ATGGTAGG
ds9mer	ATGGTAGGC
ds10mer	ATGGTAGGCA
ds12mer	ATGGTAGGCAAC
ds14mer	ATGGTAGGCAACCA
ds15mer	ATGGTAGGCAACCAT
ds20mer	ATGGTAGGCAACCATGTAGT

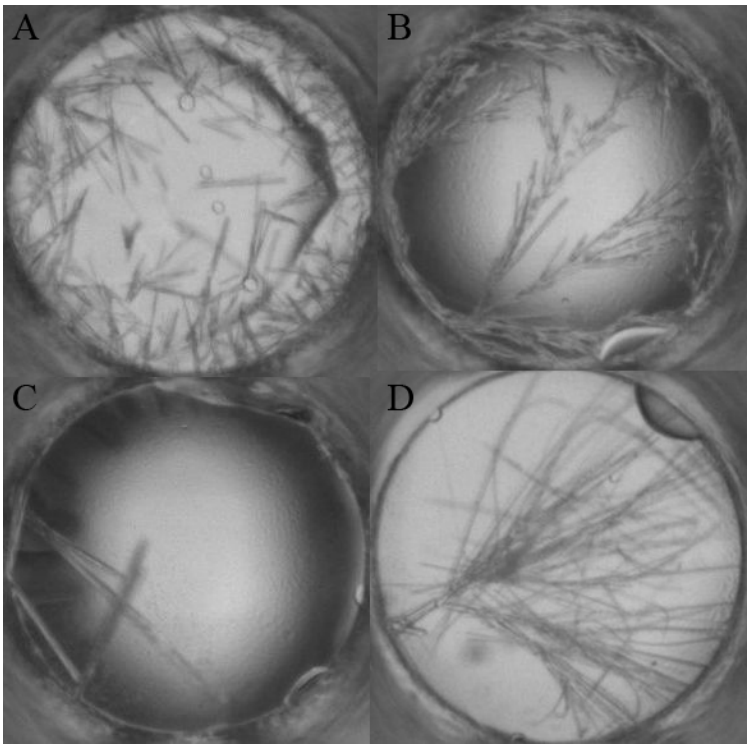


Figure 27. Co-crystals of xMcm10-ID and ds5mer. Crystals obtained from HWI-HTS with a mixture of 500µM xMcm10-ID and 600µM ds5mer. Conditions of these crystals are: (**A**), 0.1M CAPS pH 10.0, 0.1M NH_4Cl, and 40% PEG 4000, (**B**), 0.1M HEPES pH 7.0 and 30% PEG 6000, (**C**), 0.1M Bis-tris pH 6.5 and 28% PEG MME 2000, (**D**), 0.15M KBr and 30% PEG MME 2000.

In addition to screening at HWI-HTS, crystallization trials were pursued with in-house screens. Crystals obtained were optimized by varying DNA length, and the buffer components (Figure 28A,B), which yielded crystals that diffracted to ~4.2 Å (Figure 28C) on a laboratory X-ray source. The data were processed with HKL2000 (Otwinowski and Minor 1997) and initial phases were obtained with molecular replacement using unliganded Mcm10-ID as a search model in the program Phaser (Storoni, McCoy et al. 2004). The molecular replacement model and phases were refined in REFMAC (Collaborative Computational Project 1994), and composite omit electron density maps were generated in CNS (Brunger, Adams et al. 1998). The protein could be readily identified in the electron density, although no density corresponding to DNA could be discerned.

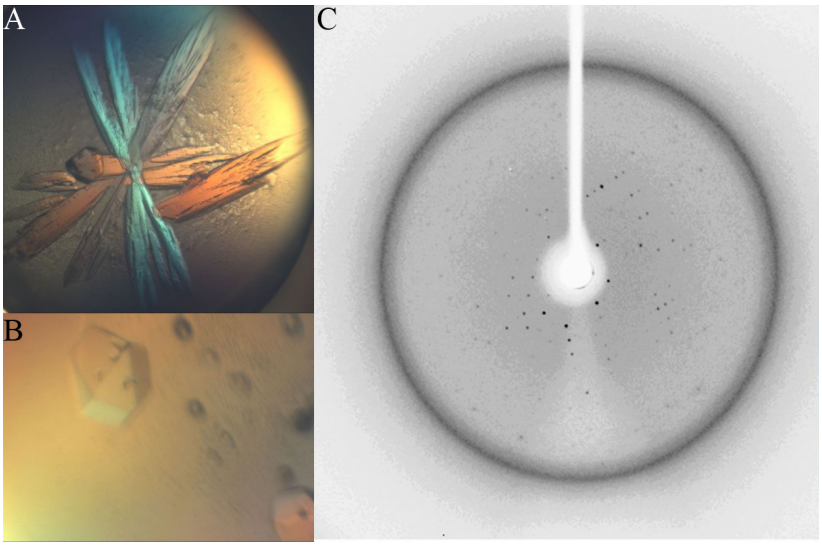


Figure 28. Crystallization and diffraction of xMcm10-ID/ds5mer co-crystals. Crystals obtained with a mixture of 500µM xMcm10-ID and 600µM ds5mer from a PEG vs. pH screen. (**A**), 0.1M HEPES pH 7.5 and 14% PEG 8000, (**B**), 0.1M HEPES pH 6.8 and 12% PEG 8000. (**C**), A 120 second exposure on the Proteum PT135 CCD using a Bruker Microstar X-ray generator.

The atomic model, which consists of two Mcm10-ID molecules in the asymmetric unit (ASU), was refined to 5.0 Å resolution to a crystallographic residual of 0.237 (R_{free} =

0.330) (Table 8). The two protomers in the ASU are arranged in a head-to-tail orientation such that the OB-folds are facing one another while the zinc fingers are on opposite sides. Loops L12 and L23 from each protomer pack against one another, while the zinc finger loop of each protomer packs against the N-terminal helix on the back side of the opposite protomer (Figure 29A). Interestingly, electron density corresponding to the degraded C-terminus was discovered. Building in this 10-residue peptide resulted in a 0.3% and 2.0% decrease in the R_{work} and R_{free}, respectively. The electron density for this C-terminal peptide was only found next to one protomer. The peptide packed against the N-terminal helix, which is adjacent to the zinc-finger from the second protomer in the ASU. The model built for the peptide is shown as yellow sticks in Figure 29A.

Table 8 Data collection and refinement statistics

	dsDNA Crystal
Data collection	
Space group	$P3_1$
Cell dimensions	
a, *b*, *c* (Å)	96.131, 96.131, 63.844
α, β, γ (°)	90.0, 90.0, 120.0
Resolution (Å)	50.0-4.2 (4.35-4.2)
R_{sym} or R_{merge}	17.1 (47.1)
$I / \sigma I$	8.17 (2.62)
Completeness (%)	99.0 (98.9)
Redundancy	10.4 (9.0)
Refinement	
Resolution (Å)	5.0
No. reflections	2824
R_{work} / R_{free}	23.77 / 33.00
No. atoms	
Protein	2924
Ligand/ion	2
Water	0
B-factors	
Protein	53.63
Ligand/ion	41.76
Water	0
R.m.s. deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.634

Values in parentheses are for highest-resolution shell.

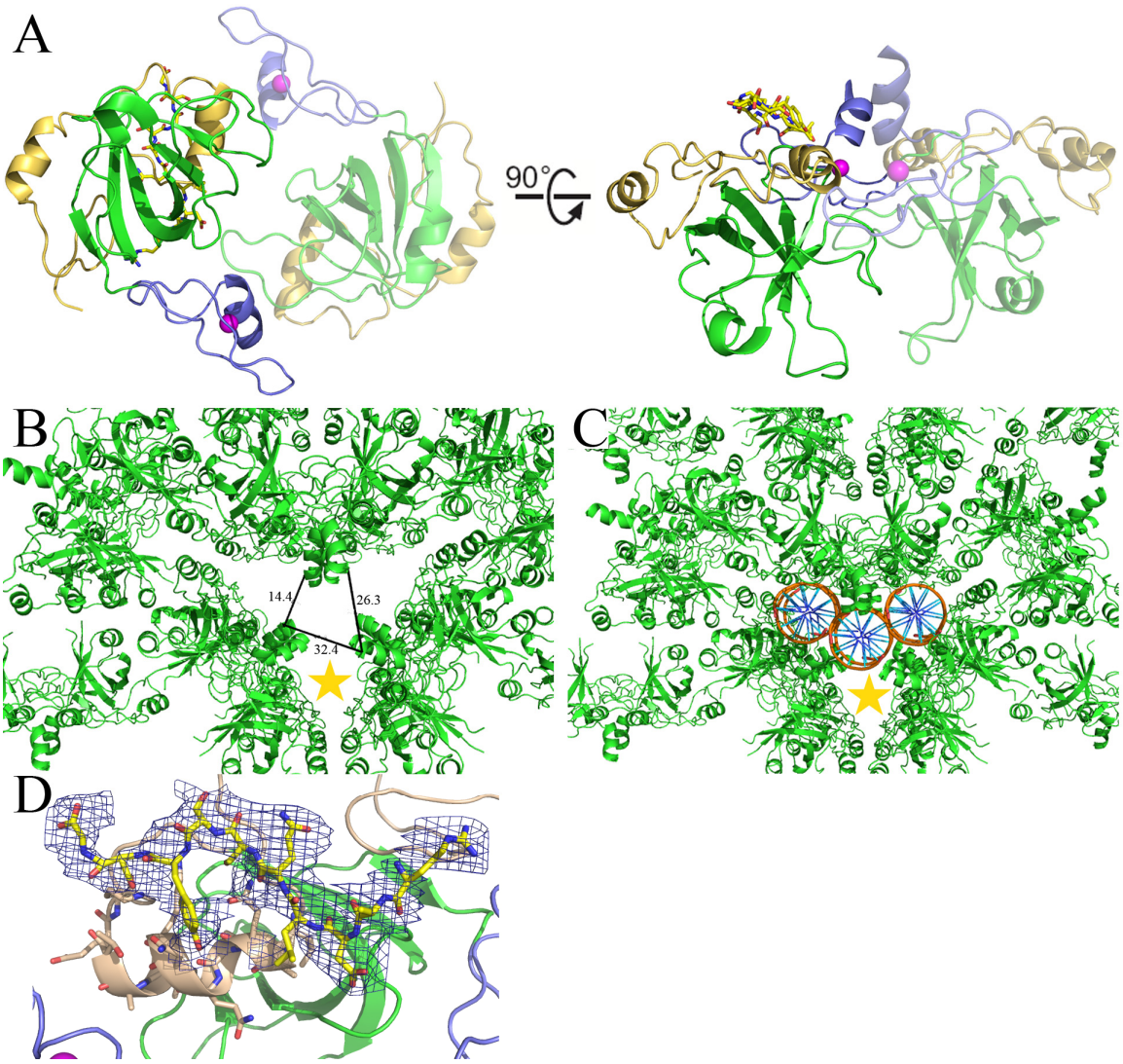


Figure 29. Crystal structure and packing of a $P3_1$ form of xMcm10-ID. (**A**), xMcm10-ID was crystallized in the presence of dsDNA in the $P3_1$ space group and is rendered as a ribbon diagram. Two orthogonal views show the relative orientation of the protomers with respect to one another. The OB-fold is colored green, the zinc finger blue with a magenta Zn^{2+} sphere, the N-terminal α-helical/coil region is colored gold, and the C-terminal peptide is rendered as yellow sticks. (**B**), A view of the crystal lattice packing with protein shown as green ribbons. One of the areas of low density occurs on a symmetry axis and the perimeter is measured (14.4Å x 26.3Å x 32.4Å). (**C**), The same view as in **B**, but three 10 nucleotide ideal B-DNA molecules have been placed in the areas of weak density around the symmetry axis. The area below the three DNA molecules, denoted by a star, was not large enough to model a B-DNA molecule without significant clashes. (**D**), Composite omit electron density corresponding to the C-terminal peptide.

Although no electron density for DNA was observed, due to the crystal packing there are three cavities along a symmetry axis with sufficient room to accommodate dsDNA (Figure 29B,C). Additionally, it has been reported in the literature that when

proteins bind DNA in a non-sequence specific way, there is not always complete electron density for the DNA (Brownlie, Ceska et al. 1997; Bell and Lewis 2001; Shaw, Tempel et al. 2008). It has been hypothesized that this lack of DNA density is due to static disorder in the DNA. That is, the DNA binds in multiple registers along the protein surface, which blurs and weakens the electron density for the DNA.

More recently, a similar crystal packing has been observed in which DNA was necessary for crystallization but no DNA was located in the electron density. Thus, the DNA binding protein was hypothesized to bind dsDNA that would be positioned along symmetry axes (Bowles, Metz et al. 2008). This evidence suggests that it is possible for the structures shown in figure 29A to accommodate a DNA duplex generated by 2-fold crystallographic symmetry. Figure 30A shows the locations of the dsDNA molecules in the ASU, if the DNA were present in the crystal lattice as depicted in figure 29B,C. One DNA molecule is not close enough to make any contacts with xMcm10-ID except at the tips of the Zn-finger recognition helix. The other two molecules occupy the space on either side of the Zn-finger recognition helix, within interaction distance of the extended Zn-finger loop (yellow dsDNA) or the N-terminal helix on the backside of the OB-fold (purple dsDNA).

In order to provide important insight into the manner in which Mcm10 interacts with origin DNA during initiation, a comparison of known OB-folds bound to dsDNA was performed. Other OB-fold containing proteins that have been crystallized bound to dsDNA and dsRNA include tRNA synthetases, ribosomal proteins, and helicases (Cavarelli, Eriani et al. 1994; Eiler, Dock-Bregeon et al. 1999; Wimberly, Brodersen et al. 2000; Klein, Schmeing et al. 2001; Singleton, Scaife et al. 2001). These structures

were examined and superimposed on one protomer from the crystal structure presented here. In almost all cases, the dsDNA or dsRNA was distorted or unwound, positioning the single-stranded segment in the OB-fold cleft (Figure 30B). Comparison of the modeled DNAs from these crystal structures with the predicted binding mode from the $P3_1$ crystal structure shows no similarities, suggesting that either Mcm10-ID has a novel dsDNA binding mode, or that there was no dsDNA present in the crystals and that it only aided in the crystallization of the protein itself.

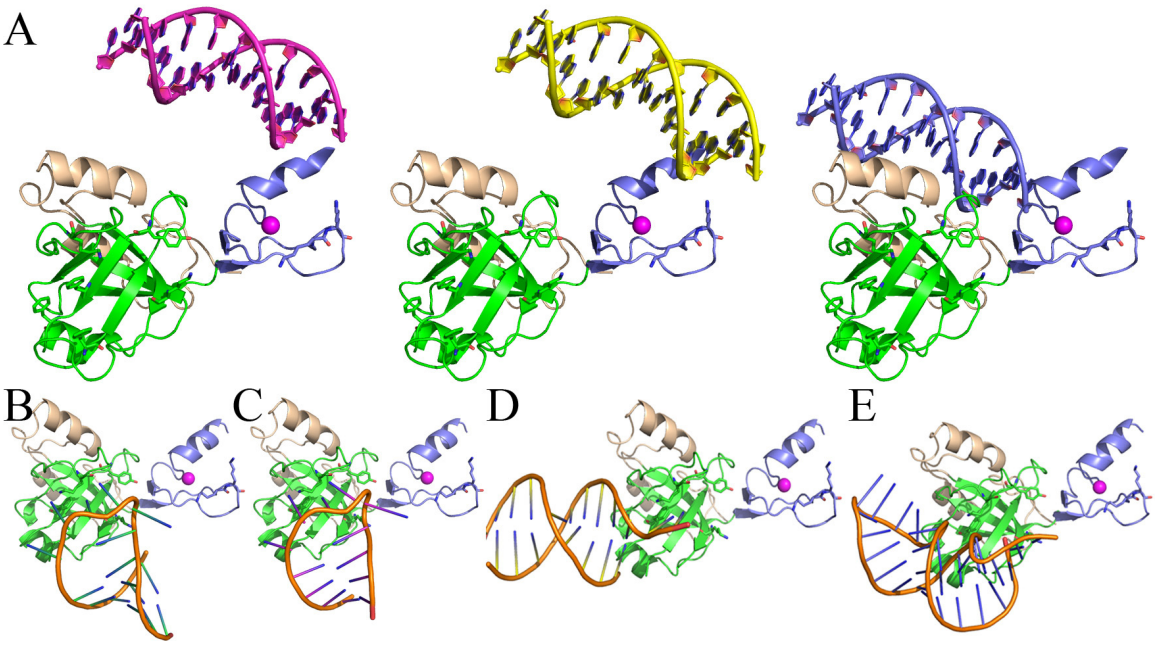


Figure 30. dsDNA bound to OB-folds. (**A**), Predicted modes of dsDNA binding by the $P3_1$ crystal form of xMcm10-ID based on the unoccupied space along the crystallography symmetry axis. (**B-E**), The OB-fold of xMcm10-ID superimposed onto the OB-fold of other crystal structures bound to dsDNA or dsRNA to show the location of these molecules with respect to the OB-fold. (**B**), Yeast aspartyl-tRNA synthetase, PDB ID: 1ASZ (Cavarelli, Eriani et al. 1994); (**C**), *E. coli* aspartyl-tRNA synthetase, 1COA (Eiler, Dock-Bregeon et al. 1999); (**D**), RecG, 1GM5 (Singleton, Scaife et al. 2001); and (**E**), 30S ribosomal S17 protein 1J5E (Wimberly, Brodersen et al. 2000).

The dsDNA modeling data suggest that Mcm10 could be involved in distorting or denaturing the dsDNA to expose the ssDNA needed for the helicase to begin unwinding the origin to form functional replication forks. This idea complements the

inconsistencies in dsDNA binding constants obtained by fluorescence anisotropy experiments (data not shown). Although salt was able to affect DNA binding, due to the large electrostatic component of xMcm10-ID's interaction with DNA (chapter 4), it is possible that melting of the duplex might also contribute to the variation of binding constants calculated in the experiments. Further investigation of this interaction needs to be pursued in order to shed more light onto this interaction.

PCNA Docking onto xMcm10-ID

It was recently demonstrated that scMcm10 is di-monoubiquitinated, and that Ub-Mcm10 is the only form of the protein that interacts with PCNA (Das-Bradoo, Ricke et al. 2006). These results confirmed reports of Mcm10 being modified in a cell-cycle dependent manner (including ubiquitination during late G1 and S phases) (Izumi, Yatagai et al. 2001). Using yeast-two-hybrid and pull-down experiments, Das-Bradoo and colleagues demonstrated an interaction between scPCNA residues Y133 and A251 with scMcm10 residues L242 and Y245, respectively. These residues align with xMcm10-ID I321 and F324, and thus this interaction was further investigated.

To first determine whether the unmodified domain would interact, affinity chromatography pull-down assays were attempted but were unsuccessful (data not shown). In the absence of biochemical interaction data, this interaction was probed *in silico* using docking programs DOT (Eyck, Mandell et al. 1995), Hex (Ritchie and Kemp 2000), ZDOCK (Chen and Weng 2002), and RosettaDock (Gray, Moughon et al. 2003). Models from these programs were compared for lowest energy function and visually inspected for most reasonable model. The corresponding best 10 models were compared

to crystal structures of known proteins bound to PCNA (Figure 31A,B). DOT and RosettaDock return the best docked models and were thus used for the rest of this study.

Interestingly, two models placed I321 and F324 near the interdomain connector loop (IDCL) and C-terminal tail, commonly involved in protein-protein interactions (containing the Y133 and A251 residues, respectively). It is interesting to note that, in these two models, the C-terminus of β1 within the OB-fold of Mcm10-ID lies in the exact same position as many of the PIP box containing ligands (Figure 31C). In this orientation, F324 is 10.3Å away from A252 of PCNA and I321 is 17.1Å away from Y133 of PCNA. Although these distances aren't meaningful in terms of direct interactions, it is possible that a di-monoubiquitinated form of Mcm10 might be structurally rearranged such that these two residues can be closer to PCNA. This interaction between modified PCNA and Mcm10 is very intriguing and paves the way for research into this interaction.

Recent studies have demonstrated that the interaction between pol α and Mcm10 is important to normal cell cycle progression (Fien, Cho et al. 2004; Ricke and Bielinsky 2004; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007). Another study has shown, in budding yeast, that the interaction between PCNA and Mcm10 is also important for normal cellular growth (Das-Bradoo, Ricke et al. 2006). The central role of Mcm10 in these studies suggests that Mcm10 could function to recruit pol α to DNA but also to juggle it on the lagging strand as pol δ is recruited for the polymerase switch (Nick McElhinny, Gordenin et al. 2008).

