Modulation of MCM helicase activity

Rabab Satti

PhD

University of York

Biology

October 2011

Abstract:

The minichromosome maintenance (MCM) proteins are essential for the initiation and progression of DNA replication. MCMs are believed to function as the replicative helicase in archaea and eukaryotes. In the present studies, human MCM and archaeal MCM from *Methanococcus maripaludis* were purified from *E.coli* and assembled as complexes. The six MCMs from human co-purified over multiples steps including three chromatographic steps when co-expressed in *E.coli* suggesting that they form a heterohexamer complex. Reconstitution of *in vitro* helicase activity of the human complex and individual subunits of *M. maripaludis* were described. Furthermore the effect of phosphorylation on human MCM complex and *M. maripaludis* MCM A were also tested and the effect of phosphorylation was shown to be inhibitory on helicase activities. Different reaction conditions were monitored on MCM helicase activity and the most interesting was the effect of salts on helicase assays which showed that glutamate has a significant positive effect while chloride ions inhibit helicase activity. Putative kinases were purified from *M. maripaludis* and *Methanocaldococcus jannaschii* and trials were conducted to assign them a role in regulating the DNA initiation process.

TABLE OF CONTENTS:

LIST OF TABLES:

LIST OF FIGURES:

ACKNOWLEDGEMENTS:

I would firstly like to thank my supervisor, Dr James Chong for his guidance and support throughout my PhD. I would like to extend my thanks to past and present colleagues of the Chong lab, especially Richard Parker for his help in protein purification experiments.

I would also like to thank Prof. C. I. Edvard Smith, Dr Beston F. Nore and Alamdar Baloch for hosting me in Karolinska Institute for seven weeks during which I have learned a lot of science.

Ahsan has played the role of an angel sent by God. Without his mathematical and computational skills, my thesis in its current form might not have been the same. Special thanks to Afzal for valuable tips in Adobe programming. I had certainly enjoyed my experimental work with Dr Abbas, Chitvan and Lina. Thanks to Prof. Tony Wilkinson and Dr Daniela Barilla as members of my training committee. No words will suffice for me to express my gratitude to Julie Knox for her valuable advices. And thanks to Hilary Jones for being so supportive especially during the time of failure.

On a personal level I would like to thank my friends Noreen, Shamshad, Maya, Shujaat, Ian, Dr Dzmitry, Seema, Hamda, Sumra, Mai, Nimra, Tahreem, Dr Anila, and Aiden for their encouragement over the years. Without you it would have been so lonely. The most importantly I want to thanks my family (Mum, Dad, Manny, Seemab, Deeya, Ashi, Mustehsin and Mohsin) for their constant love and support. A thanks also goes to Ashaz for keeping me sane and entertained throughout my writing up journey. In fact, I am obliged to many people who helped me to reach this goal. The list is too numerous to mention everyone's name.

This project was funded by the HEC Pakistan PhD studentship.

AUTHOR'S DECLARATION:

The work in this thesis is entirely my own, except for the human MCM co-purification that was carried out in conjunction with Dr Richard Parker. None of the work described herein has been presented previously.

Chapter 1: Introduction

1.1 DNA replication overview

DNA replication is an essential process in all living organisms. The faithful transmission of genetic material requires precise coordination and regulation. DNA replication is a biological process that occurs in all living organisms. In this process DNA molecules must duplicate accurately and be segregated to daughter cells (Albers et al., 2002). This process is initiated at particular points in the DNA, known as origins, which are targeted by proteins that separate the two strands and initiate DNA synthesis (Berg et al., 2002). Origins contain DNA sequences recognized by replication initiator proteins for examples DnaA in *E. coli* and the origin recognition complex in yeast (Weigel et al., 1997). These initiator proteins recruit other proteins to separate the two strands and initiate replication forks. To dissect this multistep process of DNA replication, it is important to find out all the proteins involved in this process and the role of each protein in each of these steps. There are a number of proteins already known that are involved in this process but still there are gaps in our knowledge of all proteins and their functions. These proteins are different in the three different domain of life. The work of Carl Woese in late 1970s (Woese and Fox, 1977) led to the biological classification of life into three domains that divide cellular life forms into bacteria, archaea and eukarya (Figure 1.1). The realization that living organisms can be divided into three main lineages, led to a quest to discover the molecular fundamentals that distinguish them.

1.1.1 Bacterial replication

The replicon hypothesis postulates two basic elements for the initiation of replication: the initiator, a trans-acting substance, DnaA protein, and the cis-acting replicator, which we now call the replication origin, oriC (Jacob and Brenner, 1963). The initiator protein DnaA is found in all bacteria analyzed so far. The bacterial chromosome typically contains a single origin of replication *oriC.* Binding of DnaA to such an array, origin recognition, is the