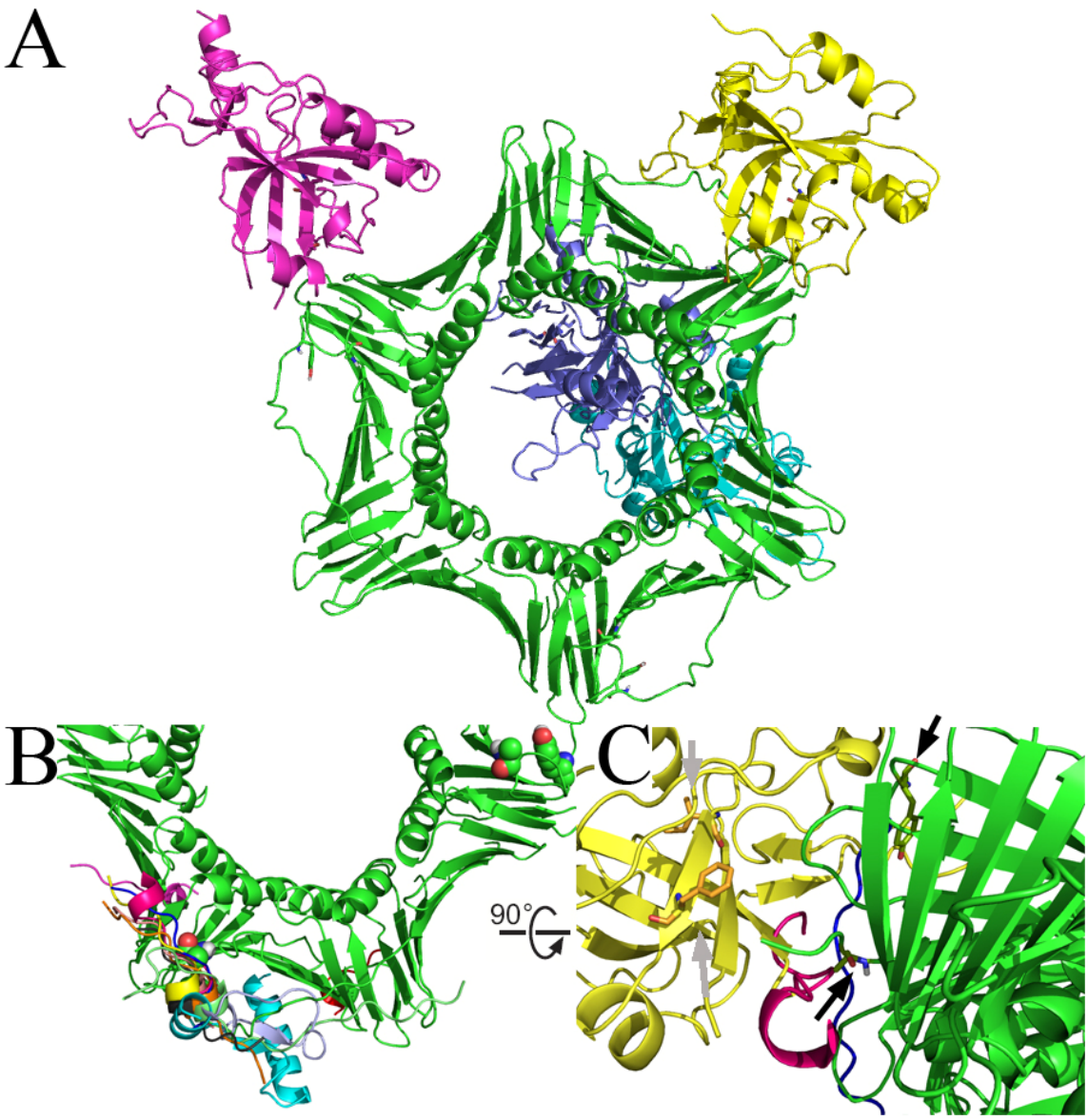


Figure 31. xMcm10-ID and PCNA docking. (**A**), Representative 4 of the best scoring docked models from Hex (yellow and magenta) (Ritchie and Kemp 2000) and RosettaDock (cyan and blue) (Gray, Moughon et al. 2003). (**B**), Overlay of crystal structures of ligands bound to PCNA (Pink; PDB ID 1U76 (Bruning and Shamoo 2004), Purple; PDB ID 1RXZ (Chapados, Hosfield et al. 2004), Blue; PDB ID 1RXM (Chapados, Hosfield et al. 2004), Yellow; PDB ID 2OD8 (Vijayakumar, Chapados et al. 2007), White; PDB ID 1YYP (Appleton, Brooks et al. 2006), Orange; PDB ID 1VYJ (Kontopidis, Wu et al. 2005), Black; PDB ID 1U7B (Bruning and Shamoo 2004), Light Green; PDB ID 1UL1 (Sakurai, Kitano et al. 2005), Cyan; PDB ID 1DML (Zuccola, Filman et al. 2000), Red; PDB ID 1ISQ (Matsumiya, Ishino et al. 2002), Magenta; PDB ID 2IZO (Dore, Kilkenny et al. 2006)). The residues that interact with scMcm10 (Y133 and A252) are shown as CPK models (Das-Bradoo, Ricke et al. 2006). (**C**), Image in **B** rotated forward 90º with all ligands removed except FEN-1 (magenta). xMcm10-ID is shown as a yellow cartoon with each PCNA interacting residue shown as sticks and denoted with grey arrows (F324 and I321). The interdomain connector loop of PCNA (containing Y133) is colored blue and the c-terminus (containing A252) is colored light green with each residue shown as sticks and denoted by black arrows.

Docking xMcm10-ID into EM Density

A recent publication has utilized negative stain electron microscopy (EM) to determine the structure and oligomeric state of full-length hMcm10 (Okorokov, Waugh et al. 2007). A 16Å electron density map demonstrated that hMcm10 forms a hexameric ring reminiscent of the helicase Mcm2-7 (Figure 31). Thus, the authors used secondary structure predictions to align hMcm10 with known helicase structures. The hMcm10-ID aligned with an archaeal MCM helicase, while the hMcm10-CTD aligned with SV40 T-antigen. Using this secondary structure alignment, the authors placed these crystal structures into their EM density (Figure 32A).

Since the crystal structure of xMcm10-ID is now known (chapters 3 and 4), the docking program Situs Colores (Wriggers, Milligan et al. 1999) was used to place this domain into the EM density maps. The 10 best models produced were docked into the lower tier of density and two were in the same lobe as the archaeal MCM (Figure 32A,B). The placement of xMcm10-ID into the EM density is interesting because the C-terminus points toward the upper tier of electron density and thus, the CTD would occupy the upper tier. However the upper tier is smaller in volume and it would be difficult for a domain of larger mass (30 kDa for the CTD compared to 24 kDa for the ID) to occupy such a small volume. Additionally, there is no reasonable density for the NTD, which is predicted to be a rod-like coiled-coil, if the upper tier is occupied by the CTD. Thus it is possible that the authors might have imaged density for another protein such as the prevalent GroEL protein. Another problematic feature of the EM density is that it is very porous (i.e. it contains many thin leaflets). While the overall shape resembles other

hexameric rings, the many compartments and pores do not resemble any known ring-shaped proteins.

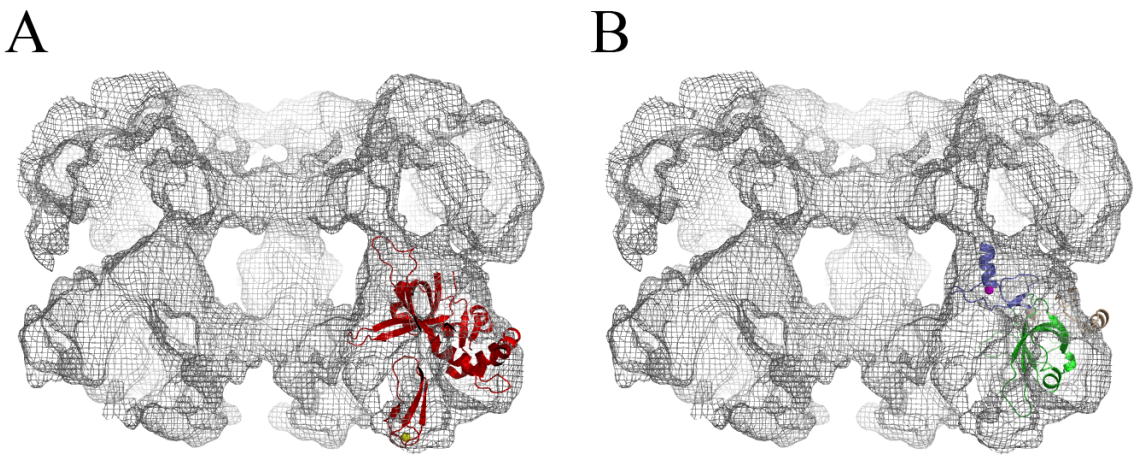


Figure 32. xMcm10-ID docked into 16Å EM density. (**A**), EM density map (Okorokov, Waugh et al. 2007) with archaeal MCM subunit (PDB ID 1LTL monomer) docked as in above paper. (**B**), Same view of EM density with xMcm10-ID docked with Situs Colores (Wriggers, Milligan et al. 1999) showing similar placement into the density map.

Mcm10 has been shown to oligomerize by several groups using several techniques (Cook, Kung et al. 2003; Fien and Hurwitz 2006; Okorokov, Waugh et al. 2007; Robertson, Warren et al. 2008). One such study suggested that Mcm10 forms a dimer in addition to higher molecular weight aggregates (Robertson, Warren et al. 2008). To address this disparity, we investigated the structure and oligomeric state of full-length xMcm10 by EM in collaboration with Melanie Ohi. Purified Mcm10 was adsorbed onto a glow-discharged carbon-coated copper grid, stained with uranyl formate (0.7% wt/vol), and analyzed. Other than the heptameric GroEL ring, no ring-shaped particles corresponding to hexameric Mcm10 were identified (Figure 33). The most noticeable particles were amorphous and much smaller than the GroEL ring, with a molecular weight estimate to be 100-200 kDa, consistent with a monomeric or dimeric form of xMcm10. Although this result agrees with the AUC data presented in chapter 2, further

work is needed in this area to determine the oligomeric state of Mcm10 and its significance in DNA replication. It is an attractive possibility that Mcm10 could dimerize, poising itself to interact with both leading and lagging strands.

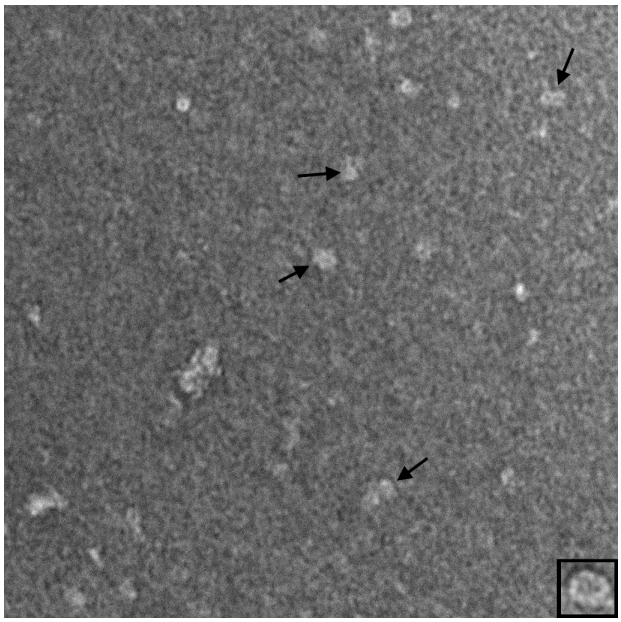


Figure 33. Electron micrograph of xMcm10. An electron micrograph of negatively stained Mcm10 taken with a Philips Tecnai T12 electron microscope at an acceleration voltage of 80 kV with a magnification of x44,000. A representative ring-shaped particle is shown in the inset and density believed to correspond to xMcm10 is labeled with arrows.

Ubiquitination of xMcm10

Using prediction servers such as Phyre (Kelley and Sternberg 2009) and Meta server (Ginalski, Elofsson et al. 2003), Mcm10-CTD shares sequence homology with zinc binding motifs from a family of proteins involved in the ubiquitination pathway. Known as RING fingers, these motifs are commonly found in E3 ubiquitin ligases such as human double minute 2 (HDM2) and coordinate two zinc ions through a C_3HC_4 or $C_2H_2C_4$ cassette (Kostic, Matt et al. 2006). A variant of this motif, the PHD finger for Plant Homeo Domain, is found in proteins that interact with modified histones (Pascual, Martinez-Yamout et al. 2000). In the absence of a structure of Mcm10-CTD, the possibility of Mcm10 acting as an E3 to auto-ubiquitinate itself in order to facilitate the

interaction with PCNA is quite intriguing. To test this hypothesis, full length xMcm10 was incubated in an *in vitro* ubiquitination reaction containing the recombinantly produced proteins E1, E2-UbcH5a, and ubiquitin (BostonBiochem). Briefly, 2.5 μM Mcm10 was added to 5 mM ATP, 52 nM E1, 600 nM E2-UbcH5a, and 50 μM ubiquitin in a 20 μL reaction volume. The reactions were incubated for 60 minutes in a 30 ºC water bath upon activation with ATP. The process was stopped with the addition of SDS-loading buffer and heating at 95 ºC for 15 minutes. The samples were then resolved on a 4-12% gradient SDS-PAGE gel and visualized by Coomassie staining.

The results of this experiment demonstrate that Mcm10 is not capable of auto-ubiquitination under these conditions (Figure 34). This does not, however, rule out the possibility that Mcm10 could ubiquitinate another target protein. Due to the plethora of E3 ligases and their numerous targets, it is difficult to determine the target protein of an E3. Additional experiments are needed to determine if a different E2 functions to facilitate Mcm10 auto-ubiquitination. This experiment could be repeated with several other commercially available E2 enzymes such as UbcH1-4, UbcH5b, UbcH5c, UbcH6-13, and Use1. An alternative approach is to map the endogenous ubiquitination sites on Mcm10 using mass spectrometry. Current work is underway to accomplish this (A. Bielinsky, personal communication).

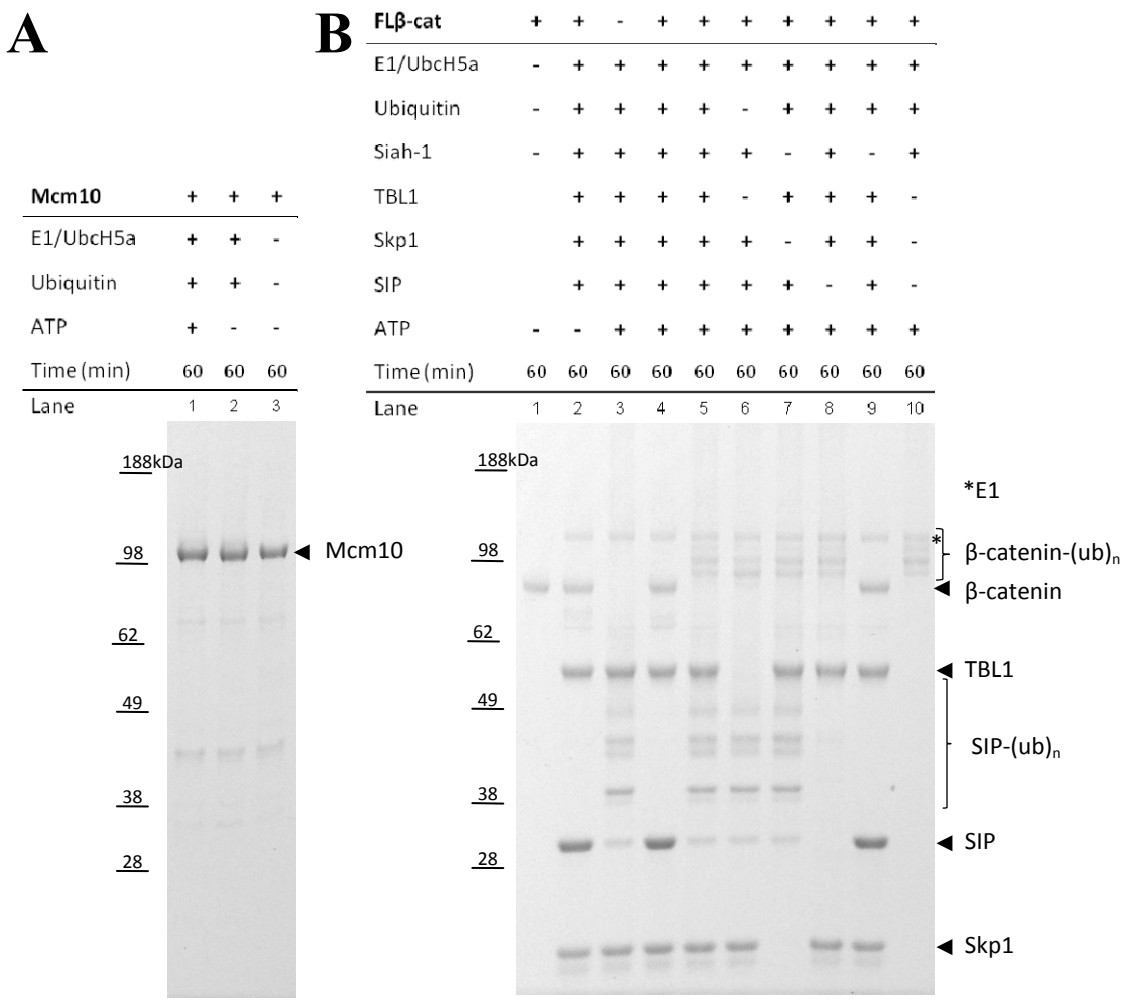


Figure 34. xMcm10 does not auto-ubiquitinate. (**A**), Incubation of Mcm10 (lanes 1-3) with ubiquitination pathway components E1 and E2-UbcH5a (lanes 1 and 2), ubiquitin (lanes 1 and 2), and ATP (lane 1) does not result in supershifted bands. (**B**), A control gel of known ubiquitinated proteins, SIP and β-catenin with and without necessary components of the ubiquitination pathway. Data courtesy of Yoana Dimitrova.

Preliminary Crystallization of xMcm10-ID-CTD

In order to examine the larger context in which Mcm10 functions, a construct containing both the ID and CTD was designed. This dual domain construct containing xMcm10-ID and –CTD was purified as described in chapter 4, mixed with a 1.2 fold molar excess of ssDNA (see 25mer described in chapter 2), and submitted to the HWI-HTS screening facility. Several hits were obtained at HWI-HTS, but none could be

reproduced or optimized using in-house vapor diffusion screens. Thus microbatch under oil trays were set up, mimicking the conditions used at HWI-HTS. Only one condition yielded a crystal: 100 mM Tris-HCl pH 8.5 and 25% (v/v) glycerol (Figure 35A). This crystal was captured in a cryo-loop, flash frozen, and examined for X-ray diffraction at a synchrotron source (APS, LS-CAT ID-F). Only solvent diffraction corresponding to ice was observed (Figure 35B). Further optimization will be needed to verify if this is a protein crystal and perhaps lead to a crystal structure of Mcm10-ID-CTD bound to ssDNA.

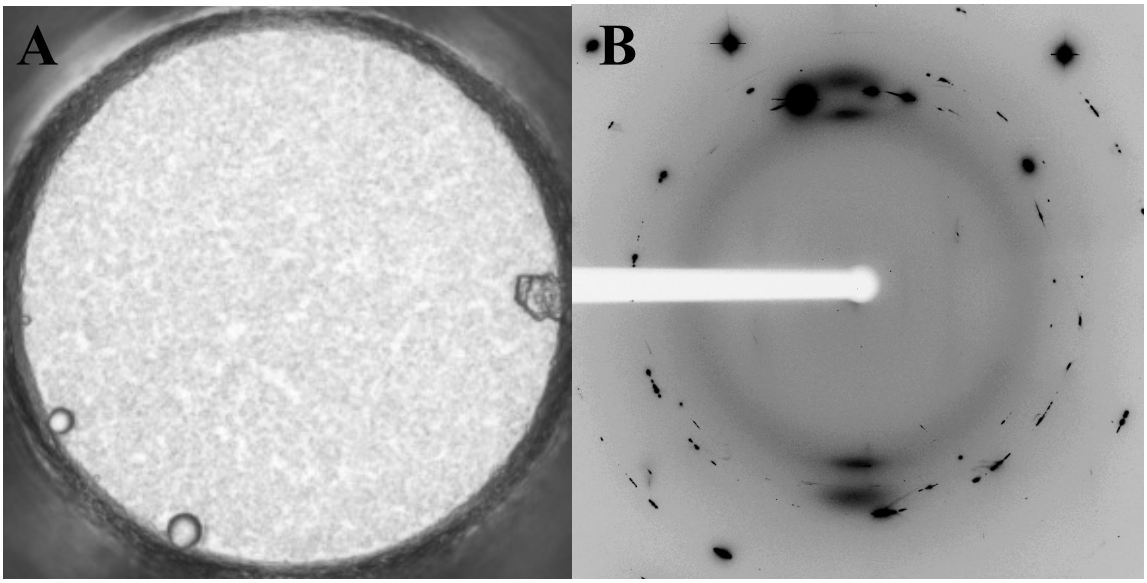


Figure 35. Crystallization and diffraction of Mcm10-ID-CTD/ssDNA crystal. (**A**), Image of a crystal obtained from HWI-HTS that was optimized by microbatch under oil. (**B**), The resulting diffraction pattern obtained from the crystal in A.

Possible Roles of Mcm10 in DNA Replication

The DNA binding activity of Mcm10, its putative role in the recruitment of pol α to chromatin, and its essential role in the activation of Mcm2-7 suggest Mcm10's function may directly manipulate DNA during unwinding. It was recently determined in *Xenopus* egg extracts that when the replicative helicase and polymerase machinery are

decoupled by the addition of aphidicolin, Mcm10 stays associated with the polymerase but not the helicase. This excludes Mcm10 from playing a role in DNA unwinding during fork progression and suggests that Mcm10 remains associated with the polymerases as a protein fold stabilizing factor (Pacek, Tutter et al. 2006). These data do not rule out a role for Mcm10 in origin melting prior to fork unwinding. Such origin melting is required for the recruitment of RPA and pol α, which have been shown to be dependent on the presence of Mcm10 and suggests that Mcm10 could play a role in origin melting. Another possible rationale for DNA binding by Mcm10 is to recruit downstream factors, such as pol α, directly onto DNA. It is apparent that a more thorough understanding of the Mcm10-DNA interaction is needed to clarify the significance of this function during the emergence of an active replisome.

The structure-function analyses presented here suggest that Mcm10 serves as recruitment and scaffolding protein at origins of replication, and the following model begins to emerge. Mcm10 is recruited to origins of replication at the onset of S-phase via interactions with the pre-RC components, ORC and Mcm2-7 (Wohlschlegel, Dhar et al. 2002; Sawyer, Cheng et al. 2004). Through its interactions with Cdc7/Dbf4 (DDK), Mcm10 stimulates the kinase activity leading to phosphorylation of Mcm2-7, which is required to activate the helicase (Lee, Seo et al. 2003). In addition to activating Mcm2-7, DDK is also required for the recruitment of Sld3, TopBP1, Sld2, GINS, and Cdc45 (Yabuuchi, Yamada et al. 2006; Tanaka, Umemori et al. 2007). At this point the origin DNA is denatured so that the helicase can begin unwinding DNA. It is likely that the large number of proteins present at the origin exert stress on the duplex DNA. This stress could cause destabilization of the duplex DNA and result in origin denaturation and RPA

recruitment (Zou and Stillman 2000). Loading of RPA then facilitates, in a Mcm10-dependent manner, the loading of pol α. However, the proper loading of pol α requires the concerted effort of Mcm10, And-1, and Cdc45 for recruitment, and the rearrangement of RPA on ssDNA (Walter and Newport 2000; Ricke and Bielinsky 2004; Arunkumar, Klimovich et al. 2005; Zhu, Ukomadu et al. 2007).

In order for these interactions to occur, it is possible that Mcm10 utilizes a hand-off mechanism to facilitate the progression of replication, similar to the mechanism proposed for RPA, DNA polymerases, and helicases (Yuzhakov, Kelman et al. 1999; Yuzhakov, Kelman et al. 1999; Davey and O'Donnell 2000; Kowalczykowski 2000; Mer, Bochkarev et al. 2000; Stauffer and Chazin 2004; Jiang, Klimovich et al. 2006). Mcm10 is ideally poised both temporally and spatially to function analogously to T-ag by displacing RPA through interactions with its OB-fold, while loading pol α. Similarities between Mcm10 and Simian Virus 40 Large T-angtigen (T-ag) further support this notion. It is likely that Mcm10 is recruited to origins by the helicase and acts to recruit pol α right before RPA is recruited (Homesley, Lei et al. 2000; Wohlschlegel, Dhar et al. 2002). T-ag is itself a helicase that recruits pol α, through interactions with the same region on p180, and RPA (Collins and Kelly 1991; Dornreiter, Erdile et al. 1992; Dornreiter, Copeland et al. 1993; Taneja, Nasheuer et al. 2007). Not only does this suggest that T-ag could hijack or bypass Mcm10, but it also suggests that Mcm10 could also be playing a role in RPA recruitment to origins of replication. T-ag has been hypothesized to be involved in the displacement of RPA from ssDNA via its origin binding domain, and the simultaneous loading of pol α onto DNA via its helicase domain (Ott, Rehfuess et al. 2002; Arunkumar, Klimovich et al. 2005). To examine this

possibility, NMR experiments are currently underway to examine the possible interactions between Mcm10-ID and either RPA70AB or p68N.

Another avenue currently being pursued is the effect of multiple domains on the activity of Mcm10. Chapter 4 introduced the Mcm10-ID+CTD construct and demonstrated the cooperative effect that these two domains exhibit in both DNA binding and p180 binding. Currently, crystallization trials with this construct are being performed as well as co-crystallization trials with ssDNA. As more structural information becomes available from the CTD and the ID-CTD construct, a more comprehensive model of Mcm10 interactions, and their implications for eukaryotic replication will become clearer.

Although these findings will provide much insight, they will need to be extended to the intact protein. Because full-length Mcm10 is unstable when expressed in *E. coli*, it should be expressed in another system such as insect cells. However, Mcm10 is insoluble when expressed as a His-tagged protein in insect cells (E.M. Warren & B. F. Eichman, unpublished results). Thus, another construct containing a different tag should be designed, or Mcm10 could be co-expressed with a binding partner such as p180N or the complete pol α complex. With a system such as this, one could optimize purification conditions to perform EM or small-angle X-ray scattering (SAXS) on the complex to visualize the interaction between intact Mcm10 and pol α. These methods would also facilitate the determination of the oligomeric state of Mcm10. Information on the oligomeric state would be the crucial next step in determining the role of Mcm10 at replication forks. In addition to the dimerization data presented in chapter 2, previous data suggest there are approximately two Mcm10 molecules loaded per origin in *Xenopus*

egg extracts, yielding a cellular concentration of ~16 ng/μL (Wohlschlegel, Dhar et al. 2002).

Thus it is intringing to imagine Mcm10 dimerizing at the emerging replication fork, where it would then be positioned to recruit pol α as well as δ and ε, all of which have been shown to interact with Mcm10 (Kawasaki, Hiraga et al. 2000; Fien, Cho et al. 2004; Ricke and Bielinsky 2004; Pacek, Tutter et al. 2006). The molecular anatomy of the replisome in *Xenopus laevis* egg extracts was dissected and demonstrated that when the replication fork stalls due to aphidicolin treatment, Mcm10 remains associated with the polymerases rather than the helicase and its cofactors (Cdc45 and GINS) (Pacek, Tutter et al. 2006). Thus, in the context of dimeric Mcm10, one polymerase could be loaded per strand by Mcm10. Alternatively, in a manner similar to that proposed for T-antigen, one molecule of the pol α complex could be loaded per replication fork (Huang, Weisshart et al. 1998). Thus Mcm10 would be functioning as a juggler of polymerases at the replication fork.

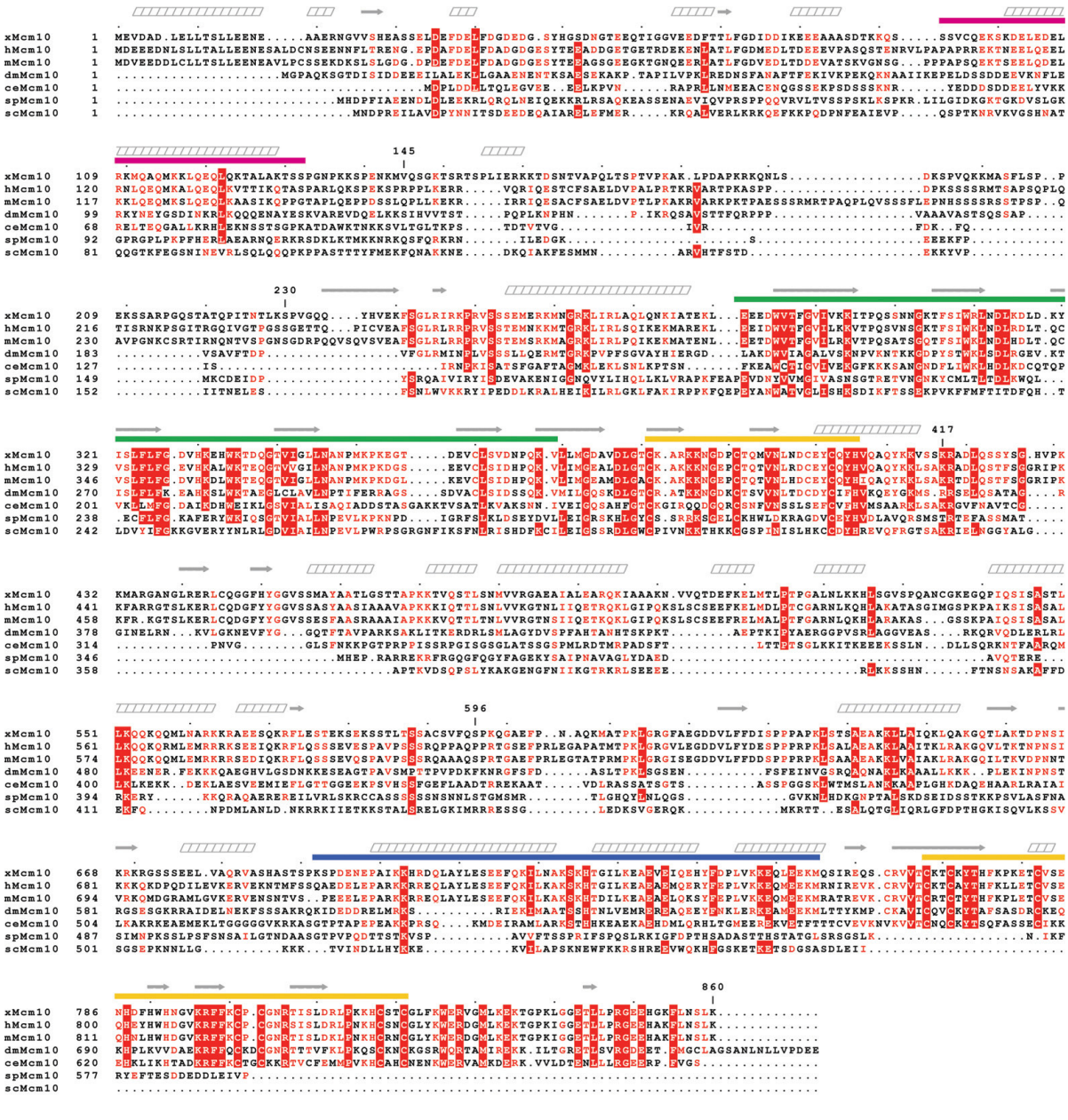
The separation of the Mcm10-polymerase complex (primosome) from the helicase complex (unwindosome) provides for a Mcm10-mediated mechanism in which DNA damage checkpoints can be activated. Under conditions of replication fork stalling or double-strand breaks (DSBs), the helicase machinery separates from the polymerase machinery (Byun, Pacek et al. 2005). This results in hyper-unwinding of DNA and RPA has been shown to bind the excess ssDNA to serve as a signal for recruitment of ATR-ATRIP (Zou and Elledge 2003) and the Rad9-Rad1-Hus1 (9-1-1) complex (Majka, Binz et al. 2006). These components provide a localized signal which stimulates the phosphorylation of Chk1 by ATR, thereby activating the checkpoint (Ball, Myers et al.

2005; Byun, Pacek et al. 2005). Additionally, it was demonstrated that although DNA unwinding is necessary for checkpoint activation, it is not sufficient and that additional DNA synthesis by pol α is needed (Michael, Ott et al. 2000; Byun, Pacek et al. 2005). This line of evidence provides the possibility of a direct link between Mcm10 and the DNA damage checkpoint. Such a link should be investigated because knockdown of Mcm10 has been shown to trigger Chk1 phosphorylation (Park, Bang et al. 2008). Additionally Mcm10 is able to recruit and stabilize pol α at the replication fork and post-translationally modified Mcm10 interacts with PCNA, a structural analog to the 9-1-1 complex (Das-Bradoo, Ricke et al. 2006; Ricke and Bielinsky 2006; Chattopadhyay and Bielinsky 2007; Zhu, Ukomadu et al. 2007; Dore, Kilkenny et al. 2009).

In summary, Mcm10 is required for eukaryotic DNA replication, although its role has not yet been fully elucidated. Mcm10 interacts with many proteins involved in DNA replication, namely, the Mcm2-7 helicase and polymerase α-primase. Upon replisome encounter of DNA damage, Mcm10 prefers to stay with the polymerase rather than the helicase. In addition to these interactions, Mcm10 interacts with the leading and lagging strand polymerases and PCNA. However the functional significance of these interactions is presently not understood. The emerging model and hypotheses expressed in this chapter suggest the major role of Mcm10 is to recruit and rearrange proteins at sites of replication. This role suggests that a malfunction in Mcm10 would lead to genome instability and therefore Mcm10 may prove to be an important therapeutic target.

APPENDIX A

DOMAIN ARCHITECTURE AND BIOCHEMICAL CHARACTERIZATION OF VERTEBRATE MCM10*



* The work presented in this chapter was published in 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) *J Biol Chem.* **283**, 3338-3348.

Figure A1, continued. Mcm10 sequence alignment. Primary sequence alignment of Mcm10 proteins from *Xenopus laevis* (x), *Homo sapiens* (h), *Mus musculus* (m), *Drosophila melanogaster* (de), *Caenorhabditis elegans* (ce), *Schizosaccharomyces pombe* (sp), and *Saccharomyces cerevisiae* (sc). Conserved residues are shown in red letters and invariant or strongly conserved residues are highlighted with a red background. Predicted secondary structural elements are shown above the sequence in grey (α-helices, leaning boxes; β-strands, arrows). Predicted structural motifs are shown as colored bars (magenta, coiledcoil; green, OB-fold; yellow, zinc motif; blue, winged helix). Sequence alignments were generated with ClustalW (Thompson, Higgins et al. 1994) and displayed using ESPript (Gouet, Courcelle et al. 1999). Secondary and tertiary structure predictions were carried out using MultiCoil, Phyre, 3D-PSSM (Berger, Wilson et al. 1995; Wolf, Kim et al. 1997; Kelley, MacCallum et al. 2000).

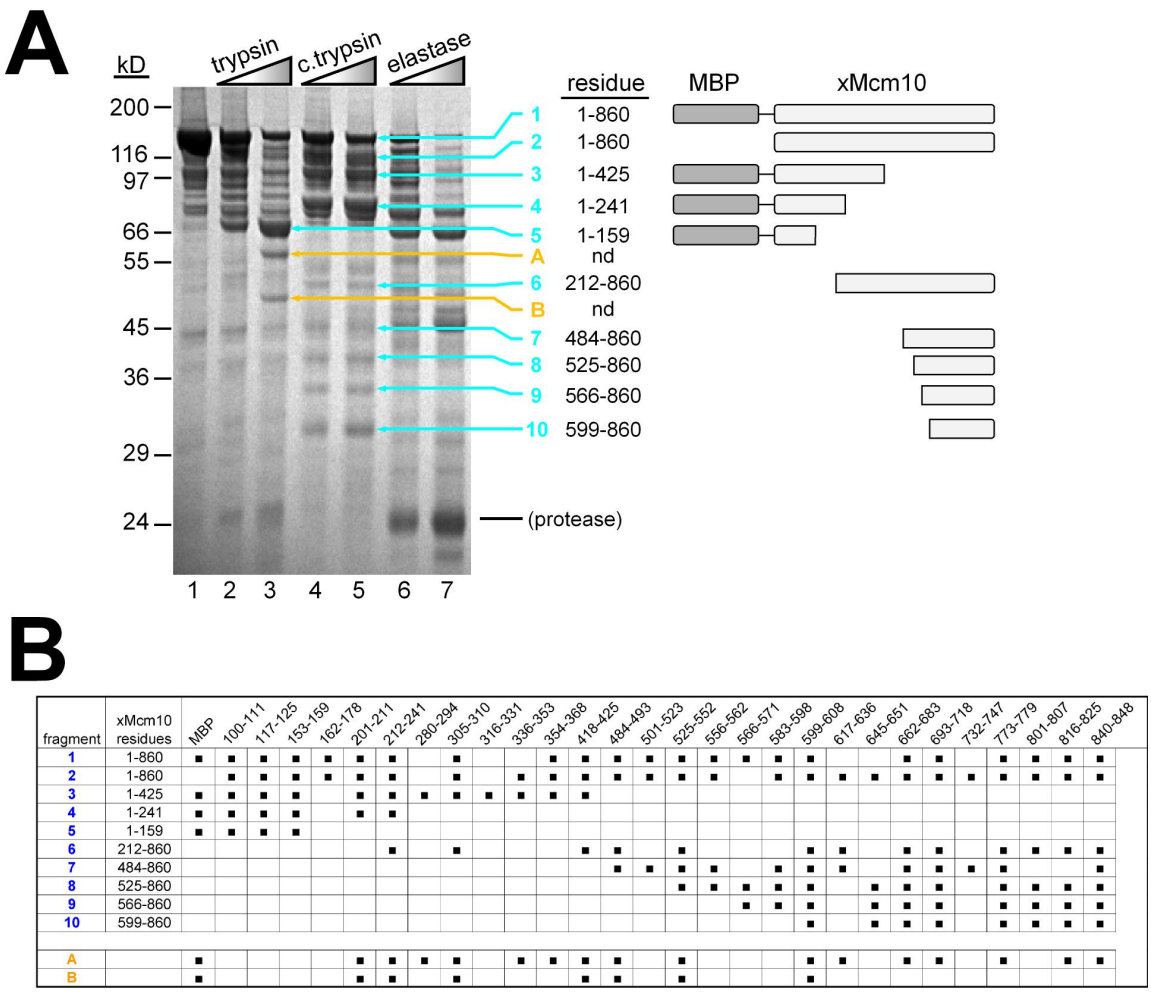


Figure A2. Identification of proteolytically sensitive regions within xMcm10. A, Same Coomassie SDS-PAGE gel shown in Figure 7B, with the major proteolytic fragments labeled 1-10, A, and B. Bands 1-10 (blue) were single species that could be unambiguously identified by mass spectrometry, while bands A and B (orange) were mixtures of several co-migrating proteins and thus could not be defined. **B**, Peptide coverage map of fragments shown in panel **A**. Each band 1-10, A, and B was excised from the gel and subjected to complete lysis with tripsin and the resulting tryptic peptides (numbers at the top of the chart) were identified by MALDI-TOF and TOF/TOF mass spectrometry. Endpoints of fragments 1-10 could be unambiguously assigned based on the recovered peptides. Peptides spanning MBP and the entire length of Mcm10 were recovered from fragments A and B.