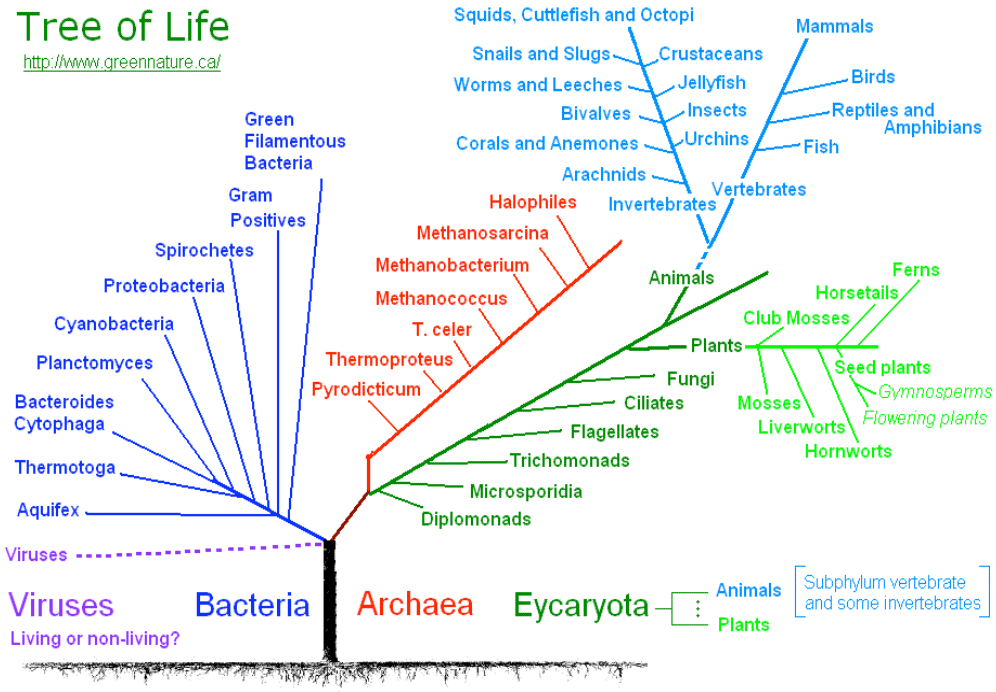


Figure 1.1 Tree of lif: This illustration of the three domains of life is like a family tree. The branches that are very far apart are not very similar. Branches that are near each other indicate closely related groups. This figure is taken from http://www.greennature.ca.

first step in the assembly of a specialized nucleoprotein complex (Echols, 1990), the initiation complex. Most of our understanding on bacterial replication comes from the extensive research on *E.coli.* DnaA protein binds to its five binding sites in oriC as a monomer, introducing a 40^{o} bend at each site (Schaper and Messer, 1995). Only DnaA complexed with ATP is active in initiation (Sekimizu et al., 1987). ATP-DnaA binds to additional sites, 6-mer ATP-DnaA boxes with the sequence 5'P-AGATCT or a close match (Speck et al., 1999; Speck and Messer, 2001). These sites are predominant in the AT-rich region, which is unwound, and subsequently stabilized due to specific binding of ATP-DnaA to single-stranded ATP-DnaA boxes. The unwound region spans 28 bp without and 44-46 bp with SSB (Single-strand binding protein) present (Krause and Messer, 1999). Since single-stranded DNA covered with SSB is a poor substrate for DnaB helicase, it has to be loaded with the help of DnaA. Two double hexamers of DnaB and the helicase loader DnaC, one double hexamer for each replication direction, are positioned by DnaA into the loop (Fang et al., 1999; Carr and Kaguni, 2001). DnaC leaves the complex immediately after or during loading, accompanied by ATP hydrolysis. This activates the helicase activity of DnaB (Wahle et al., 1989A, Wahle et al., 1989B). The two DnaB hexamers slide past each other in 5'-3' direction, and expand the bubble to about 65 nucleotides (Fang et al., 1999). Now primase can enter the complex and synthesize two leading strand primers.

The unwinding of double-stranded origin DNA and the loading of helicase into this region are one of the basic reactions in the cell cycle.

1.1.2 Archaeal replication

The status of the archaea as one of the three primary domains emphasizes the importance of understanding their molecular fundamentals. Basic DNA replication in the archaea is much similar to eukaryal replication. Archaea replicate their circular genome from a single DNA replication origin mostly as do bacteria, even though they may use eukaryotic-like proteins to do so (Myllykallio and Forterre, 2000).

All archaeal genomes sequenced to date contain at least one gene with homology to both ORC1 (Origin recognition complex subunit 1) and Cdc6 (Cell Division Cycle 6) (Barry and Bell, 2006). Origin Recognition Complex is a multi-subunit DNA binding complex (6 subunits) that binds in all eukaryotes in an ATP-dependent manner to origins of replication. The subunits of this complex are encoded by the ORC1, ORC2, ORC3, ORC4, ORC5 and ORC6 genes. Cdc6 is part of the pre-replicative complex (pre-RC) and is required for loading Minichromosome Maintenance (Mcm) proteins onto the DNA, an essential step in the initiation of DNA synthesis. The ORC1/Cdc6 proteins were identified to act as the origin recognition and binding proteins in archaea (Lundgren et al., 2004; Robinson et al., 2004). Structures show that the C-terminal region of ORC1/Cdc6 contains a winged-helix (WH) domain, and sequence alignments show that this is conserved throughout archaeal and eukaryotic Cdc6 proteins (Liu et al., 2000; Singleton et al., 2004). All archaeal genomes sequenced to date have at least one Mcm homologue (Majernik et al., 2004). Archaeal Mcm has shown efficient DNA binding and helicase activity *in vitro* (Kelman et al., 1999; Chong et al., 2000). For archaea, little is known about Mcm loading. There is apparently no homologue of DnaC or Cdt1 (Cyclin dependent transcript 1) in archaea, suggesting that the ORC1/Cdc6 proteins may perform the functions carried out by ORC, Cdc6, and Cdt1 in eukaryotes. Binding of archaeal ORC1/Cdc6 proteins to origins is apparently ATP independent (Robinson et al., 2004), despite the fact that they have a functional AAA+ domain (Grabowski and Kelman, 2001; Singleton et al., 2004). ATP hydrolysis by Cdc6 may be important for Mcm loading. The GINS (Go, Ichi, Nii, San) complex is essential in yeast and interacts with Mcm and CDC45 (Cell Division Cycle 45), which is an essential protein required to the initiation of DNA replication (Kubota et al., 2003; Gambus et al., 2006). More recently, it was shown that GINS is necessary for the inclusion of Mcm in replisome progression complexes, which include several replication and checkpoint proteins, (Gambus et al., 2006). An archaeal homologue of GINS was also identified, suggesting a role in pairing Mcm activity on the leading strand with primase activity on lagging strand (Marinsek et al., 2006). Despite our emerging understanding of archaeal

replication, little is known about its mechanisms of regulation.

1.1.3 Eukaryotic replication

DNA replication in eukaryotes is much more complicated than in bacteria or archaea, although there are many similar aspects. Eukaryotic cells can only initiate DNA replication at a specific point in the cell cycle. The cell-division cycle is the series of events that takes place in a cell leading to its division and duplication (replication). In eukaryotes, the budding yeast *Saccharomyces cerevisiae* DNA replication origins have been extensively studied and the best characterised eukaryotic origins (Bell and Dutta, 2002).

1.1.3.1 Stepwise assembly of replisome

In budding yeast, replication origins are bound throughout the cell cycle by a conserved six-subunit protein complex known as ORC like archeae. ORC protein complex is bound at replication origins characterised by the presence of ARS (autonomously replicating sequences) throughout the cell cycle, allowing replicative proteins access to the ARS (Huberman et al., 1987; Fangman and Brewer, 1991). During the G1 -phase of the cell cycle, ORC is bound by the Cdc6 protein, which then recruits two additional factors, Cdt1 and the Mcm helicase, to form the pre-RC (pre-replicative complex) (Diffley et al., 1994; Speck et al., 2005; Chen et al., 2007). Assembly of pre-RC takes place in late M and early G1 phase when the cyclin dependent kinase (CDK) activity is low. The high CDK activity prevents pre-RC assembly by phosphorylating components of pre-RC and inhibiting their activities (Diffley, 2004). Subsequent to pre-RC assembly, replication initiation is promoted by the action of S-CDK (S-phase CDK) and a second protein kinase, DDK (Cdc7-Dbf4) (Dutta, and Bell, 1997; Bell and Dutta, 2002). S-CDK main substrates in yeast are two essential proteins, Sld2 and Sld3, (Tanaka et al, 2007; Zegerman, and Diffley, 2007). Phosphorylation of Sld3 by S-CDK blocks entry into S-phase. The effect of the phosphorylation of Sld3 is to generate a binding site for another essential replication factor, Dpb11 (Tanaka et al., 2007; Zegerman and Diffley, 2007) that is equivalent

to TOPBP1 in human cells. The N-terminal Dpb11 BRCT domain is responsible for binding to phosphorylated Sld3. The importance of phosphorylation in mediating Dpb11 binding is shown by the fact that fusion of non-phosphorylatable Sld3 to Dpb11 produces a protein that is able to rescue loss-of-function mutations in both genes and partially bypass the requirement for S-CDK activity to initiate S-phase (Zegerman and Diffley, 2010). Sld3 also associates with another essential replication factor, Cdc45, which is likely to be recruited to the pre-RC via its interactions with Mcm (Figure 1.2).

The second essential S-CDK substrate in yeast is Sld2 (equivalent to RecQL4 in human cells) (Tanaka et al., 2007; Zegerman and Diffley, 2007). This protein is phosphorylated on multiple sites by S-CDK, but Thr84 phosphorylation is crucial for Sld2 function (Masumoto et al., 2002). The effect of Thr84 phosphorylation is also to create a binding site for Dpb11, via the C-terminal Dpb11 BRCT domains. Sld2 is reportedly (Tanaka et al., 2007) part of a complex (termed the pre-loading complex or pre-LC) that includes Dpb11, the leading strand polymerase Polε and GINS, suggesting that the function of Dpb11 in bridging Cdc45–Sld3 with Sld2 is to recruit the pre-LC components to the origin to activate the Mcm helicase (Figure 1.2).

Once the replisome is activated, the replication forks move off bidirectionally from the origin DNA into the surrounding sequences. ORC remains bound at the origin, apparently intact, at least in yeast cells, whereas Mcm, Cdc45, GINS and Polε move with the fork (Aparicio et al., 1997; Aparicio, et al., 1999; Kanemaki et al., 2003; Takayama, et al., 2003; Kanemaki, and Labib, 2006). The remainder of the factors described above dissociate from the replisome and in some cases are degraded or excluded from the nucleus.

Unlike yeast or bacterial origins, metazoan origins appear to have very low ARS activity in general even though some origins show ARS activity (Ariizumi et al., 1993; Taira et al., 1994; Berberich et al., 1995). Replication initiation sites appear to occur within particular regions of chromosomes