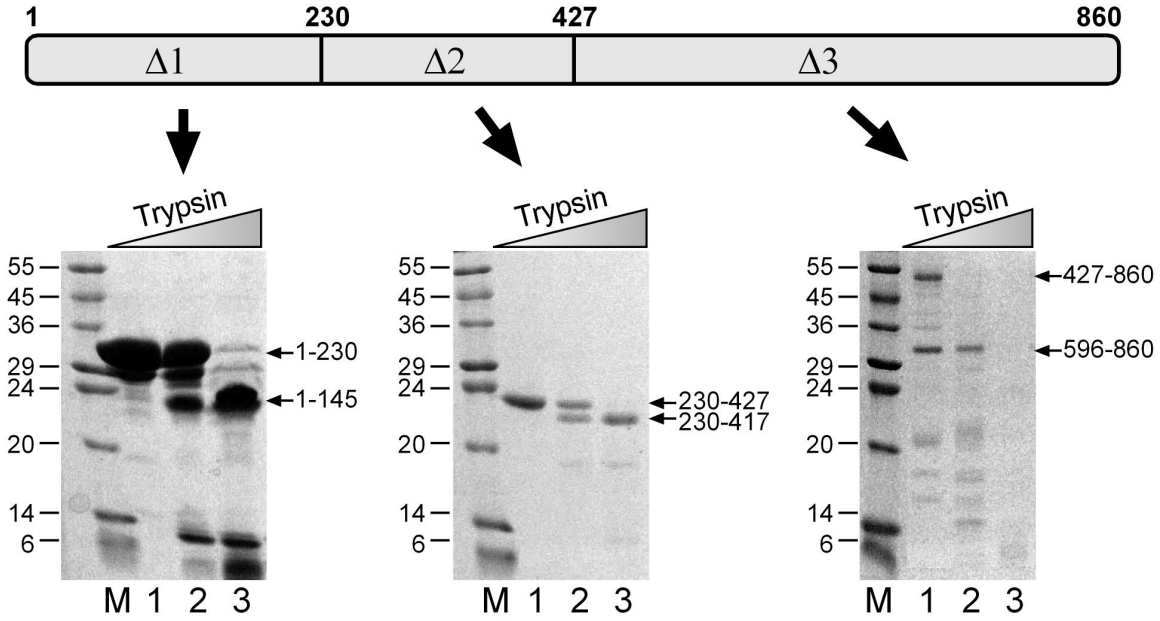


Figure A3. Identification of xMcm10 domains. Three truncation mutant proteins (Δ1, Δ2, Δ3) were purified (lane 1) and subjected to limited proteolytic digestion (lanes 2, and 3) with increasing amounts of trypsin (shown), chymotrypsin, elastase, and endo-GluC. 50-200 pmol xMcm10 mutant was incubated with 1 and 10 ng elastase (lanes 2 and 3, respectively) for 30 min at 37° C. Intact masses of proteolytic products were identified by MALDI mass spectrometry of each reaction mixture. N-terminal sequences were identified by Edman degradation of individual bands from the gel. Lane M, molecular weight standards.

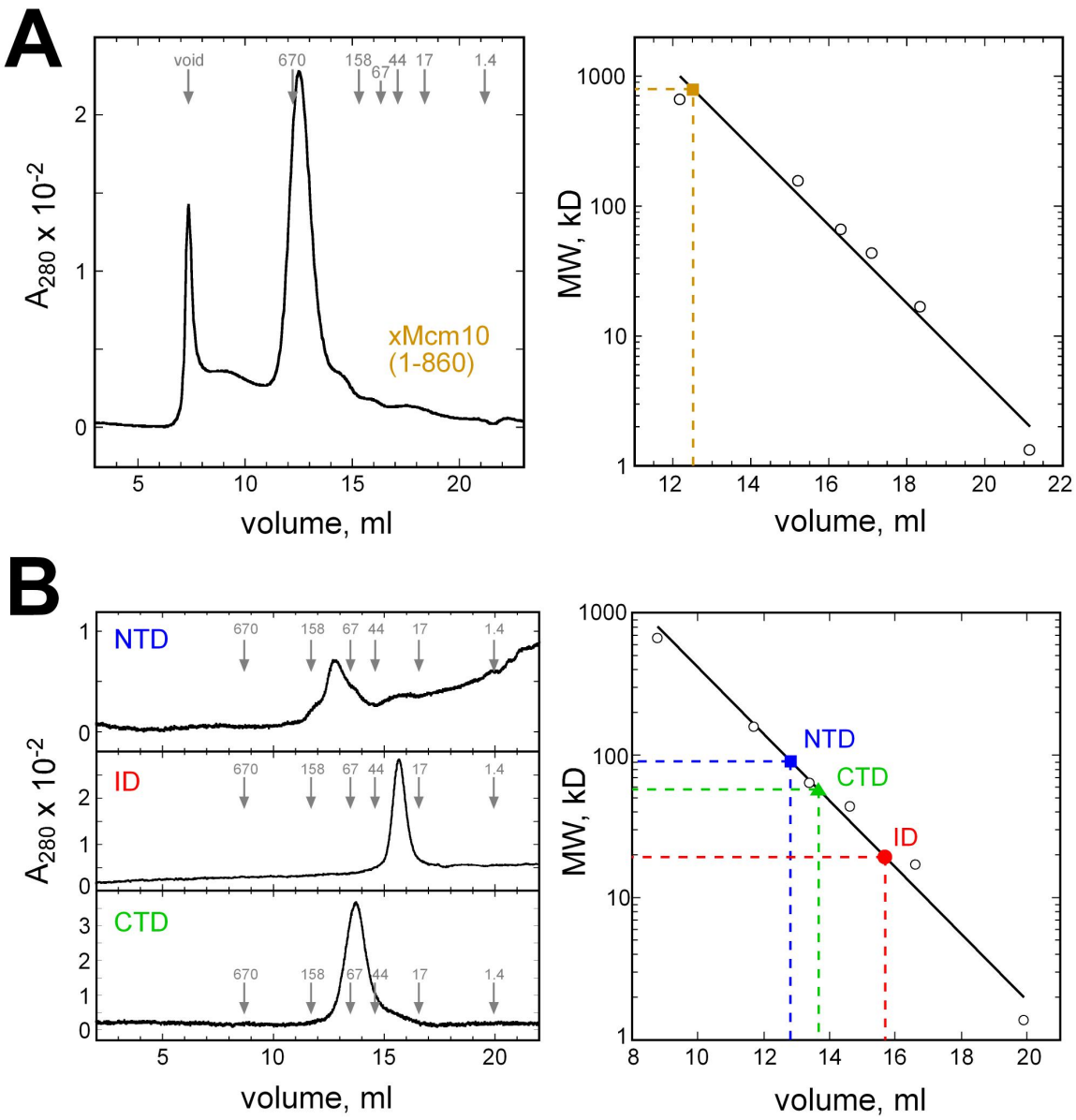


Figure A4. Gel filtration analysis of xMcm10. The left panels show gel filtration chromatograms of full-length xMcm10 (**A**) and individual xMcm10 domains (**B**). Elution volumes of molecular weight standards are marked by gray arrows. The standard curves are shown on the right, with elution volumes for xMcm10 (brown square), xMcm10-NTD (blue square), ID (red circle), and CTD (green triangle) superimposed. Molecular weights calculated from primary sequences are as follows: xMcm10, 95.4 kD; NTD, 16.2 kD; ID, 22.7 kD; CTD, 30.1 kD.

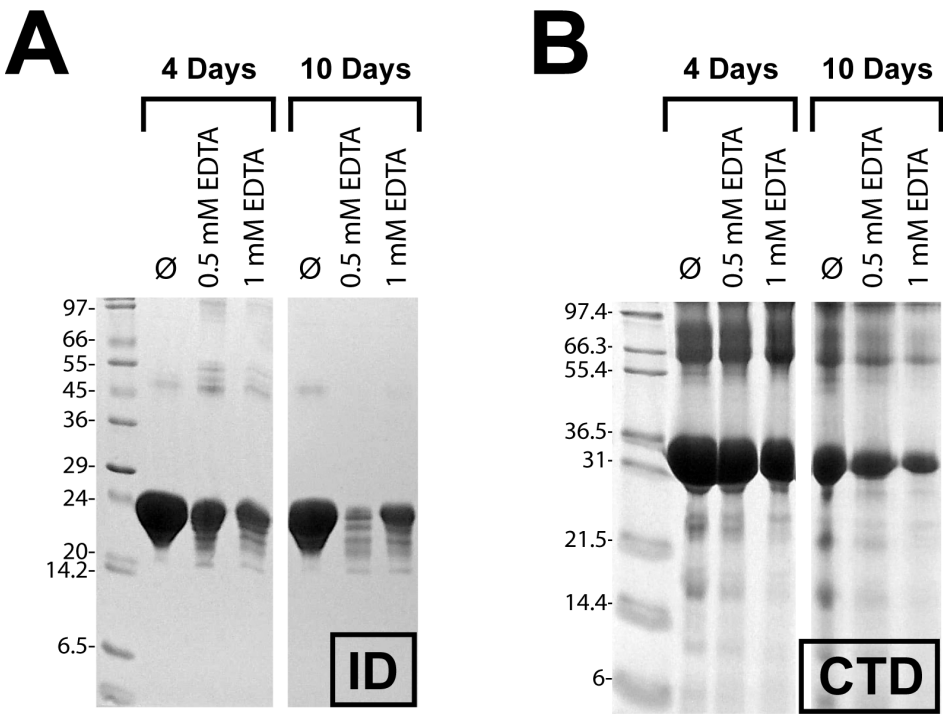


Figure A5. EDTA affects the stability of xMcm10-ID and xMcm10-CTD. An SDS-PAGE gel showing the effect of EDTA on the stability of xMcm10-ID (**A**) and -CTD (**B**). Samples were incubated at room temperature for a period of 10 days alone and in the presence of 0.5 mM or 1.0 mM EDTA. Samples were taken on day 4 and frozen at -80ºC to be resolved by SDS-PAGE on day 10.

APPENDIX B

STRUCTURAL BASIS FOR DNA BINDING BY REPLICATION INITIATOR MCM10*

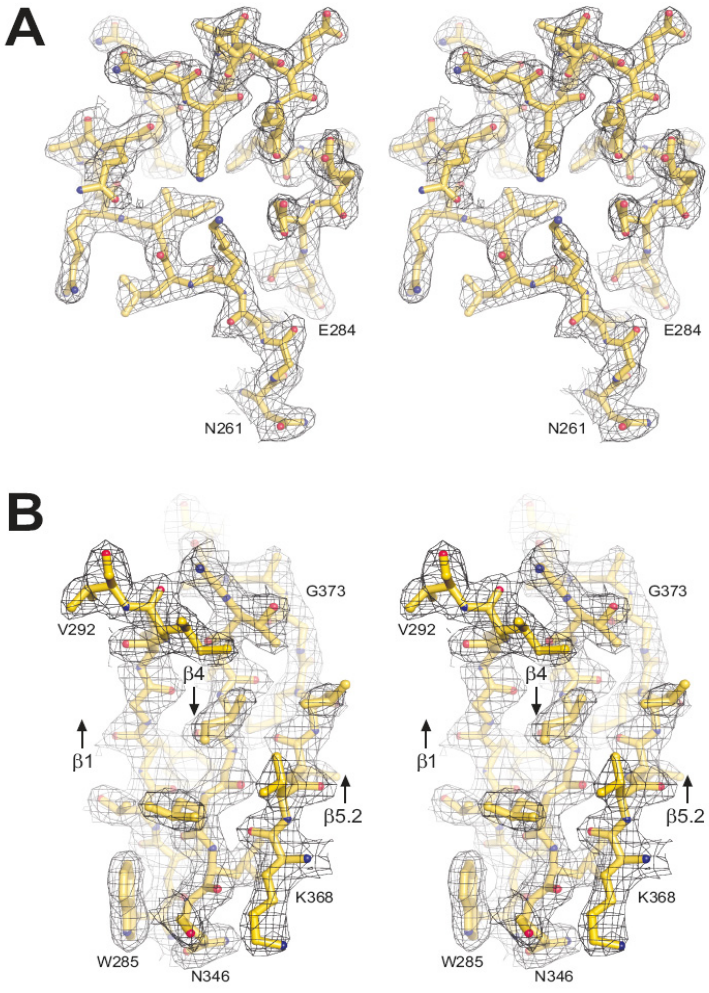


Figure B1. Crystallographic model of Mcm10-ID. Stereodiagrams of two representative sections of the final refined model (gold sticks) are shown superimposed onto composite omit electron density maps contoured at 1σ. (A) α-helical region Asn261-Glu284. (B) Antiparallel β-sheet formed from strands β1, β4, and β5.2 (arrows).

* The work presented in this chapter was published in Warren, E. M., Vaithiyalingam, S.R., 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): 1892-1901.

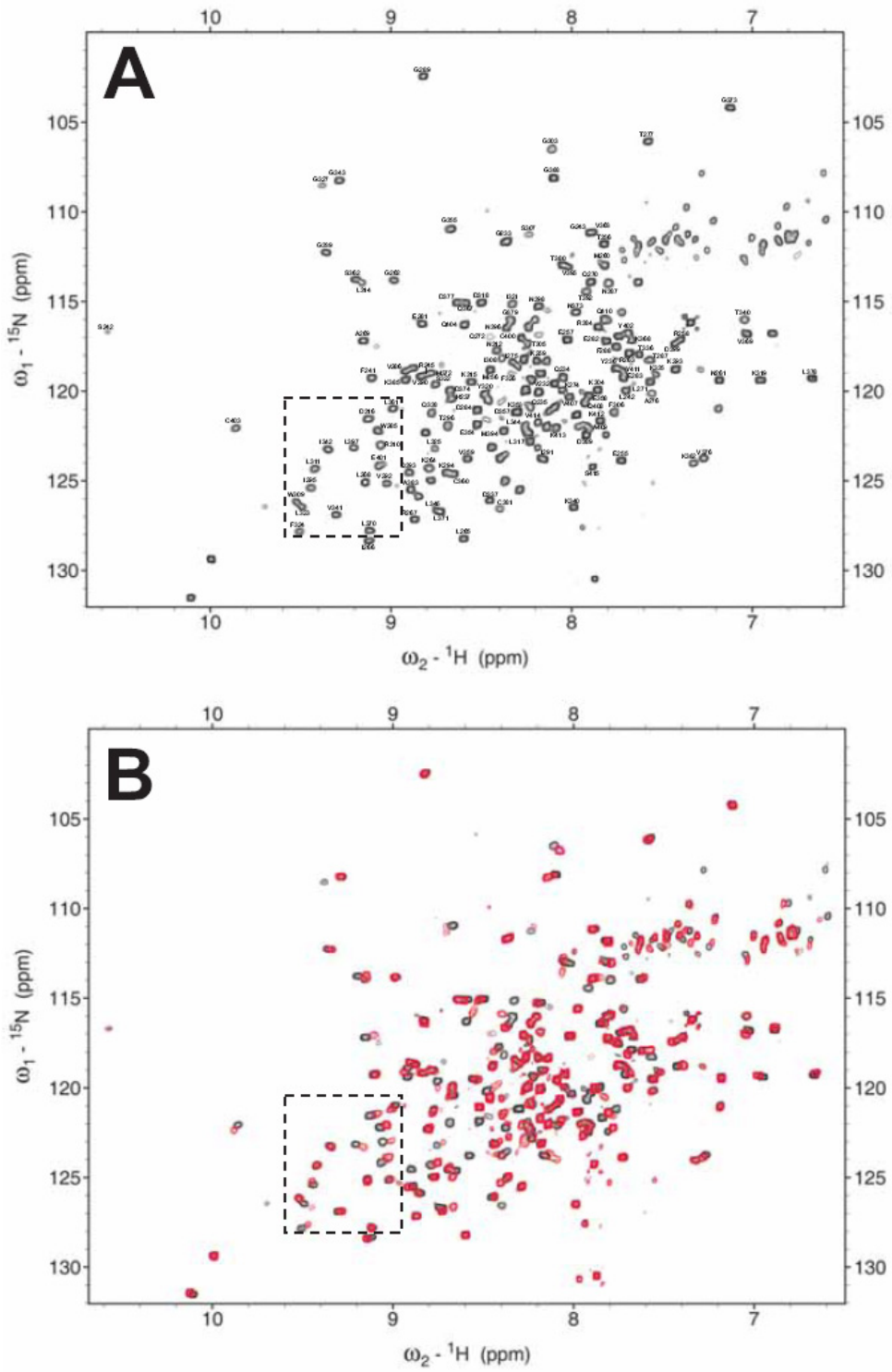


Figure B2. **NMR chemical shift assignments and perturbation by ssDNA binding.** (A) ^{1}H-^{15}N HSQC spectrum of Mcm10-ID with assigned chemical shifts labeled by residue number. (B) ^{1}H-^{15}N HSQC spectrum for Mcm10-ID in the absence (black) and presence (red) of ssDNA. The spectrum of the complex was acquired at a protein:DNA ratio of 1:1. The region of the spectra shown in Figure 17 is boxed.

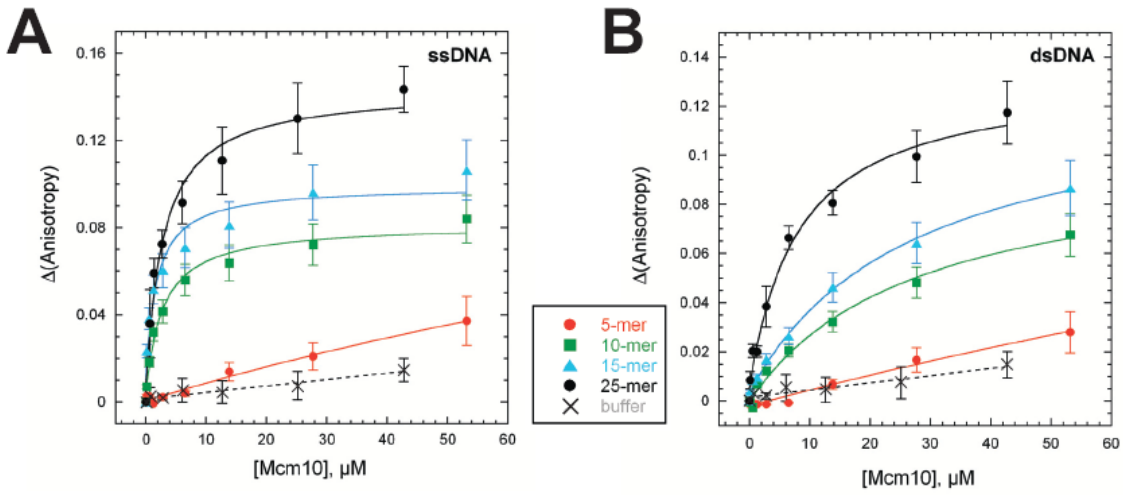


Figure B3. **Dependence of DNA length on Mcm10-DNA binding.** (A) Binding of Mcm10 to fluorescently-labeled ssDNA (A) and dsDNA (B) oligonucleotides of varying lengths was monitored by an increase in fluorescence anisotropy as a function of protein concentration. Red circles, 5mer; green squares, 10mer; blue triangles, 15mer; black circles, 25mer. A negative control in which only buffer was added to 25mer DNA is shown as black crosses. Isotherms represent the average from three independent measurements. K_d values shown in Figure 18D were determined by fitting the data using the equation $\Delta A = A_f[\text{Mcm10}]/(K_d+[\text{Mcm10}])$, in which A_f represents the anisotropy at saturation.

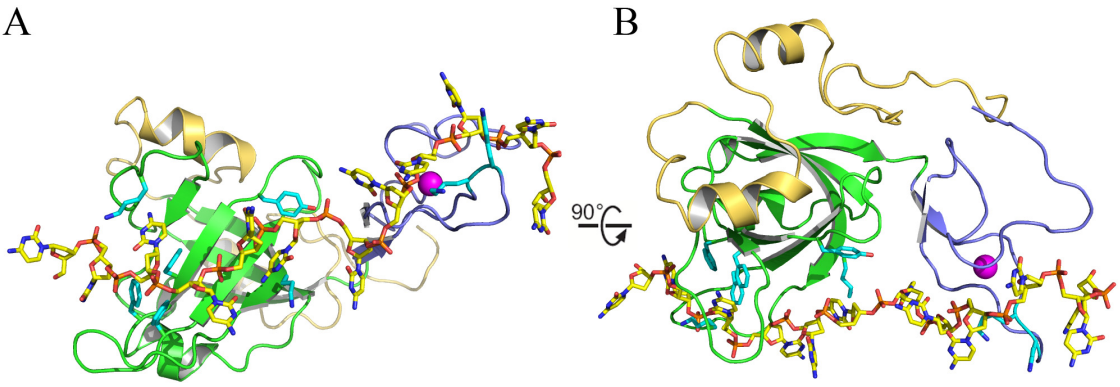


Figure B4. HADDOCK ssDNA docking. (A) ssC9mer was docked onto the xMcm10-ID crystal structure using restraints from NMR titrations as well as mutagenic DNA binding data. Residues shown to be important for DNA binding are rendered as cyan sticks to show their location with respect to the docked DNA. (B) Structure from (A) rotated forward by 90º.