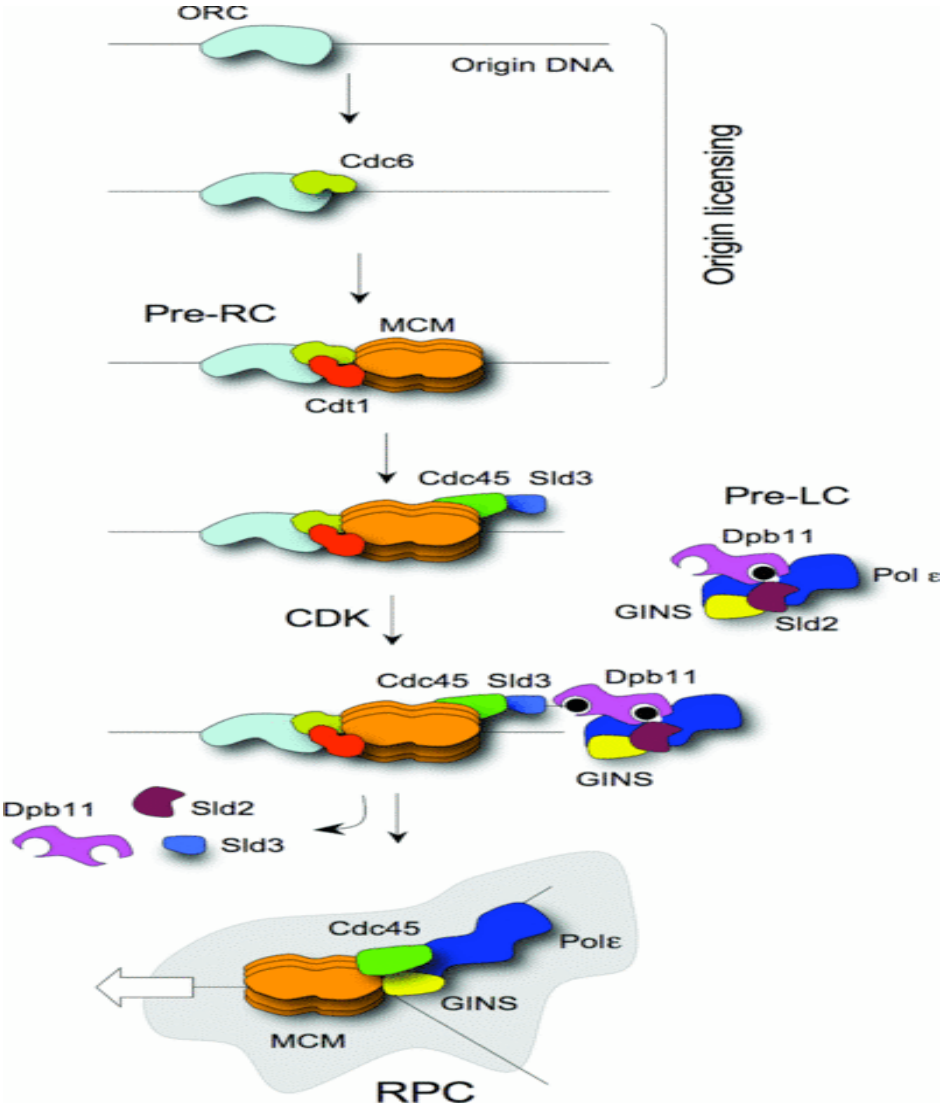


Figure 1.2 Model for stepwise replisome assembly in budding yeast: In G_1, Cdc6 binds to ORC and recruits the Mcm helicase and Cdt1 to form the pre-RC. Cdc45–Sld3 appears to associate with pre-RC-bound origins in G_1. Sld3 is phosphorylated by S-CDK, allowing it to bind to Dpb11 and recruit the pre-LC. The pre-LC contains Dpb11, Sld2, polymerase Pol ε and the GINS. Cdc45, Mcm and GINS form a tight complex (CMG) that moves with the replication fork. Assembled around a single CMG complex is the multi-protein replisome progression complex (RPC). The replisome is shown moving to the left (indicated by block arrow): Pol ε is shown synthesizing the leading strand (MacNeill, 2010).

called “initiation zones” in higher eukaryotes (Hamlin et al., 2008).

A cartoon representation comparing the important factors involved in replication initiation in the three domain of life is summarized in table 1.1

1.2. Helicases as essential replication factors

Helicases are motor proteins that move directionally along a nucleic acid phosphodiester backbone, separating two annealed nucleic acid strands (i.e., DNA, RNA, or RNA-DNA hybrid) using energy derived from ATP hydrolysis (Gorbalenya et al., 1988). Helicases are present in all domains of life and play important roles in DNA replication by unwinding the duplex. A crucial step in the initiation of DNA replication involves the loading of helicase proteins onto the duplex. In *E.coli*, DnaC is the protein responsible for loading the hexameric DnaB helicase onto the origin bound by DnaA (Davey el al., 2002). The final step is the activation of DnaB helicase so that the unwinding of the duplex allows entry of the DNA polymerase machinery (Leatherwood, 1998; Mariorano and Mechali, 2002). In eukaryotes, Cdc6 and Cdt1 are responsible for loading of hexameric Mcm proteins onto the origin while very little is known about the archaeal loading machinery but the replicative helicase, the Mcm complex is conserved within archaea. All Mcms share a great deal of homology, particularly in the central domain of the protein, which spans nearly 200 amino acids and includes the NTP binding and hydrolysis domain (Hall and Matson, 1999).

1.3 The Mcm2-7 complex as a DNA helicase

The Mcm2-7 complex is the main candidate to be the DNA helicase associated with replication in archaea and eukarya, although there are still certain gaps in our knowledge about the method by which they function. Experimental work on *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and *Xenopus laevis* showed that the Mcms are nuclear proteins required for DNA replication (Gibson et al., 1990; Chen at al., 1992; Maiti and Sinha 1992; Kubota et al., 1995). Soon after the their discovery as a

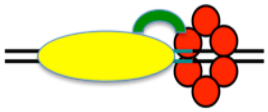
	Bacteria	Archaea	Eukaryotes
 Genomic replication origin	1	1 - 3	Many (10 - 100,000)
 Origin binding protein	DnaA	Orc / Cdc6-like	ORC1-6 Cdc6
 Helicase and loader	DnaB DnaC	MCM (usually 1)	MCM2-7 Cdt1
 Accessory proteins		GINS	Cdc45 GINS

Table 1.1 A cartoon representation of factors involved in DNA replication: Summary of important factor that involved in the process of initiation of DNA in the three domains of life.

nuclear protein in yeast, their homologues were identified in mammalian cells (Thommes at al., 1992). It was then hypothesized that Mcms could be a component of the replication licensing factor (RLF) activity defined when G1 chromatin is licensed to replicate and G2 is barred from this licensing process to prevent unscheduled DNA synthesis, which limits DNA replication to once per cycle (Blow and Laskey, 1988). This hypothesis was confirmed when the Mcm complex was purified from the RLF isolated from *Xenopus* egg extracts (Chong et al., 1995). Later six Mcm proteins were identified, given the names of Mcm2-Mcm7 and showed that they were conserved in all eukaryotes (Chong et al., 1996; Kearsey et al., 1996). Like all other helicases, the Mcm subunits contain a conserved central AAA+ motif (discussed in section 1.5.1) and the general structure of the Mcm complex resembles other DNA helicases.

It has also been found in *Xenopus* eggs that if the Mcm complex is inactivated then chromosome unwinding during S phase is inhibited (Madine at al., 2000). Early data suggested that the Mcms play an important role in DNA initiation but more recent studies show that Mcms also play a role in replication elongation, transcription, DNA repair and chromatin remodelling (Prokhorova and Blow, 2000; Labib et al., 2000; Labib et al., 2001; Forsburg, 2004). DNA helicase activity was first identified in the human Mcm4/6/7 complex of around 600 KDa containing approximately equal amounts of each Mcm suggesting a dimer of trimers. Cross-linking studies suggests that this hexamer consisting of two molecules each of Mcm4, Mcm6 and Mcm7 (Ishimi, 1997).

It is clear from a variety of experiments in different systems that Mcms associate with one another, perhaps forming a variety of complexes with differing stoichiometries. A large Mcm protein complex of around 450 – 600 kDa has been detected in extracts from budding yeast (Lei et al., 1996), fission yeast (Adachi et al., 1997), *Drosophila* (Su et al., 1997), *Xenopus* (Hendrickson et al., 1996; Kubota et al., 1997; Thommes et al., 1997) and mammalian (Kimura et al., 1996; Richter and Knippers, 1997) cell extracts. Some studies support the idea that the Mcm2-7 complex represents a hetero-

hexamer of Mcm proteins, where each Mcm type is present in equal stoichiometry. Immunoprecipitation with antibodies to a specific Mcm protein in *Xenopus* extracts precipitates all six Mcm proteins in approximately equal amounts (Kubota et al., 1997).

Other reports suggest that Mcm proteins may interact to form more heterogeneous complexes, consisting of tetramers or hexamers with different compositions (Lei et al., 1996; Su et al., 1997; Thommes et al., 1997). The Mcm proteins can be isolated as several stable subassemblies including Mcm 2/3/4/5/6/7 (hexamer), Mcm 2/4/6/7 (tetramer), Mcm 4/6/7 (trimer) and Mcm3/5 (dimer) complexes (Lei et al., 1996; Thommes et al., 1997; Ishimi, 1997; Holthoff et al., 1998; Sherman and Forsburg, 1998) but no helicase activity was detected in these complexes (Davey et al., 2003).

The Mcm2-7 hexamer (Mcm 2/3/4/5/6/7) has been shown recently to possess anionic dependent helicase activity *in vitro* (Bochman and Schwacha, 2008). *In vitro* studies using recombinant Mcm proteins from *S. pombe* indicated that Mcm2 could inhibit the helicase activity of the Mcm4/6/7 complex (Ishimi, 1997; Lee and Hurwitz, 2000). The binding of the Mcm3/5 complex to Mcm4/6/7 also inhibits the DNA helicase activity of Mcm4/6/7 complex by converting its double trimer structure into a hetero-tetramer or hetero-pentamer (Sato et al., 2000). So far two different complexes are known to play a role in unwinding dsDNA (Kanter et al., 2008), Mcm4/6/7 (Ishimi et al., 1997; You et al., 1999) and Mcm2-7 (Bochman et al., 2008). Mcm6 from pea has shown to function as a DNA helicase as a single subunit (Tran et al., 2010).

1.4 Phylogenetic analysis of Mcm

The Mcm complex was first discovered in budding yeast as mutants that were unable to maintain small chromosome-like structures (Maine et al., 1984). The complete genome sequence of *S. cerevisiae* indicates that there are six Mcm helicase encoding genes (Kearsey and Labib, 1998) and the phylogentic analysis suggested that there are six distinct classes of Mcm

proteins ranging from yeast to man (Chong et al., 1996). Identification of Mcm proteins in higher eukaryotes initially came from the detection of a murine protein related to *S. cerevisiae* Cdc 46/ Mcm5 (Hennessy, 1991) and the isolation of human P1 protein (homologous to Mcm 3) that co-purified with DNA polymerase α (Thommes et al., 1992) indicated the presence of Mcm in humans whilst the considerable sequence conservation of the family has made it easy to identify other higher eukaryotic homologues (Coxon et al., 1992; Hu et al., 1993; Sabelli et al., 1996). Higher eukaryotic Mcms have also emerged by the characterization of mRNAs or antigens that are specifically associated with proliferating cells (Todorov et al., 1994; Feger et al., 1995; Starborg et al., 1995; Sykes and Weiser, 1995) or by screening for mutants affecting cell proliferation during development in *Drosophila* (Treisman et al., 1995) or *Arabidopsis* (Springer et al., 1995). Phylogenetic comparison of eukaryotic Mcms sequences shows that the six classes of Mcm proteins are approximately equally related (Figure 1.3). The tendency of Mcm proteins to interact with each other has made it possible in some cases to co-purify and characterize proteins in the family (Kubota et al., 1997; Kimura et al., 1996). Higher eukaryotic Mcms are also present in archaea. Although archaea are prokaryotes, their replication machinery and the proteins participating in the initiation of DNA replication are more similar to those found in eukarya than bacteria (Grabowski and Kelman, 2003).

1.4.1 Archael Mcm

In all eukaryotes six Mcm proteins are found but this is not true for archaea. All archaeal species sequenced to date contain at least one homologue of the eukaryotic Mcm2–7 family (reviewed in Kelman and Kelman, 2003; Enemark and Joshua-Tor, 2008) and evidence is accumulating that Mcm is also the archaeal replicative helicase. Most of the archaea have usually one Mcm homologue but the *Methanocaldococcus jannaschii* and *Methanococcus maripaludis* S2 genomes contains four Mcms (Bult et al., 1996; Hendrickson et al., 2004), while *Methanothermobacter thermautotrophicus* has just a single Mcm that forms a double hexamer