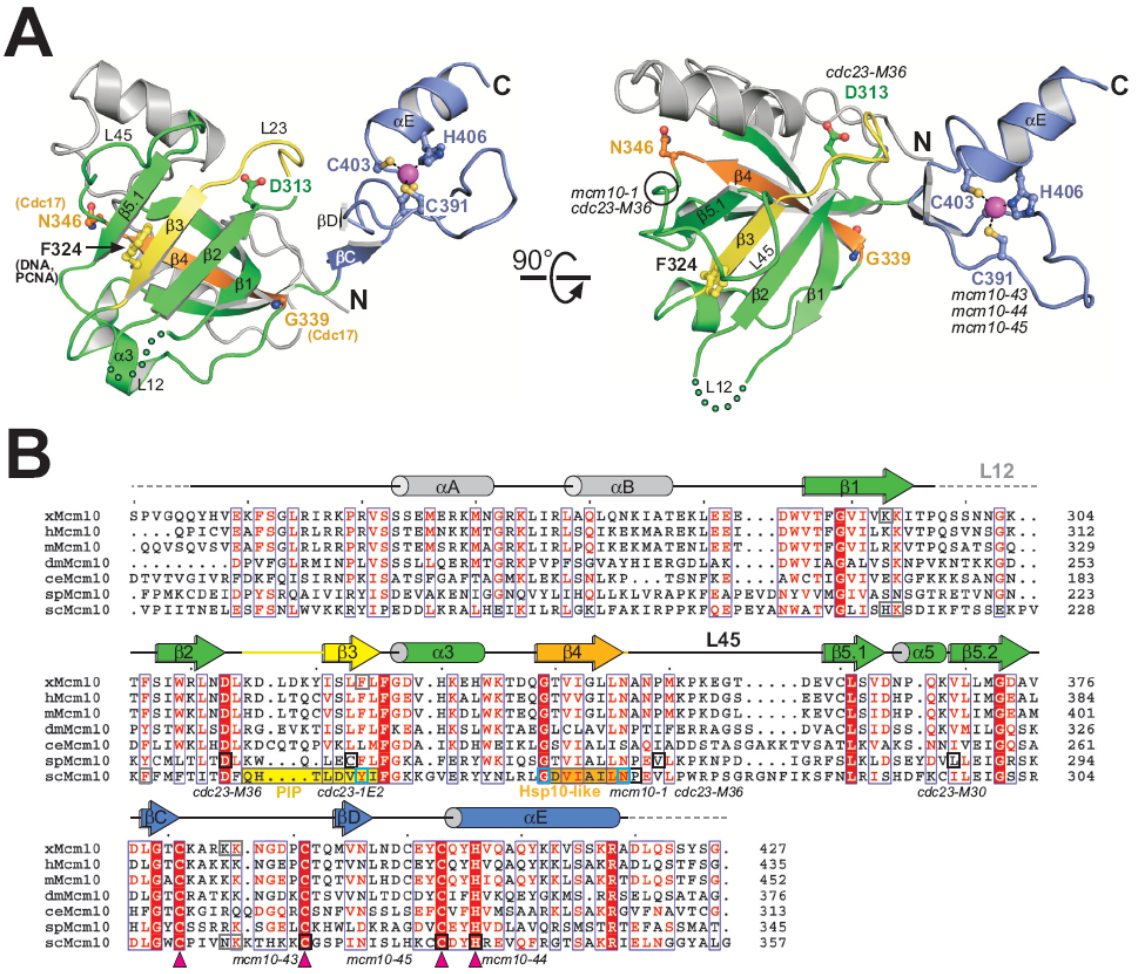


Figure B5. The locations of important residues within Mcm10-ID. (A) The same two views of the Mcm10-ID crystal structure as Figure 15, with residues important for Mcm10 function labeled. For clarity, DNA binding residues K293 and E385/386 identified in the present work are not shown in the structure. Molecules affected by mutation of highlighted Mcm10 residues are labeled in parentheses on the left image, and names and positions of genetic mutations identified in yeast are labeled in italics on the right image. Putative PIP box (Das-Bradoo, Ricke et al. 2006) and Hsp10-like motifs (Ricke and Bielinsky 2006) in scMcm10 are highlighted yellow and orange, respectively. (B) The sequence alignment of the conserved ID from known Mcm10 proteins is shown together with schematic secondary structural elements and colored as in panel A. Residues identified from genetic screens in budding (mcm10) and fission yeast (cdc23) that affect cell growth and DNA replication are highlighted with black boxes. scMcm10 residues that interact with Cdc17 and PCNA are highlighted with blue boxes. Mutations that affect xMcm10 binding to DNA *in vitro* or that increase the sensitivity of *S. cerevisiae* to hydroxyurea are highlighted with grey boxes. x, *Xenopus laevis*; h, *Homo sapiens*; m, *Mus musculis*; dm, *Drosophila melanogaster*; ce, *Caenorhabditis elegans*; sp, *Schizosaccharomyces pombe*; sc, *Saccharomyces cerevisiae*.

Table B1. Multidimensional NMR experiments and acquisition parameters

Experiment Name	Dimension	NS	Time (h:min)	Number of Increments (t1×t2×t3)	Magnetic Field (MHz)	Carrier Frequency (ppm)
^{15}N HSQC[a]	2D	8	0:40	1024(H)×256(N)	800	4.7(H); 117.5(N)
HNCA[b]	3D	16	20:48	1024(H)×40(N)×96(CA)	600	4.7(H); 117.5(N); 52.5(CA)
HNCACB[b]	3D	32	51:20	1024(H)×40(N)×120(CB)	600	4.7(H); 117.5(N); 41.0(CB)
CBCACONH[b]	3D	16	20:48	1024(H)×40(N)×96(CB)	600	4.7(H); 117.5(N); 41.0(CB)
(H)C(CO)NH-TOCSY[b]	3D	32	41:36	1024(H)×40(N)×96(CB)	600	4.7(H); 117.5(N); 41.0(C)
HNCO[b]	3D	8	4:26	1024(H)×40(N)×40(CO)	600	4.7(H); 117.5(N); 175(CO)

(a) ^{15}N-enriched samples

(b) ^{15}N,^{13}C-enriched samples

APPENDIX C

CHARACTERIZATION OF PHYSICAL INTERACTIONS BETWEEN MCM10, DNA POLYMERASE α, AND DNA*

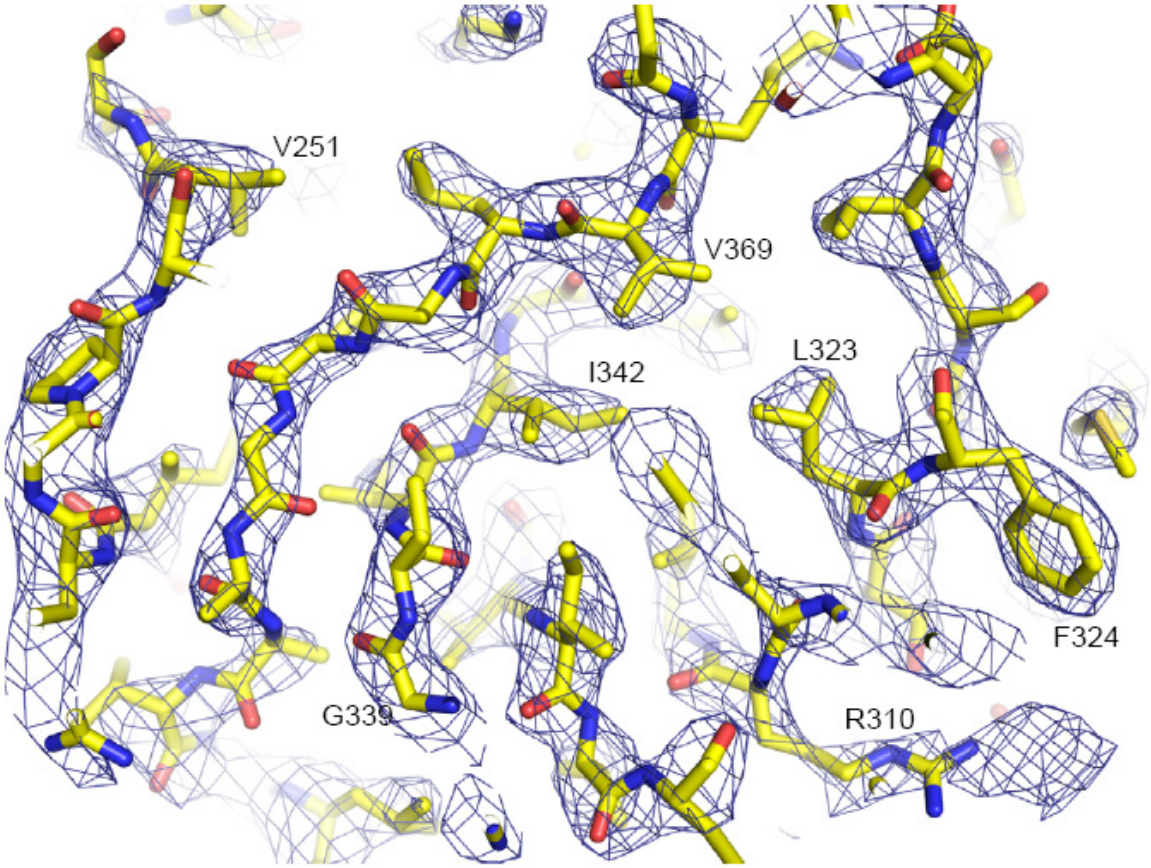


Figure C1. Fit of crystallographic model to electron density. Shown is a representative section of the final refined protein model (sticks) superimposed onto a $2F_o$-F_c composite omit electron density map contoured at 1σ. Several residues are labeled as landmarks.

* The work presented in this chapter was submitted as Warren, E. M., Huang, H., Fanning, E., Chazin, W. J., and Eichman, B. F., (2009) *J Biol Chem.* In press.

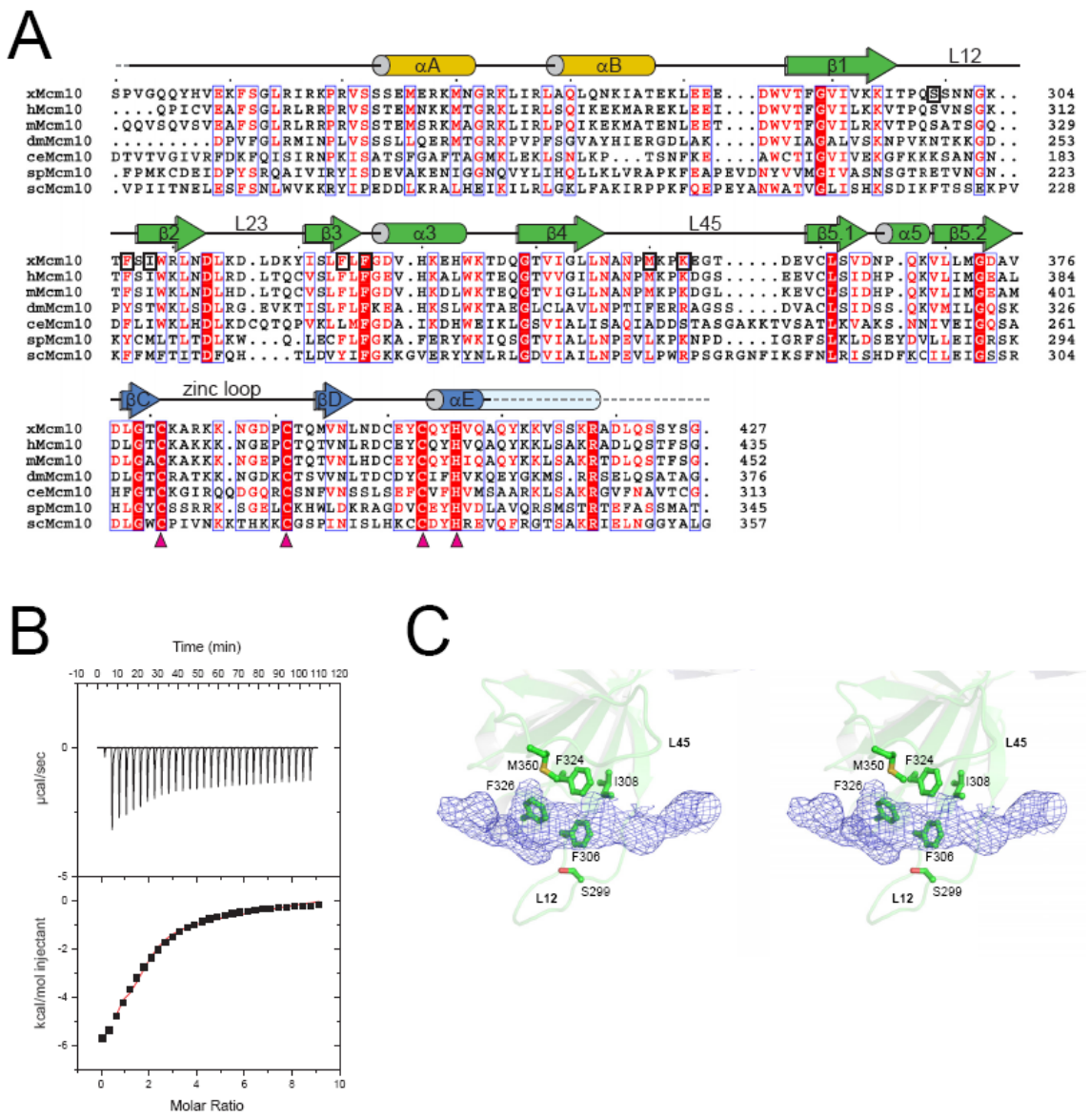


Figure C2. **Nature of the interactions between Mcm10-ID and ssDNA.** (**A**), Mcm10-ID sequence alignment showing secondary structure elements and DNA binding residues (black boxes) from the Mcm10-ID/ssDNA complex. Protein regions not observed in the electron density are depicted by dashed lines (coil regions) or a light blue helix (αE). (**B**), Isothermal titration calorimetry measurements for 25mer ssDNA titrated into Mcm10-ID at 21ºC. *Upper panel*, raw ITC data for sequential injections of $p180^{189-323}$; *lower panel*, integrated heat responses (squares) fit with a single site binding model (continuous line). The following parameters were obtained from the fit: K_d, 27 ± 0.3 µM; ΔH, -9.8 kcal/mol; TΔS, -3.6 kcal/mol. (**C**), Stereoview of the ssDNA binding site on the Mcm10-ID OB-fold. Annealed omit electron density contoured at 3σ is shown as blue mesh, and ssDNA-contacting residues are rendered as sticks.

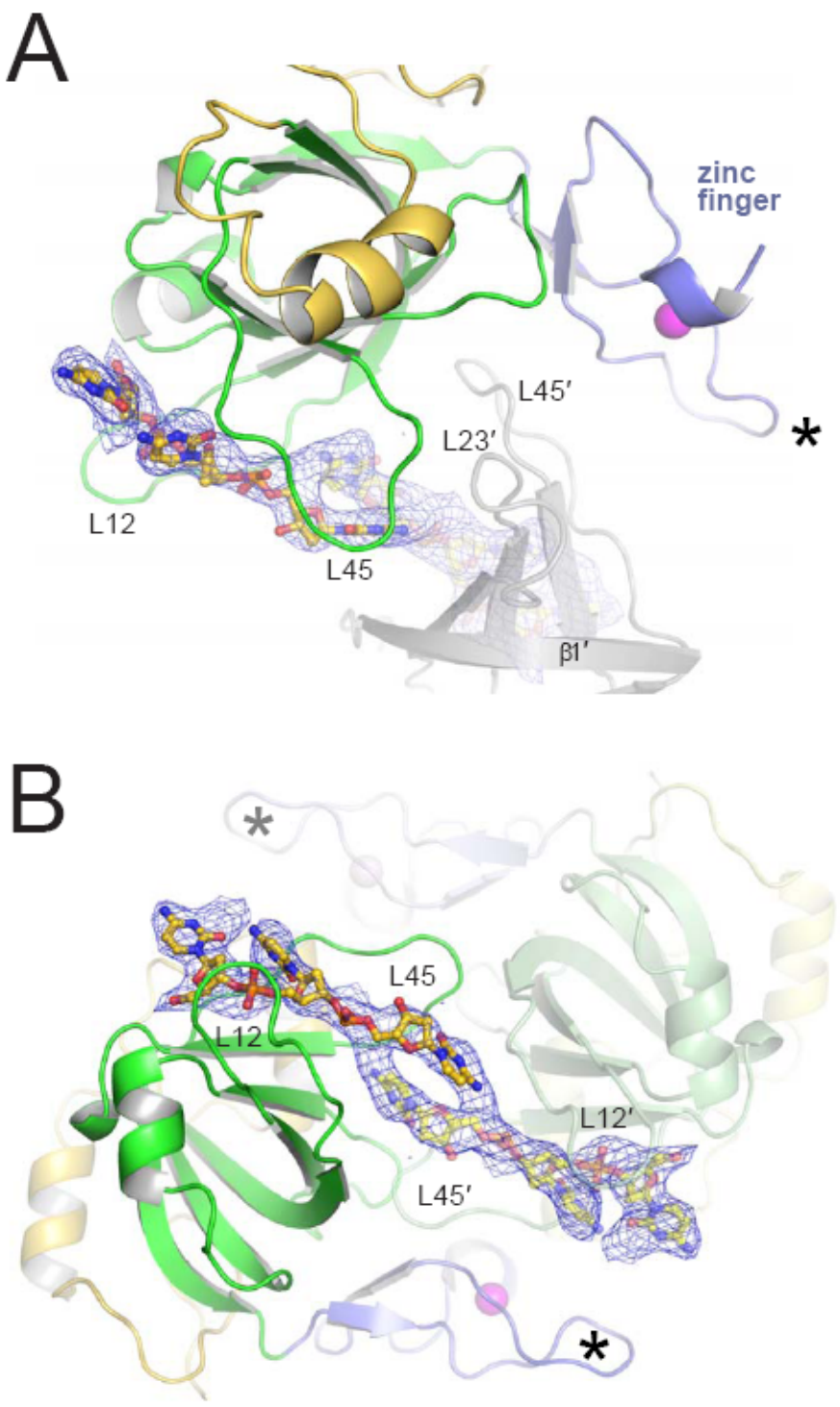


Figure C3. **Crystal packing of the Mcm10-ID/ssDNA complex.** (**A**). Crystal structure of the Mcm10-ID/ssDNA complex colored as in Figure 20, shown with the protein OB-fold (grey) and ssDNA (yellow carbons) from a symmetry-related complex. An annealed omit map for ssDNA contoured at 3σ is shown. The position of the zinc loop previously implicated in DNA binding is highlighted with an asterisk. (**B**), A different angle is shown to highlight the packing of the loops, precluding the ssDNA from interacting with the Zn-finger. The symmetry-related protein/DNA complex is dimly colored with yellow DNA carbons.

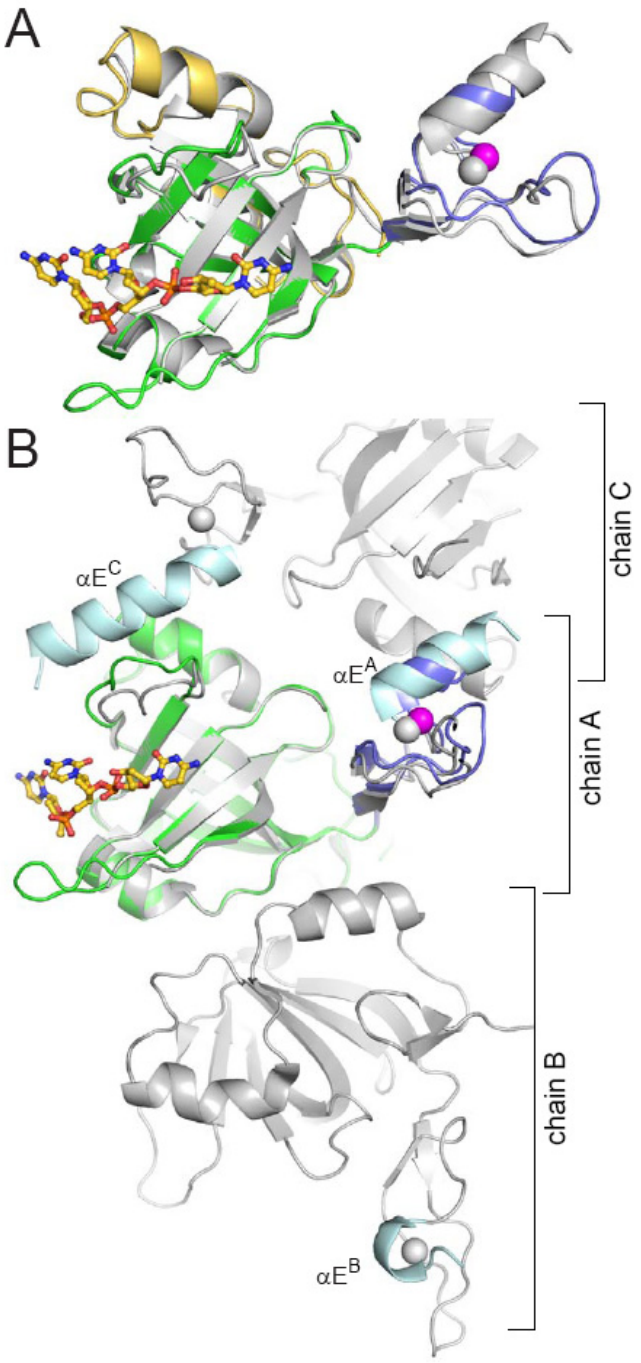


Figure C4. Differences in unliganded and ssDNA bound Mcm10-ID as a result of crystal packing interactions. (**A**), Superposition of the Mcm10-ID/ssDNA complex, colored by motif as in Figure 20, with chain A of the unliganded crystal structure (PDB ID 3EBE, Warren 2008) in grey. (**B**), Protein interactions stabilize the zinc finger helix (αE) in unliganded Mcm10-ID. The Mcm10-ID/ssDNA complex (green OB-fold, dark blue zinc finger) is superimposed onto chain A of the unliganded structure, from which the entire asymmetric unit (chains A, B, C) is shown in grey with αE helices colored light blue. Zinc finger helices from chains A and C are more ordered, and are forming contacts with a symmetry-related protomer in the crystal, whereas αE from chain B is disordered and does not make intermolecular contacts.

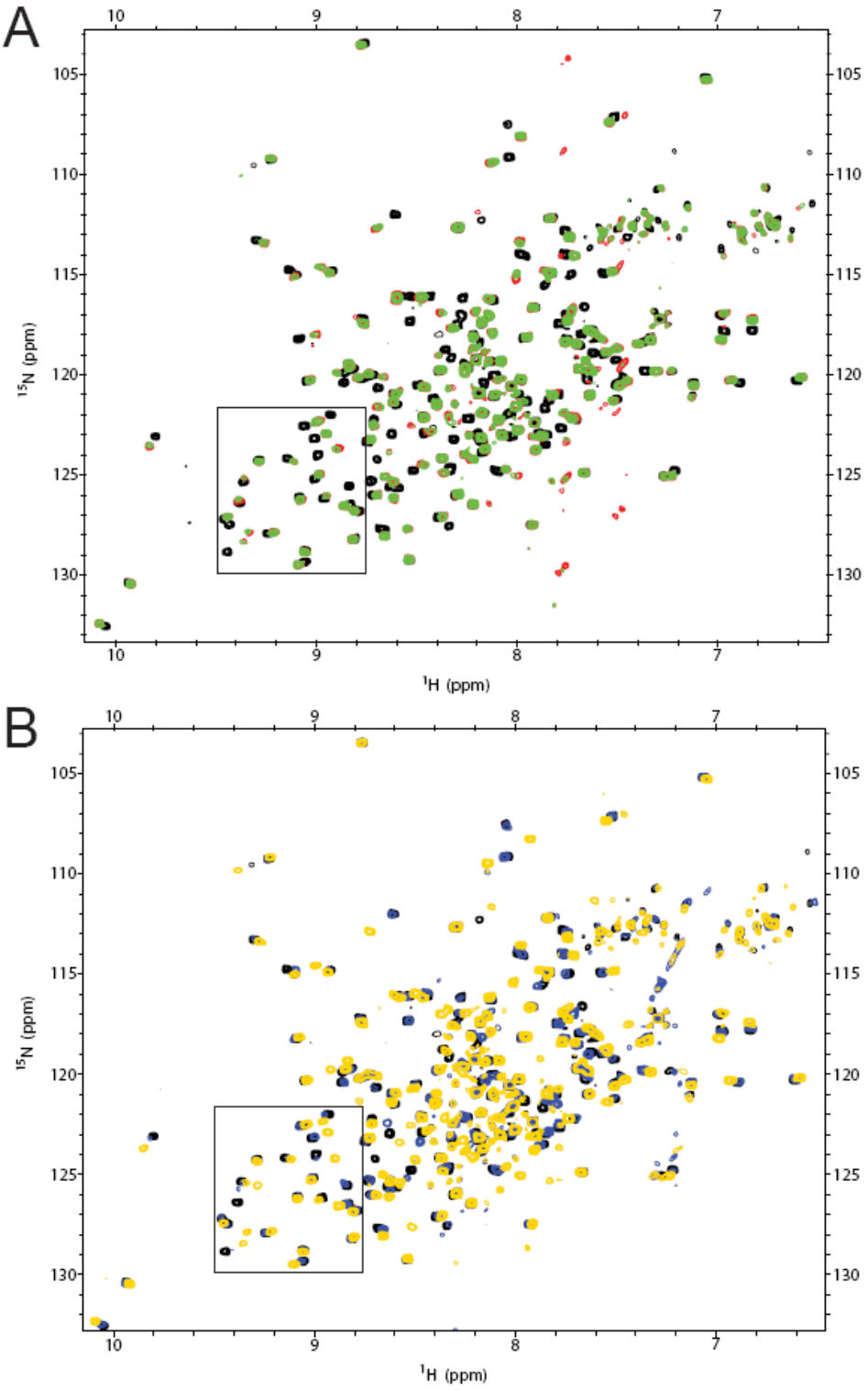


Figure C5. Competition for Mcm10-ID binding by ssDNA and p180$^{189-323}$. (**A**), ^{15}N-^{1}H HSQC spectrum for Mcm10-ID alone (black), 1:1 ratio of Mcm10-ID:ssDNA (red), and a 1:1:1 ratio Mcm10-ID:ssDNA:p180$^{189-323}$ (green). (**B**), The reverse titration with Mcm10-ID alone (black) mixed in a 1:1 ratio of p180$^{189-323}$ (blue), and then ssDNA in a 1:1:1 ratio (gold). The region of the spectra shown in Figure 23 is boxed.

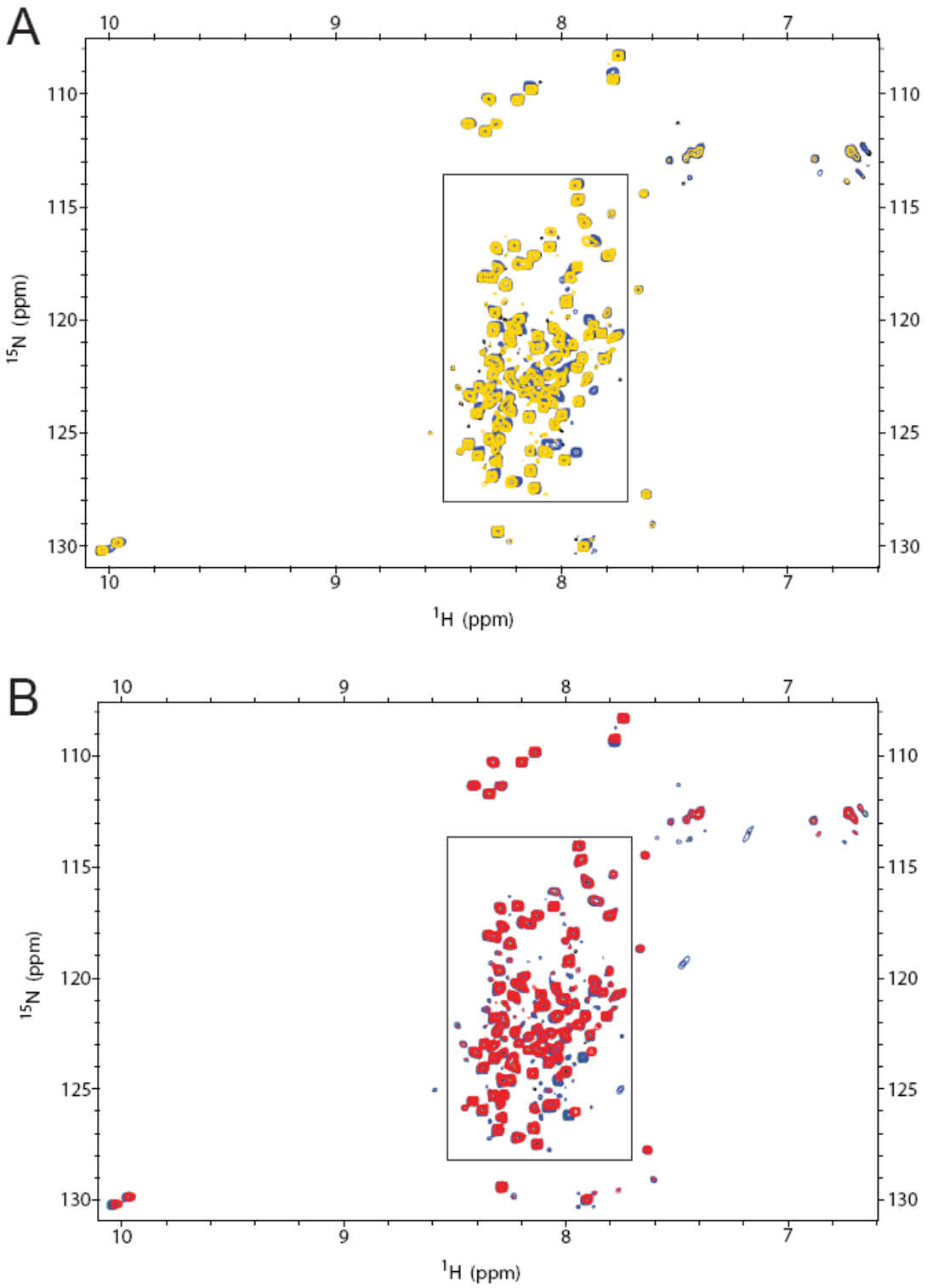


Figure C6. DNA-induced release of Mcm10-ID from p180$^{189-323}$. (**A**), ^{15}N-^{1}H HSQC spectrum for ^{15}N-enriched p180$^{189-323}$ (black), 1:1 molar ratio of p180$^{189-323}$:Mcm10-ID (blue), and a 1:1:1 ratio of p180$^{189-323}$:Mcm10-ID:ssDNA (gold). (**B**), The reverse titration with p180$^{189-323}$ alone (black), 1:1 molar ratio of ssDNA (green), a 1:1:0.67 molar ratio of p180$^{189-323}$:ssDNA: Mcm10-ID (blue), and then a 1:1:1 molar ratio of ^{15}N-p180$^{189-323}$:ssDNA:Mcm10-ID (red). The region of the spectra shown in Figure 23 is boxed.

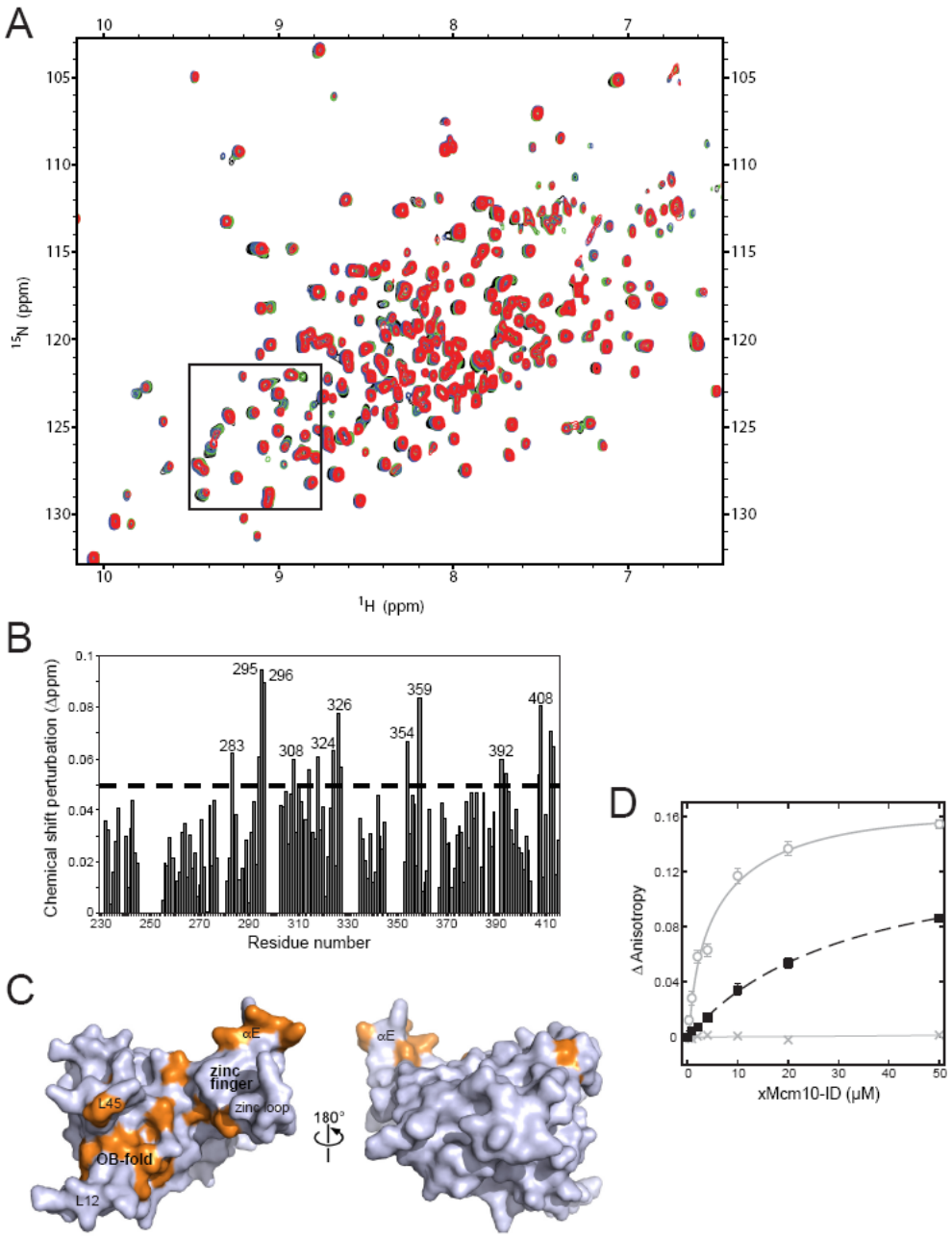


Figure C7. Binding of p180$^{286\text{-}310}$ to Mcm10-ID. (**A**), ^{15}N-^{1}H HSQC spectra from ^{15}N-enriched Mcm10-ID performed at Mcm10-ID:p180$^{286\text{-}310}$ ratios of 1:0 (black), 1:0.25 (green), 1:0.5 (blue), and 1:1 (red). (**B**), Quantitation of chemical shift perturbations of ^{15}N-enriched Mcm10-ID upon addition of 1:1 molar ratio of p180$^{286\text{-}310}$. The dashed line represents 1 standard deviation above the mean. A shift of zero indicates an unassigned residue. (**C**), Surface representation of Mcm10-ID showing that residues exhibiting a significant shift in response to p180$^{286\text{-}310}$ (orange) predominate on the ssDNA binding face of the protein. (**D**), Fluorescence anisotropy titration in which Mcm10-ID was added to FITC-labeled p180$^{286\text{-}310}$ (black boxes). Titration of MTS-fluorescein-p180$^{189\text{-}323}$ with Mcm10-ID (circles) and buffer only (crosses) from Figure 21 are shown for reference. The error bars represent the standard deviations from the average values from three independent measurements. The curve fits are non-linear regression of the data as described in Materials and Methods of chapter 4.

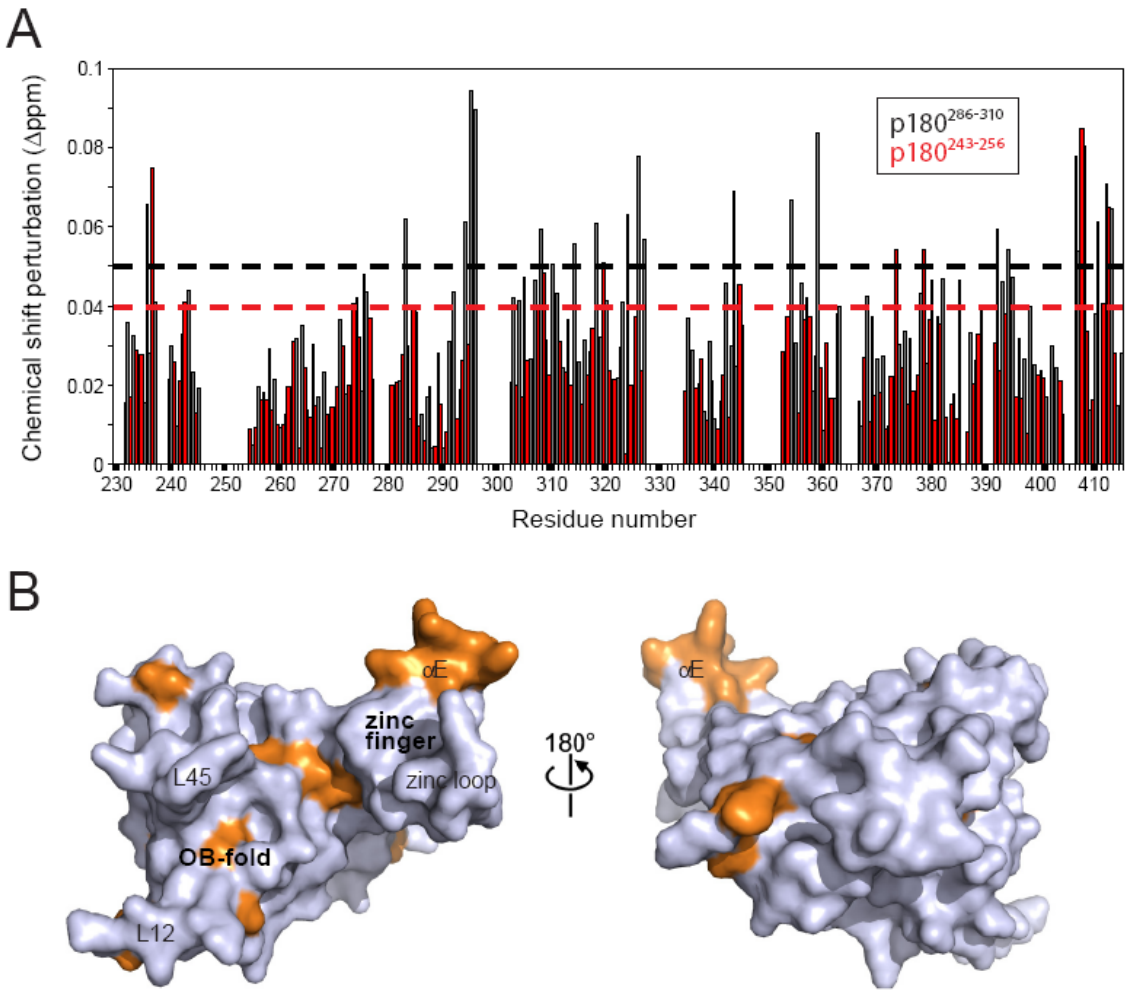


Figure C8. The p180$^{243-256}$ peptide does not bind specifically to Mcm10. (**A**), A comparison of the magnitudes of chemical shift perturbations of ^{15}N-enriched Mcm10 resulting from addition of unlabeled p180$^{243-256}$ (red bars) and p180$^{286-310}$ (grey bars). Dashed lines represent 1 standard deviation above the mean shift perturbation for all residues. (**B**), A surface representation of Mcm10-ID, with orange highlighting those residues that exhibit a significant shift (above the dashed line in panel A) from the p180$^{243-256}$ titration.