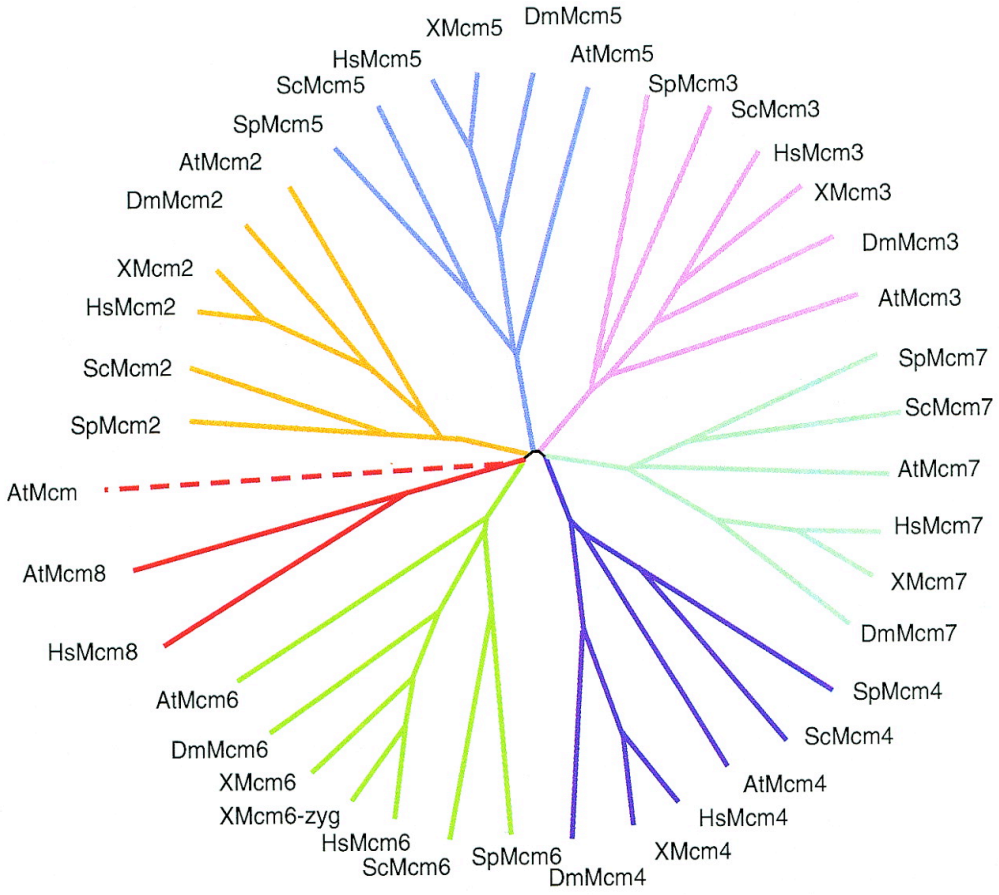


Figure 1.3 Phylogenetic tree of eukaryotic Mcms assembled using ClustalX: This figure was generated using PHYLODENDRON and taken from the Forsburg, 2004. Colours correspond to the seven Mcm subfamilies. *Schizosaccharomyces pombe* Mcm: SpMcm; *Saccharomyces cerevisiae* Mcm: ScMcm; Homo sapiens (Human) Mcm: HsMcm; *Xenopus laevis* Mcm: Xmcm; *Arabidopsis thaliana* Mcm: AtMcm; *Drosophila melanogaster* Mcm: DmMcm.

complex and has an ATP-independent DNA-binding activity which can distinguish between single- and double- stranded DNA and has 3'-5' helicase activity with the requirement of ATP hydrolysis (Chong et al., 2000). The first structural information on the archaeal Mcm came from electron microscope (EM) reconstructions studies of the Mcm protein from the archaeon *Methanothermobacter thermautotrophicus* (Chong et al., 2000; Yu et al., 2002). Furthermore it was suggested that the helicase formed heptameric rings and helical filaments (Yu et al., 2002; Chen et al., 2005) compared to the initial study where it was shown to form a double hexamer structure in solution (Chong et al., 2000) and supported by electron microscopy (Gomez-Llorente et al., 2005) and structural studies (Fletcher et al., 2003). Subsequent EM studies with the same Mcm protein revealed structural polymorphism. It was found that in addition to helical filaments and heptameric rings the protein also formed double heptamers, hexamers and double hexamers, octamers and open rings (Pape et al., 2003; Gomez-Llorente et al., 2005; Costa et al., 2006 a,b; Jenkinson et al., 2009). The single Mcm from *M. thermautotrophicus* was the first archaeal Mcm that was shown to have DNA binding, ATPase and 3'-5' helicase activities (Kelman et al., 1999; Chong et al., 2000; Shechter et al., 2000). Helicase activities were then demonstrated for Mcms from *S. solfataricus* (Carpentieri et al., 2002), *Archaeoglobus fulgidus* (Grainge et al., 2003) and *Aeropyrum pemix* (Atanassova and Grainge, 2008). A near-full-length *Sulfolobus solfataricus* Mcm hexamer has generated based on the 6-fold symmetry of the N-terminal *Methanothermobacter thermautotrophicus* (*Mth*) hexamer (Brewster et al., 2008).

1.4.2 Eukaryotic Mcms

All eukaryotic Mcm sequences obtained thus far appear to be homologous to one or other of the six *S. cerevisiae* Mcm proteins, suggesting that there were six distinct Mcm genes in a primordial eukaryote. Recently, two more additional distantly related proteins Mcm8 and Mcm9 (Maiorano et al., 2005; Maiorano et al., 2006) have been identified, however, these are not present in all eukaryotes (Gozuacik et al 2003, Lutzmann et al., 2005).

Mcm9 only exists in vertebrates (Maiorano et al., 2006). Mcm8 functions in the elongation step of DNA replication as a helicase and stimulates the processivity of DNA polymerases at replication sites. Mcm8 does not associate with the soluble Mcm2-7 complex and binds chromatin upon initiation of DNA synthesis (Maiorano et al., 2005). It has been demonstrated that Mcm9 binds to chromatin in an ORC-dependent manner and is required for the recruitment of the Mcm2-7 helicase onto chromatin (Lutzmann and Mechali, 2008). There has been no direct interaction detected between these newly discovered Mcms and the Mcm2-7 helicase family (Maiorano et al., 2005; Lutzmann and Mechali, 2008).

Mcm1 and Mcm10 do not belong to this family but are conserved in higher eukaryotes. Mcm1 is a transcription factor that may play a role in DNA synthesis while Mcm10 appears to be directly required to initiate DNA synthesis (Maiorano et al., 2006).