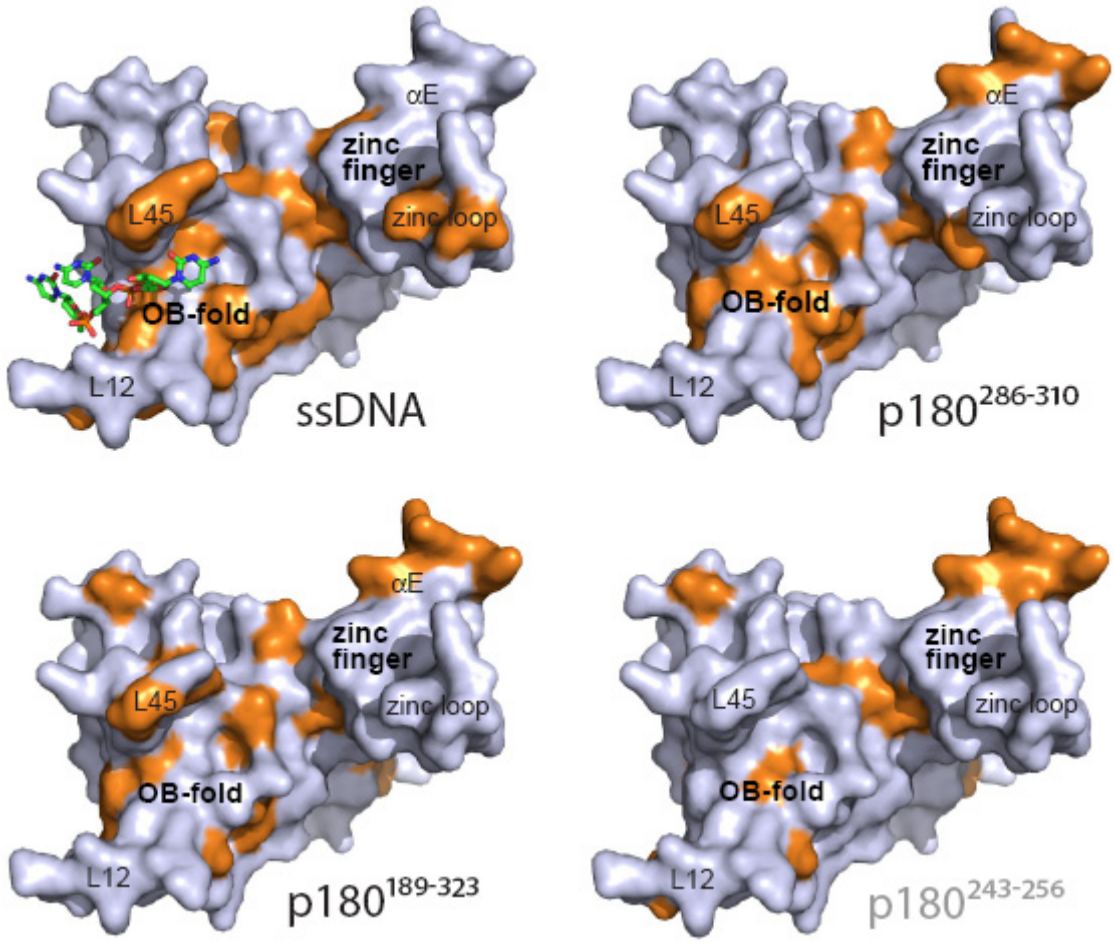


Figure C9. ssDNA and the N-terminal domain of p180 share the same binding site on Mcm10-ID. The protein from the Mcm10-ID/ssDNA co-crystal structure is rendered as a solvent accessible surface. Mcm10 residues exhibiting significant NMR chemical shifts perturbations in ^{15}N-^{1}H HSQC spectra are highlighted orange. Maps of p180 fragment binding were determined in the present work, and the ssDNA map is from Warren *et al* (2008) Structure *16*, 1892-1901, and is shown here for comparison. Residues 297-302 in the L12 loop were not assigned in the NMR spectra, and thus were unable to be measured for perturbation.

REFERENCES

Afonine, P. V., R. W. Grosse-Kunstleve, et al. (2005). "A robust bulk-solvent correction and anisotropic scaling procedure." Acta Crystallogr D Biol Crystallogr **61**(Pt 7): 850-5.

Allen, J. B., Z. Zhou, et al. (1994). "The SAD1/RAD53 protein kinase controls multiple checkpoints and DNA damage-induced transcription in yeast." Genes Dev **8**(20): 2401-15.

Altman, A. L. and E. Fanning (2001). "The Chinese hamster dihydrofolate reductase replication origin beta is active at multiple ectopic chromosomal locations and requires specific DNA sequence elements for activity." Mol Cell Biol **21**(4): 1098-110.

Anumanthan, G., S. K. Halder, et al. (2006). "Oncogenic serine-threonine kinase receptor-associated protein modulates the function of Ewing sarcoma protein through a novel mechanism." Cancer Res **66**(22): 10824-32.

Aparicio, O. M., A. M. Stout, et al. (1999). "Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication." Proc Natl Acad Sci U S A **96**(16): 9130-5.

Appleton, B. A., J. Brooks, et al. (2006). "Crystal structure of the cytomegalovirus DNA polymerase subunit UL44 in complex with the C terminus from the catalytic subunit. Differences in structure and function relative to unliganded UL44." J Biol Chem **281**(8): 5224-32.

Araki, H., S. H. Leem, et al. (1995). "Dpb11, which interacts with DNA polymerase II(epsilon) in Saccharomyces cerevisiae, has a dual role in S-phase progression and at a cell cycle checkpoint." Proc Natl Acad Sci U S A **92**(25): 11791-5.

Arcus, V. (2002). "OB-fold domains: a snapshot of the evolution of sequence, structure and function." Curr Opin Struct Biol **12**(6): 794-801.

Arezi, B., B. W. Kirk, et al. (1999). "Interactions of DNA with human DNA primase monitored with photoactivatable cross-linking agents: implications for the role of the p58 subunit." Biochemistry **38**(39): 12899-907.

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

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

Arunkumar, A. I., V. Klimovich, et al. (2005). "Insights into hRPA32 C-terminal domain--mediated assembly of the simian virus 40 replisome." Nat Struct Mol Biol **12**(4): 332-9.

Arunkumar, A. I., M. E. Stauffer, et al. (2003). "Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains." J Biol Chem **278**(42): 41077-82.

Baker, T. A. and S. P. Bell (1998). "Polymerases and the replisome: machines within machines." Cell **92**(3): 295-305.

Ball, H. L., M. R. Ehrhardt, et al. (2007). "Function of a conserved checkpoint recruitment domain in ATRIP proteins." Mol Cell Biol **27**(9): 3367-77.

Ball, H. L., J. S. Myers, et al. (2005). "ATRIP binding to replication protein A-single-stranded DNA promotes ATR-ATRIP localization but is dispensable for Chk1 phosphorylation." Mol Biol Cell **16**(5): 2372-81.

Bell, C. E. and M. Lewis (2001). "Crystallographic analysis of Lac repressor bound to natural operator O1." J Mol Biol **312**(5): 921-6.

Bell, S. P. (1995). "Eukaryotic replicators and associated protein complexes." Curr Opin Genet Dev **5**(2): 162-7.

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

Berger, B., D. B. Wilson, et al. (1995). "Predicting coiled coils by use of pairwise residue correlations." Proc Natl Acad Sci U S A **92**(18): 8259-63.

Berthet-Colominas, C., L. Seignovert, et al. (1998). "The crystal structure of asparaginyl-tRNA synthetase from Thermus thermophilus and its complexes with ATP and asparaginyl-adenylate: the mechanism of discrimination between asparagine and aspartic acid." Embo J **17**(10): 2947-60.

Bhattacharya, S., A. I. Arunkumar, et al. (2004). "1H, 13C and 15N assignments of single-stranded DNA binding domains from the 70 kDa subunit of human replication protein A." J Biomol NMR **28**(2): 195-6.

Bhattacharya, S., M. V. Botuyan, et al. (2002). "Characterization of binding-induced changes in dynamics suggests a model for sequence-nonspecific binding of ssDNA by replication protein A." Protein Sci **11**(10): 2316-25.

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

Blow, J. J., P. J. Gillespie, et al. (2001). "Replication origins in Xenopus egg extract Are 5-15 kilobases apart and are activated in clusters that fire at different times." J Cell Biol **152**(1): 15-25.

Bochkarev, A., R. A. Pfuetzner, et al. (1997). "Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA." Nature **385**(6612): 176-81.

Bochkareva, E., V. Belegu, et al. (2001). "Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding." Embo J **20**(3): 612-8.

Bochkareva, E., L. Kaustov, et al. (2005). "Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A." Proc Natl Acad Sci U S A **102**(43): 15412-7.

Bochkareva, E., S. Korolev, et al. (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 A. Schwacha (2008). "The Mcm2-7 complex has in vitro helicase activity." Mol Cell **31**(2): 287-93.

Bogden, C. E., D. Fass, et al. (1999). "The structural basis for terminator recognition by the Rho transcription termination factor." Mol Cell **3**(4): 487-93.

Borden, K. L. (2000). "RING domains: master builders of molecular scaffolds?" J Mol Biol **295**(5): 1103-12.

Borowiec, J. A. and J. Hurwitz (1988). "Localized melting and structural changes in the SV40 origin of replication induced by T-antigen." Embo J **7**(10): 3149-58.

Bowles, T., A. H. Metz, et al. (2008). "Structure and DNA binding of alkylation response protein AidB." Proc Natl Acad Sci U S A **105**(40): 15299-304.

Brownlie, P., T. Ceska, et al. (1997). "The crystal structure of an intact human Max-DNA complex: new insights into mechanisms of transcriptional control." Structure **5**(4): 509-20.

Brunger, A. T., P. D. Adams, et al. (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." Acta Crystallogr D Biol Crystallogr **54 (Pt 5)**: 905-21.

Bruning, J. B. and Y. Shamoo (2004). "Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1." Structure **12**(12): 2209-19.

Bullock, P. A. (1997). "The initiation of simian virus 40 DNA replication in vitro." Crit Rev Biochem Mol Biol **32**(6): 503-68.

Byun, T. S., M. Pacek, et al. (2005). "Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint." Genes Dev **19**(9): 1040-52.

Cavarelli, J., G. Eriani, et al. (1994). "The active site of yeast aspartyl-tRNA synthetase: structural and functional aspects of the aminoacylation reaction." Embo J **13**(2): 327-37.

Challberg, M. D. and T. J. Kelly (1989). "Animal virus DNA replication." Annu Rev Biochem **58**: 671-717.

Chapados, B. R., D. J. Hosfield, et al. (2004). "Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair." Cell **116**(1): 39-50.

Chattopadhyay, S. and A. K. Bielinsky (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, R. and Z. Weng (2002). "Docking unbound proteins using shape complementarity, desolvation, and electrostatics." Proteins **47**(3): 281-94.

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

Clyne, R. K. and T. J. Kelly (1995). "Genetic analysis of an ARS element from the fission yeast Schizosaccharomyces pombe." Embo J **14**(24): 6348-57.

Collaborative Computational Project, N. (1994). "The CCP4 suite: programs for protein crystallography." Acta Crystallographica **D50**: 760-763.

Collins, K. L. and T. J. Kelly (1991). "Effects of T antigen and replication protein A on the initiation of DNA synthesis by DNA polymerase alpha-primase." Mol Cell Biol **11**(4): 2108-15.

Collins, K. L., A. A. Russo, et al. (1993). "The role of the 70 kDa subunit of human DNA polymerase alpha in DNA replication." Embo J **12**(12): 4555-66.

Conaway, R. C. and I. R. Lehman (1982). "A DNA primase activity associated with DNA polymerase alpha from Drosophila melanogaster embryos." Proc Natl Acad Sci U S A **79**(8): 2523-7.

Conaway, R. C. and I. R. Lehman (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., G. Kung, et al. (2003). "A novel zinc finger is required for Mcm10 homocomplex assembly." J Biol Chem **278**(38): 36051-8.

Copeland, W. C. (1997). "Expression, purification, and characterization of the two human primase subunits and truncated complexes from Escherichia coli." Protein Expr Purif **9**(1): 1-9.

D'Urso, G., B. Grallert, et al. (1995). "DNA polymerase alpha, a component of the replication initiation complex, is essential for the checkpoint coupling S phase to mitosis in fission yeast." J Cell Sci **108 (Pt 9)**: 3109-18.

D.A. Case, T. A. D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M. Merz, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C. Schafmeister, W.S. Ross, and P.A. Kollman (2006). AMBER 9, University of California, San Francisco.

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

Daughdrill, G. W., G. W. Buchko, et al. (2003). "Chemical shift changes provide evidence for overlapping single-stranded DNA- and XPA-binding sites on the 70 kDa subunit of human replication protein A." Nucleic Acids Res **31**(14): 4176-83.

Davey, M. J. and M. O'Donnell (2000). "Mechanisms of DNA replication." Curr Opin Chem Biol **4**(5): 581-6.

Dean, F. B., M. Dodson, et al. (1987). "ATP-dependent formation of a specialized nucleoprotein structure by simian virus 40 (SV40) large tumor antigen at the SV40 replication origin." Proc Natl Acad Sci U S A **84**(24): 8981-5.

DePamphilis, M. L., Ed. (2006). DNA Replication and Human Disease. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Dore, A. S., M. L. Kilkenny, et al. (2006). "Structure of an archaeal PCNA1-PCNA2-FEN1 complex: elucidating PCNA subunit and client enzyme specificity." Nucleic Acids Res **34**(16): 4515-26.

Dore, A. S., M. L. Kilkenny, et al. (2009). "Crystal structure of the rad9-rad1-hus1 DNA damage checkpoint complex--implications for clamp loading and regulation." Mol Cell **34**(6): 735-45.

Dornreiter, I., W. C. Copeland, et al. (1993). "Initiation of simian virus 40 DNA replication requires the interaction of a specific domain of human DNA polymerase alpha with large T antigen." Mol Cell Biol **13**(2): 809-20.

Dornreiter, I., L. F. Erdile, et al. (1992). "Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen." Embo J **11**(2): 769-76.

Dornreiter, I., A. Hoss, et al. (1990). "SV40 T antigen binds directly to the large subunit of purified DNA polymerase alpha." Embo J **9**(10): 3329-36.

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

Eiler, S., A. Dock-Bregeon, et al. (1999). "Synthesis of aspartyl-tRNA(Asp) in Escherichia coli--a snapshot of the second step." Embo J **18**(22): 6532-41.
Eliasson, R. and P. Reichard (1978). "Replication of polyoma DNA in isolated nuclei. Synthesis and distribution of initiator RNA." J Biol Chem **253**(20): 7469-75.
Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-32.
Eyck, L. F. T., J. Mandell, et al. (1995). "Surveying molecular interactions with DOT." In Proceedings of the 1995 ACM/IEEE Supercomputing Conference: 506--517.
Fanning, E., V. Klimovich, et al. (2006). "A dynamic model for replication protein A (RPA) function in DNA processing pathways." Nucleic Acids Res **34**(15): 4126-37.
Fanning, E. and R. Knippers (1992). "Structure and function of simian virus 40 large tumor antigen." Annu Rev Biochem **61**: 55-85.
Fien, K., Y. S. Cho, et al. (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 J. Hurwitz (2006). "Fission yeast Mcm10p contains primase activity." J Biol Chem **281**(31): 22248-60.
Finerty, P. J., Jr. and B. L. Bass (1999). "Subsets of the zinc finger motifs in dsRBP-ZFa can bind double-stranded RNA." Biochemistry **38**(13): 4001-7.
Fletcher, R. J., B. E. Bishop, et al. (2003). "The structure and function of MCM from archaeal M. Thermoautotrophicum." Nat Struct Biol **10**(3): 160-7.
Foiani, M., F. Marini, et al. (1994). "The B subunit of the DNA polymerase alpha-primase complex in Saccharomyces cerevisiae executes an essential function at the initial stage of DNA replication." Mol Cell Biol **14**(2): 923-33.
Francis, L. I., J. C. Randell, et al. (2009). "Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation." Genes Dev **23**(5): 643-54.
Gai, D., D. Li, et al. (2004). "Insights into the oligomeric states, conformational changes, and helicase activities of SV40 large tumor antigen." J Biol Chem **279**(37): 38952-9.
Gai, D., R. Zhao, et al. (2004). "Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen." Cell **119**(1): 47-60.
Gambus, A., R. C. Jones, et al. (2006). "GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks." Nat Cell Biol **8**(4): 358-66.
Garcia de la Torre, J. G. and V. A. Bloomfield (1981). "Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications." Q Rev Biophys **14**(1): 81-139.
Garg, P. and P. M. Burgers (2005). "DNA polymerases that propagate the eukaryotic DNA replication fork." Crit Rev Biochem Mol Biol **40**(2): 115-28.
Gilbert, D. M. (2001). "Making sense of eukaryotic DNA replication origins." Science **294**(5540): 96-100.
Ginalski, K., A. Elofsson, et al. (2003). "3D-Jury: a simple approach to improve protein structure predictions." Bioinformatics **19**(8): 1015-8.
Gouet, P., E. Courcelle, et al. (1999). "ESPript: analysis of multiple sequence alignments in PostScript." Bioinformatics **15**(4): 305-8.

Gozuacik, D., M. Chami, et al. (2003). "Identification and functional characterization of a new member of the human Mcm protein family: hMcm8." Nucleic Acids Res **31**(2): 570-9.
Grallert, B. and P. Nurse (1997). "An approach to identify functional homologues and suppressors of genes in fission yeast." Curr Genet **32**(1): 27-31.
Gray, J. J., S. E. Moughon, et al. (2003). "Protein-protein docking predictions for the CAPRI experiment." Proteins **52**(1): 118-22.
Gregan, J., K. Lindner, et al. (2003). "Fission yeast Cdc23/Mcm10 functions after pre-replicative complex formation to promote Cdc45 chromatin binding." Mol Biol Cell **14**(9): 3876-87.
Hall, D. R., D. G. Gourley, et al. (1999). "The high-resolution crystal structure of the molybdate-dependent transcriptional regulator (ModE) from Escherichia coli: a novel combination of domain folds." Embo J **18**(6): 1435-46.
Hart, E. A., J. A. Bryant, et al. (2002). "Fission yeast Cdc23 interactions with DNA replication initiation proteins." Curr Genet **41**(5): 342-8.
Heikinheimo, P., J. Lehtonen, et al. (1996). "The structural basis for pyrophosphatase catalysis." Structure **4**(12): 1491-508.
Hennessy, K. M., A. Lee, et al. (1991). "A group of interacting yeast DNA replication genes." Genes Dev **5**(6): 958-69.
Herendeen, D. R. and T. J. Kelly (1996). "DNA polymerase III: running rings around the fork." Cell **84**(1): 5-8.
Holm, L. and C. Sander (1993). "Protein structure comparison by alignment of distance matrices." J Mol Biol **233**(1): 123-38.
Homesley, L., M. Lei, et al. (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.
Hsi, K. L., W. C. Copeland, et al. (1990). "Human DNA polymerase alpha catalytic polypeptide binds ConA and RCA and contains a specific labile site in the N-terminus." Nucleic Acids Res **18**(21): 6231-7.
Huang, S. G., K. Weisshart, et al. (1998). "Stoichiometry and mechanism of assembly of SV40 T antigen complexes with the viral origin of DNA replication and DNA polymerase alpha-primase." Biochemistry **37**(44): 15345-52.
Hubscher, U., G. Maga, et al. (2002). "Eukaryotic DNA polymerases." Annu Rev Biochem **71**: 133-63.
Hudson, B. P., M. A. Martinez-Yamout, et al. (2004). "Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d." Nat Struct Mol Biol **11**(3): 257-64.
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., K. Yanagi, et al. (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., F. Yatagai, et al. (2001). "Cell cycle-dependent proteolysis and phosphorylation of human Mcm10." J Biol Chem **276**(51): 48526-31.
Izumi, M., F. Yatagai, et al. (2004). "Localization of human Mcm10 is spatially and temporally regulated during S phase." J Biol Chem.

Jacob F, B. S., Cuzin F. (1963). Cold Spring Harbor Symp. Quant. Biol. , Cold Spring Harbor, NY.

Jiang, X., V. Klimovich, et al. (2006). "Structural mechanism of RPA loading on DNA during activation of a simple pre-replication complex." Embo J **25**(23): 5516-26.

Kadrmas, J. L. and M. C. Beckerle (2004). "The LIM domain: from the cytoskeleton to the nucleus." Nat Rev Mol Cell Biol **5**(11): 920-31.

Kamimura, Y., Y. S. Tak, et al. (2001). "Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae." Embo J **20**(8): 2097-107.

Kawasaki, Y., S. Hiraga, et al. (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.

Kearsey, S. E. and K. Labib (1998). "MCM proteins: evolution, properties, and role in DNA replication." Biochim Biophys Acta **1398**(2): 113-36.

Kelley, L., R. MacCallum, et al. (2000). "Enhanced genome annotation using structural profiles in the program 3D-PSSM." J. Mol. Biol **299**: 499-520.

Kelley, L. A. and M. J. Sternberg (2009). "Protein structure prediction on the Web: a case study using the Phyre server." Nat Protoc **4**(3): 363-71.

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

Kinoshita, Y., E. M. Johnson, et al. (2008). "Colocalization of MCM8 and MCM7 with proteins involved in distinct aspects of DNA replication." Microsc Res Tech **71**(4): 288-97.

Klein, D. J., T. M. Schmeing, et al. (2001). "The kink-turn: a new RNA secondary structure motif." Embo J **20**(15): 4214-21.

Klug, A. and J. W. Schwabe (1995). "Protein motifs 5. Zinc fingers." Faseb J **9**(8): 597-604.

Kontopidis, G., S. Y. Wu, et al. (2005). "Structural and biochemical studies of human proliferating cell nuclear antigen complexes provide a rationale for cyclin association and inhibitor design." Proc Natl Acad Sci U S A **102**(6): 1871-6.

Kostic, M., T. Matt, et al. (2006). "Solution structure of the Hdm2 C2H2C4 RING, a domain critical for ubiquitination of p53." J Mol Biol **363**(2): 433-50.

Kowalczykowski, S. C. (2000). "Some assembly required." Nat Struct Biol **7**(12): 1087-9.

Krishna, S. S., I. Majumdar, et al. (2003). "Structural classification of zinc fingers: survey and summary." Nucleic Acids Res **31**(2): 532-50.