1.5 Biochemical properties of the Mcm proteins

The Mcms belong to a distinct subgroup of the large AAA+ ATPase family of proteins, which has many cellular functions (Davey et al., 2002). AAA+ ATPases generally form large ATP-dependent complexes. Mcms can be divided into three different domains (Figure. 1.4) which are highly conserved; an N-terminal domain which contains a Zn-finger motif, a central AAA+ (ATPases Associated with various cellular Activities) domain containing Walker A and Walker B motifs and a C-terminal domain containing a winged helix-turn-helix motif (Jenkinson and Chong, 2006).

1.5.1 The AAA+ domain

The central AAA+ domain can be divided into an α/β subdomain (the functional core of the enzyme) and an α subdomain mediating inter-subunit contacts. The α/β subdomain contains the Walker A, Walker B and arginine finger signature sequence motifs typical of AAA+ ATPases (Koonin, 1993).

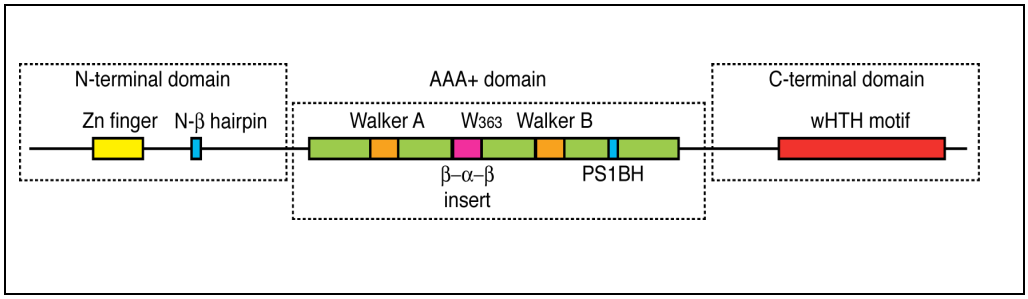


Figure 1.4 A schematic representation of the domain structure of Mcms: The Mcm cartoon is showing N-terminal domain containing zinc-finger and β-hairpin, central AAA+ ATPase domain containing Walker A and B motif, β-α-β insert, PS1BH (pre-sensor 1 beta-hairpin), and a C-terminal domain with wHTH (winged-helix-turn-helix) (Jenkinson and Chong, 2006).

Crystal structure of nearly full-length Mcm from *Sulfolobus solfactaricus* has been solved (Breswster et al., 2008). A systematic classification of AAA+ proteins has been carried out and identified a superclade which includes the Mcm proteins and is characterised by an insertion in the α/β subdomain occurring before the `sensor 1' motif and forming a β-hairpin (Iyer et al*.,* 2004). This insertion has been called the pre-sensor 1 β-hairpin (PS1BH) and defines the superclade. Point mutation analysis of the *Sulfolobus solfataricus* Mcm (*Sso*Mcm) orthologue demonstrated that this insertion has a role in DNA binding and helicase activity (McGeoch et al., 2005). Within the PS1BH superclade, a number of AAA+ proteins contain an additional insertion in the α/β subdomain, disrupting the continuity of helix 2 (Iyer et al*.,* 2004). In Mcm proteins, the h2i (helix-2 insert) has been proposed to act as a ploughshare separating the two strands during the helicase reaction (Chong 2005; Jenkinson and Chong, 2006).

1.5.2 N-terminal domain and the zinc finger

The N-terminal domain of the archeael Mcm has been shown to be important in binding to single stranded DNA, protein multimerisation and the regulation of helicase activity (Jenkinson and Chong, 2006; Kasiviswanathan et al., 2004). The conserved zinc finger motif found in the N-terminal domain of the Mcm has shown to be essential for helicase activity (Poplawski et al., 2001). The N-terminal domain has also been shown to be the site of interaction with the origin binding protein, cdc6, which modulates the activity of Mcm (Kasiviswanathan et al., 2005; Haugland et al., 2006). Mutation in the conserved loop of the N-terminal domain has no effect on DNA binding but caused decreases in ATPase and helicase activities (Sakakibara et al., 2008).

The crystal structures of the poorly conserved N-terminal portion of *Methanothermobacter thermautotrophicus* Mcm (N-*Mth*Mcm) and *Sulfolobus solfataricus* Mcm (N-*Sso*Mcm) reveal that this region influencesn the formation of dodecamers (Figure 1.5 A) and hexamers (Fletcher et al., 2003; Liu et al., 2008). The monomeric fold and the assembled hexamer

structures of the N-*Sso*Mcm and N-*Mth*Mcm are highly conserved (Fletcher et al., 2003; Liu et al., 2008). A β-hairpin structure present in the N domain of the both Mcm proteins protrudes into the central hexameric channel to form the narrowest point within the channel, possibly for interacting with DNA at a certain stage of Mcm function (Fletcher et al., 2003; McGeoch et al., 2005).

1.5.3 C-terminal domain

The last 100 amino acids of the *Mth*Mcm protein include a predicted helix–turn–helix motif (HTH), a fold typical of DNA interacting domains (Aravind and Koonin, 1999). The HTH motif appears to be present in other archaeal Mcm sequences but not in eukaryotic proteins. A deletion mutant of the *Mth*Mcm protein lacking the C-terminal domain (Δ597) showed increased dsDNA-stimulated ATP hydrolysis and increased the affinity of the mutant complex for ssDNA and dsDNA but lower helicase activity, observed by steady state fluorescence (Jenkinson and Chong, 2006). FRET (fluorescence resonance energy transfer) experiments on the *Sso*Mcm also suggested a DNA-induced conformational change involving the C-terminal domain (McGeoch et al., 2005). These observations indicate a possible role for the C-terminal domain in coupling the ATP hydrolysis and DNA spooling motion necessary for the helicase activity.