Kubota, Y., Y. Takase, et al. (2003). "A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication." Genes Dev **17**(9): 1141-52.

Kukimoto, I., H. Igaki, et al. (1999). "Human CDC45 protein binds to minichromosome maintenance 7 protein and the p70 subunit of DNA polymerase alpha." Eur J Biochem **265**(3): 936-43.

Kunkel, T. A. and P. M. Burgers (2008). "Dividing the workload at a eukaryotic replication fork." Trends Cell Biol **18**(11): 521-7.

Kwon, K., C. Cao, et al. (2003). "A novel zinc snap motif conveys structural stability to 3-methyladenine DNA glycosylase I." J Biol Chem **278**(21): 19442-6.

Labib, K. and J. F. Diffley (2001). "Is the MCM2-7 complex the eukaryotic DNA replication fork helicase?" Curr Opin Genet Dev **11**(1): 64-70.

Lao, Y., C. G. Lee, et al. (1999). "Replication protein A interactions with DNA. 2. Characterization of double-stranded DNA-binding/helix-destabilization activities and the role of the zinc-finger domain in DNA interactions." Biochemistry **38**(13): 3974-84.

LaRocque, J. R., D. L. Dougherty, et al. (2007). "Reducing DNA polymerase alpha in the absence of Drosophila ATR leads to P53-dependent apoptosis and developmental defects." Genetics **176**(3): 1441-51.

Laskowski, R. A., M. W. MacArthur, et al. (1993). "PROCHECK: a program to check the stereochemical quality of protein structures." Journal of Applied Crystallography **26**(2): 283-291.

Laue, T. M., Shah, B., Ridgeway, T.M., and Pelletier, S.L. (1992). SEDNTERP, Royal Society of Chemistry, UK.

Lee, J. H., C. J. Park, et al. (2003). "NMR study on the interaction between RPA and DNA decamer containing cis-syn cyclobutane pyrimidine dimer in the presence of XPA: implication for damage verification and strand-specific dual incision in nucleotide excision repair." Nucleic Acids Res **31**(16): 4747-54.

Lee, J. K., Y. S. Seo, et al. (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., C. Chang, et al. (2000). "Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications." Embo J **19**(5): 1119-29.

Lee, T. T., S. Agarwalla, et al. (2005). "A unique RNA Fold in the RumA-RNA-cofactor ternary complex contributes to substrate selectivity and enzymatic function." Cell **120**(5): 599-611.

Lei, M., Y. Kawasaki, et al. (1997). "Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis." Genes Dev **11**(24): 3365-74.

Lei, M. and B. K. Tye (2001). "Initiating DNA synthesis: from recruiting to activating the MCM complex." J Cell Sci **114**(Pt 8): 1447-54.

Leon, O. and M. Roth (2000). "Zinc fingers: DNA binding and protein-protein interactions." Biol Res **33**(1): 21-30.

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

Liu, Q., W. Choe, et al. (2000). "Identification of the Xenopus laevis homolog of Saccharomyces cerevisiae DNA2 and its role in DNA replication." J Biol Chem **275**(3): 1615-24.

Liu, Y., T. A. Richards, et al. (2009). "Ancient diversification of eukaryotic MCM DNA replication proteins." BMC Evol Biol **9**: 60.

Luft, J. R., R. J. Collins, et al. (2003). "A deliberate approach to screening for initial crystallization conditions of biological macromolecules." J Struct Biol **142**(1): 170-9.

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

Maine, G. T., P. Sinha, et al. (1984). "Mutants of S. cerevisiae defective in the maintenance of minichromosomes." Genetics **106**(3): 365-85.
Maiorano, D., O. Cuvier, et al. (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., M. Lutzmann, et al. (2006). "MCM proteins and DNA replication." Curr Opin Cell Biol **18**(2): 130-6.
Maiorano, D., J. Moreau, et al. (2000). "XCDT1 is required for the assembly of pre-replicative complexes in Xenopus laevis." Nature **404**(6778): 622-5.
Majka, J., S. K. Binz, et al. (2006). "Replication protein A directs loading of the DNA damage checkpoint clamp to 5'-DNA junctions." J Biol Chem **281**(38): 27855-61.
Mastrangelo, I. A., P. V. Hough, et al. (1989). "ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication." Nature **338**(6217): 658-62.
Masumoto, H., A. Sugino, et al. (2000). "Dpb11 controls the association between DNA polymerases alpha and epsilon and the autonomously replicating sequence region of budding yeast." Mol Cell Biol **20**(8): 2809-17.
Matsumiya, S., S. Ishino, et al. (2002). "Physical interaction between proliferating cell nuclear antigen and replication factor C from Pyrococcus furiosus." Genes Cells **7**(9): 911-22.
Matthews, J. M. and M. Sunde (2002). "Zinc fingers--folds for many occasions." IUBMB Life **54**(6): 351-5.
McRee, D. E. (1999). "XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density." J Struct Biol **125**(2-3): 156-65.
Mendez, J. and B. Stillman (2003). "Perpetuating the double helix: molecular machines at eukaryotic DNA replication origins." Bioessays **25**(12): 1158-67.
Meng, E. C., E. F. Pettersen, et al. (2006). "Tools for integrated sequence-structure analysis with UCSF Chimera." BMC Bioinformatics **7**: 339.
Mer, G., A. Bochkarev, et al. (2000). "Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA." Cell **103**(3): 449-56.
Merchant, A. M., Y. Kawasaki, et al. (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.
Metz, A. H., T. Hollis, et al. (2007). "DNA damage recognition and repair by 3-methyladenine DNA glycosylase I (TAG)." Embo J **26**(9): 2411-20.
Michael, W. M., R. Ott, et al. (2000). "Activation of the DNA replication checkpoint through RNA synthesis by primase." Science **289**(5487): 2133-7.
Mimura, S. and H. Takisawa (1998). "Xenopus Cdc45-dependent loading of DNA polymerase alpha onto chromatin under the control of S-phase Cdk." Embo J **17**(19): 5699-707.
Mizuno, T., N. Ito, et al. (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.

Mizuno, T., K. Yamagishi, et al. (1999). "Molecular architecture of the mouse DNA polymerase alpha-primase complex." Mol Cell Biol **19**(11): 7886-96.
Moir, D. and D. Botstein (1982). "Determination of the order of gene function in the yeast nuclear division pathway using cs and ts mutants." Genetics **100**(4): 565-77.
Moyer, S. E., P. W. Lewis, et al. (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.
Murshudov, G. N., A. A. A.A.Vagin, et al. (1997). "Refinement of Macromolecular Structures by the Maximum-Likelihood Method." Acta Crystallographica **D53**: 240-255.
Murzin, A. G., S. E. Brenner, et al. (1995). "SCOP: a structural classification of proteins database for the investigation of sequences and structures." J Mol Biol **247**(4): 536-40.
Muzi-Falconi, M., M. Giannattasio, et al. (2003). "The DNA polymerase alpha-primase complex: multiple functions and interactions." ScientificWorldJournal **3**: 21-33.
Nair, P. A., J. Nandakumar, et al. (2007). "Structural basis for nick recognition by a minimal pluripotent DNA ligase." Nat Struct Mol Biol **14**(8): 770-8.
Nasheuer, H. P., A. Moore, et al. (1991). "Cell cycle-dependent phosphorylation of human DNA polymerase alpha." J Biol Chem **266**(12): 7893-903.
Nasmyth, K. and P. Nurse (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., D. A. Gordenin, et al. (2008). "Division of labor at the eukaryotic replication fork." Mol Cell **30**(2): 137-44.
Nishitani, H., Z. Lygerou, et al. (2000). "The Cdt1 protein is required to license DNA for replication in fission yeast." Nature **404**(6778): 625-8.
Okorokov, A. L., A. Waugh, et al. (2007). "Hexameric ring structure of human MCM10 DNA replication factor." EMBO Rep **8**(10): 925-30.
Ott, R. D., C. Rehfuess, et al. (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 W. Minor (1997). "Processing of x-ray diffraction data collected in oscillation mode." Methods Enzymol. **276**: 307-326.
Pacek, M., A. V. Tutter, et al. (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 J. C. Walter (2004). "A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication." Embo J **23**(18): 3667-76.
Painter, J. and E. A. Merritt (2006). "Optimal description of a protein structure in terms of multiple groups undergoing TLS motion." Acta Crystallogr D Biol Crystallogr **62**(Pt 4): 439-50.
Paixao, S., I. N. Colaluca, et al. (2004). "Modular structure of the human lamin B2 replicator." Mol Cell Biol **24**(7): 2958-67.
Park, J. H., S. W. Bang, et al. (2008). "Knockdown of human MCM10 activates G2 checkpoint pathway." Biochem Biophys Res Commun **365**(3): 490-5.

Pascal, J. M., P. J. O'Brien, et al. (2004). "Human DNA ligase I completely encircles and partially unwinds nicked DNA." Nature **432**(7016): 473-8.
Pascual, J., M. Martinez-Yamout, et al. (2000). "Structure of the PHD zinc finger from human Williams-Beuren syndrome transcription factor." J Mol Biol **304**(5): 723-9.
Pavletich, N. P. and C. O. Pabo (1991). "Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A." Science **252**(5007): 809-17.
Petroski, M. D. and R. J. Deshaies (2005). "Function and regulation of cullin-RING ubiquitin ligases." Nat Rev Mol Cell Biol **6**(1): 9-20.
Pettersen, E. F., T. D. Goddard, et al. (2004). "UCSF Chimera--a visualization system for exploratory research and analysis." J Comput Chem **25**(13): 1605-12.
Philo, J. S. (2000). "A method for directly fitting the time derivative of sedimentation velocity data and an alternative algorithm for calculating sedimentation coefficient distribution functions." Anal Biochem **279**(2): 151-63.
Plevani, P., M. Foiani, et al. (1985). "Polypeptide structure of DNA primase from a yeast DNA polymerase-primase complex." J Biol Chem **260**(11): 7102-7.
Pursell, Z. F., I. Isoz, et al. (2007). "Yeast DNA polymerase epsilon participates in leading-strand DNA replication." Science **317**(5834): 127-30.
Raghunathan, S., A. G. Kozlov, et al. (2000). "Structure of the DNA binding domain of E. coli SSB bound to ssDNA." Nat Struct Biol **7**(8): 648-52.
Ramachandran, N., E. Hainsworth, et al. (2004). "Self-assembling protein microarrays." Science **305**(5680): 86-90.
Randell, J. C., J. L. Bowers, et al. (2006). "Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase." Mol Cell **21**(1): 29-39.
Ricke, R. M. and A. K. Bielinsky (2004). "Mcm10 regulates the stability and chromatin association of DNA polymerase-alpha." Mol Cell **16**(2): 173-85.
Ricke, R. M. and A. K. Bielinsky (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.
Ritchie, D. W. and G. J. Kemp (2000). "Protein docking using spherical polar Fourier correlations." Proteins **39**(2): 178-94.
Robertson, P. D., E. M. Warren, et al. (2008). "Domain Architecture and Biochemical Characterization of Vertebrate Mcm10." J Biol Chem **283**(6): 3338-48.
Sakurai, S., K. Kitano, et al. (2005). "Structural basis for recruitment of human flap endonuclease 1 to PCNA." Embo J **24**(4): 683-93.
Sawyer, S. L., I. H. Cheng, et al. (2004). "Mcm10 and Cdc45 cooperate in origin activation in Saccharomyces cerevisiae." J Mol Biol **340**(2): 195-202.
Schub, O., G. Rohaly, et al. (2001). "Multiple phosphorylation sites of DNA polymerase alpha-primase cooperate to regulate the initiation of DNA replication in vitro." J Biol Chem **276**(41): 38076-83.
Schuck, P. (2003). "On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation." Anal Biochem **320**(1): 104-24.
Shamoo, Y., A. M. Friedman, et al. (1995). "Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA." Nature **376**(6538): 362-6.

Shaw, N., W. Tempel, et al. (2008). "Crystal structure solution of a ParB-like nuclease at atomic resolution." Proteins: Structure, Function, and Bioinformatics **70**(1): 263-267.

Sheaff, R. J. and R. D. Kuchta (1994). "Misincorporation of nucleotides by calf thymus DNA primase and elongation of primers containing multiple noncognate nucleotides by DNA polymerase alpha." J Biol Chem **269**(30): 19225-31.

Sheaff, R. J., R. D. Kuchta, et al. (1994). "Calf thymus DNA polymerase alpha-primase: "communication" and primer-template movement between the two active sites." Biochemistry **33**(8): 2247-54.

Sheu, Y. J. and B. Stillman (2006). "Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression." Mol Cell **24**(1): 101-13.

Singleton, M. R., S. Scaife, et al. (2001). "Structural analysis of DNA replication fork reversal by RecG." Cell **107**(1): 79-89.

Smith, D. B. and K. S. Johnson (1988). "Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase." Gene **67**(1): 31-40.

Solomon, N. A., M. B. Wright, et al. (1992). "Genetic and molecular analysis of DNA43 and DNA52: two new cell-cycle genes in Saccharomyces cerevisiae." Yeast **8**(4): 273-89.

Spruck, C. H., K. A. Won, et al. (1999). "Deregulated cyclin E induces chromosome instability." Nature **401**(6750): 297-300.

Stafford, W. F., 3rd (1992). "Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile." Anal Biochem **203**(2): 295-301.

Stafford, W. F. and P. J. Sherwood (2004). "Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants." Biophys Chem **108**(1-3): 231-43.

Stahl, H., P. Droge, et al. (1986). "DNA helicase activity of SV40 large tumor antigen." Embo J **5**(8): 1939-44.

Stauffer, M. E. and W. J. Chazin (2004). "Physical interaction between replication protein A and Rad51 promotes exchange on single-stranded DNA." J Biol Chem **279**(24): 25638-45.

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

Stein, P. E., A. Boodhoo, et al. (1994). "The crystal structure of pertussis toxin." Structure **2**(1): 45-57.

Stillman, B. (1989). "Initiation of eukaryotic DNA replication in vitro." Annu Rev Cell Biol **5**: 197-245.

Storoni, L. C., A. J. McCoy, et al. (2004). "Likelihood-enhanced fast rotation functions." Acta Crystallogr D Biol Crystallogr **60**(Pt 3): 432-8.

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

Takahashi, T. S., D. B. Wigley, et al. (2005). "Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase." Trends Biochem Sci **30**(8): 437-44.

Takayama, Y., Y. Kamimura, et al. (2003). "GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast." Genes Dev **17**(9): 1153-65.

Takemura, M., S. Yoshida, et al. (2006). "Role of the second-largest subunit of DNA polymerase alpha in the interaction between the catalytic subunit and hyperphosphorylated retinoblastoma protein in late S phase." Biochim Biophys Acta **1764**(9): 1447-53.

Tanaka, S. and J. F. Diffley (2002). "Deregulated G1-cyclin expression induces genomic instability by preventing efficient pre-RC formation." Genes Dev **16**(20): 2639-49.

Tanaka, S., Y. S. Tak, et al. (2007). "The role of CDK in the initiation step of DNA replication in eukaryotes." Cell Div **2**: 16.

Tanaka, S., T. Umemori, et al. (2007). "CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast." Nature **445**(7125): 328-32.

Tanaka, T. and K. Nasmyth (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.

Taneja, P., H. P. Nasheuer, et al. (2007). "Timed interactions between viral and cellular replication factors during the initiation of SV40 in vitro DNA replication." Biochem J **407**(2): 313-20.

Theobald, D. L., R. M. Mitton-Fry, et al. (2003). "Nucleic acid recognition by OB-fold proteins." Annu Rev Biophys Biomol Struct **32**: 115-33.

Thompson, J. D., D. G. Higgins, et al. (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res **22**(22): 4673-80.

Tsurimoto, T., T. Melendy, et al. (1990). "Sequential initiation of lagging and leading strand synthesis by two different polymerase complexes at the SV40 DNA replication origin." Nature **346**(6284): 534-9.

Tye, B. K. (1999). "MCM proteins in DNA replication." Annu Rev Biochem **68**: 649-86.

Uchiyama, M. and T. S. Wang (2004). "The B-subunit of DNA polymerase alpha-primase associates with the origin recognition complex for initiation of DNA replication." Mol Cell Biol **24**(17): 7419-34.

Uson, I. and G. M. Sheldrick (1999). "Advances in direct methods for protein crystallography." Curr Opin Struct Biol **9**(5): 643-8.

Vagin, A. and A. Teplyakov (1997). MOLREP: an Automated Program for Molecular Replacement. Journal of Applied Crystallography. **30:** 1022-1025.

Vaughn, J. P., P. A. Dijkwel, et al. (1990). "Replication initiates in a broad zone in the amplified CHO dihydrofolate reductase domain." Cell **61**(6): 1075-87.

Vaziri, C., S. Saxena, et al. (2003). "A p53-dependent checkpoint pathway prevents rereplication." Mol Cell **11**(4): 997-1008.

Vijayakumar, S., B. R. Chapados, et al. (2007). "The C-terminal domain of yeast PCNA is required for physical and functional interactions with Cdc9 DNA ligase." Nucleic Acids Res **35**(5): 1624-37.

Voitenleitner, C., E. Fanning, et al. (1997). "Phosphorylation of DNA polymerase alpha-primase by cyclin A-dependent kinases regulates initiation of DNA replication in vitro." Oncogene **14**(13): 1611-5.

Voitenleitner, C., C. Rehfuess, et al. (1999). "Cell cycle-dependent regulation of human DNA polymerase alpha-primase activity by phosphorylation." Mol Cell Biol **19**(1): 646-56.

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

Vonrhein, C., E. Blanc, et al. (2006). "Automated Structure Solution With autoSHARP." Methods Mol Biol **364**: 215-30.

Waga, S. and B. Stillman (1994). "Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro." Nature **369**(6477): 207-12.

Walter, J. and J. Newport (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.

Walter, J., L. Sun, et al. (1998). "Regulated chromosomal DNA replication in the absence of a nucleus." Mol Cell **1**(4): 519-29.

Walther, A. P., X. V. Gomes, et al. (1999). "Replication protein A interactions with DNA. 1. Functions of the DNA-binding and zinc-finger domains of the 70-kDa subunit." Biochemistry **38**(13): 3963-73.

Wang, T. S. (1991). "Eukaryotic DNA polymerases." Annu Rev Biochem **60**: 513-52.

Warren, E. M., Vaithiyalingam, S.R., 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): 1892-1901.

Wedel, A. and S. Kustu (1995). "The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation." Genes Dev **9**(16): 2042-52.

Williams, D. R. and J. R. McIntosh (2005). "Mcl1p is a polymerase alpha replication accessory factor important for S-phase DNA damage survival." Eukaryot Cell **4**(1): 166-77.

Wimberly, B. T., D. E. Brodersen, et al. (2000). "Structure of the 30S ribosomal subunit." Nature **407**(6802): 327-39.

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

Wolf, E., P. S. Kim, et al. (1997). "MultiCoil: a program for predicting two- and three-stranded coiled coils." Protein Sci **6**(6): 1179-89.

Wriggers, W., R. A. Milligan, et al. (1999). "Situs: A package for docking crystal structures into low-resolution maps from electron microscopy." J Struct Biol **125**(2-3): 185-95.

Xu, X., S. Vaithiyalingam, et al. (2008). "The basic cleft of RPA70N binds multiple checkpoint proteins, including RAD9, to regulate ATR signaling." Mol Cell Biol **28**(24): 7345-53.

Yabuuchi, H., Y. Yamada, et al. (2006). "Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins." Embo J **25**(19): 4663-74.

Yang, X., J. Gregan, et al. (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.

Yoshida, K. (2005). "Identification of a novel cell-cycle-induced MCM family protein MCM9." Biochem Biophys Res Commun **331**(2): 669-74.

Yuzhakov, A., Z. Kelman, et al. (1999). "Multiple competition reactions for RPA order the assembly of the DNA polymerase delta holoenzyme." Embo J **18**(21): 6189-99.

Yuzhakov, A., Z. Kelman, et al. (1999). "Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase." Cell **96**(1): 153-63.

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

Zheng, N., B. A. Schulman, et al. (2002). "Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex." Nature **416**(6882): 703-9.

Zhu, W., C. Ukomadu, et al. (2007). "Mcm10 and And-1/CTF4 recruit DNA polymerase alpha to chromatin for initiation of DNA replication." Genes Dev **21**(18): 2288-99.

Zou, L. and S. J. Elledge (2003). "Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes." Science **300**(5625): 1542-8.

Zou, L. and B. Stillman (1998). "Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin." Science **280**(5363): 593-6.

Zou, L. and B. Stillman (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.

Zuccola, H. J., D. J. Filman, et al. (2000). "The crystal structure of an unusual processivity factor, herpes simplex virus UL42, bound to the C terminus of its cognate polymerase." Mol Cell **5**(2): 267-78.