1.5.4 Hairpins

There are three beta (β) hairpin structures located in the C terminal half of Mcm (Figure 1.5 B). The presensor 1 beta hairpin (PS1BH), which is predicted to protrude into the central channel of the complex at the intersection with the side channel (Brewster et al., 2008) has been shown to be essential for helicase activity (McGeoch et al., 2005). An insert in helix 2 that forms another β-hairpin structure between the Walker A and B motifs is unique to the Mcms and was predicted to play a role in DNA unwinding (Iyer et al., 2004; Chong, 2005). Deletion of this β-hairpin structure uncoupled ATP hydrolysis from helicase activity; with the deletion mutant

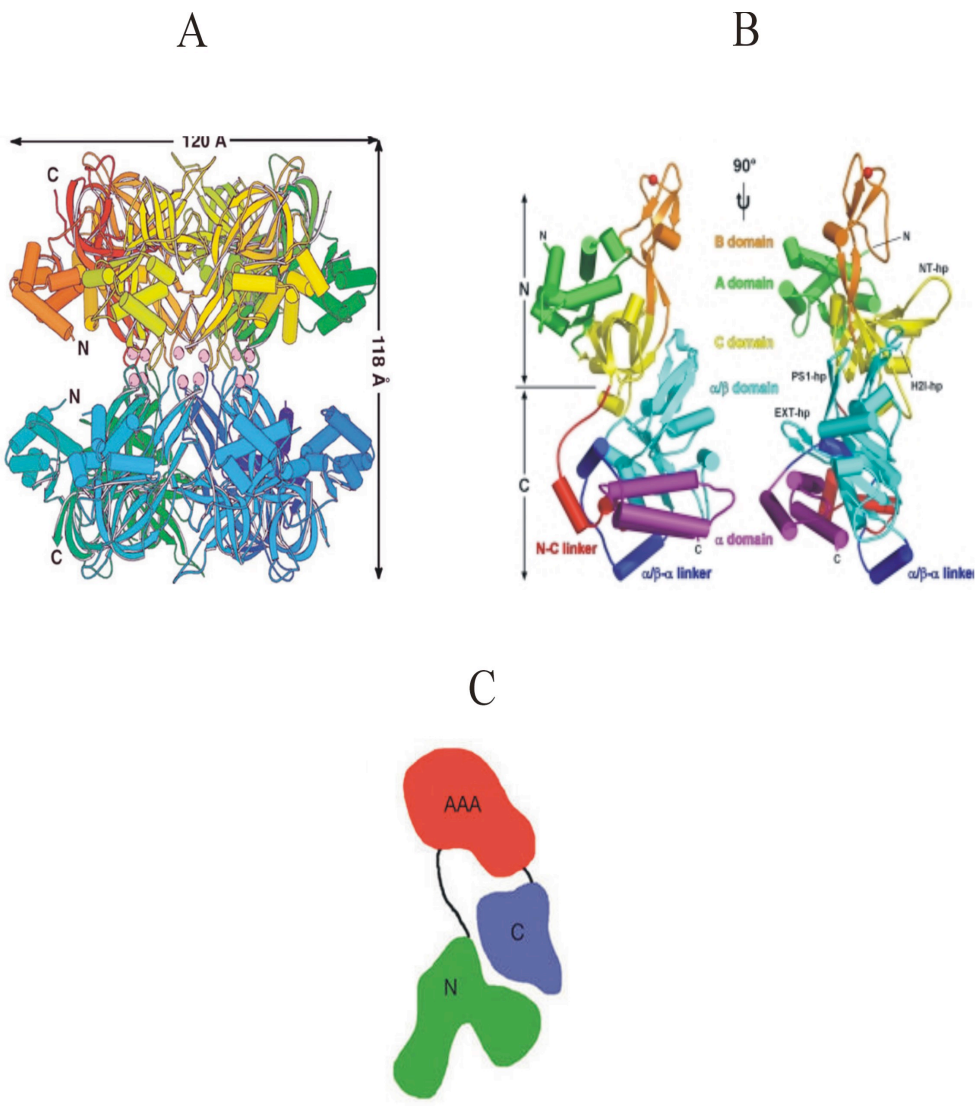


Figure 1.5 Features of the monomeric and double-hexameric Mcm structural models: (A) Dodecameric structure of the N-terminal domain of *M. thermautotrophicus* Mcm (Fletcher et al., 2003) (B) Near full length structure of a single Mcm protein from *S. solfataricus* Mcm showing hairpin structures (Brewster et al., 2008) (C) Cartoon of *M. thermautotrophicus* Mcm single subunit indicating the relative positions of N-terminal, C-terminal and AAA+ domains (Pape et al., 2003).

possessing increased ATP hydrolysis but no helicase activity (Jenkinson and Chong, 2006). The *S. solfataricus* Mcm crystal structure confirms that this β-hairpin protrudes into the central channel of hexamer, supporting the predicted mechanical role of this β-hairpin (Brewster et al., 2008). A third β-hairpin, EXT, in the C-terminal half of the protein is located close to the exit of the side channel. Mutations at the tip and the base of this β-hairpin have shown that it is required for helicase activity but not DNA binding (Moreau et al., 2007; Brewster et al., 2008). The fact that β-hairpins positioned in both the central and side channels of the complex are required for helicase activity supports the hypothesis that single stranded DNA is extruded through the side channel, as it is unwound. Three-dimentional reconstruction using electron microscopy indicates the relative positions of N-terminal, C-terminal and AAA+ domain in *M. thermautotrophicus* Mcm (Figure 1.5 C). However, despite the large amount of biochemical data generated through structural and mutational analysis, the way in which DNA is passed through the Mcm complex and exact mechanism of unwinding remain unclear.

1.6 Structure of the Mcm complex

From studies using *M. thermautotrophicus* a three dimensional reconstruction of the Mcm complex has been achieved using negatively stained particles coupled with electron microscopy. Several reports shows that the Mcm protein has a hexameric ring structure with six individual monomers arranged around a six-fold axis. The large central cavity lined by the N-terminal domains and AAA+ domains of the monomers is wide enough to accommodate dsDNA (Figure 1.6 A, B) (Sato et al., 2000; Kalpan et al., 2003; Brewster et al., 2008). A three dimensional reconstruction of *M. thermautotrophicus* Mcm shows the top and bottom halves of the molecule are asymmetrical, the top half is dome shaped whereas the bottom half has a ridged shape; the ridge corresponds to the zinc motifs on each subunit (Figure 1.6 C).

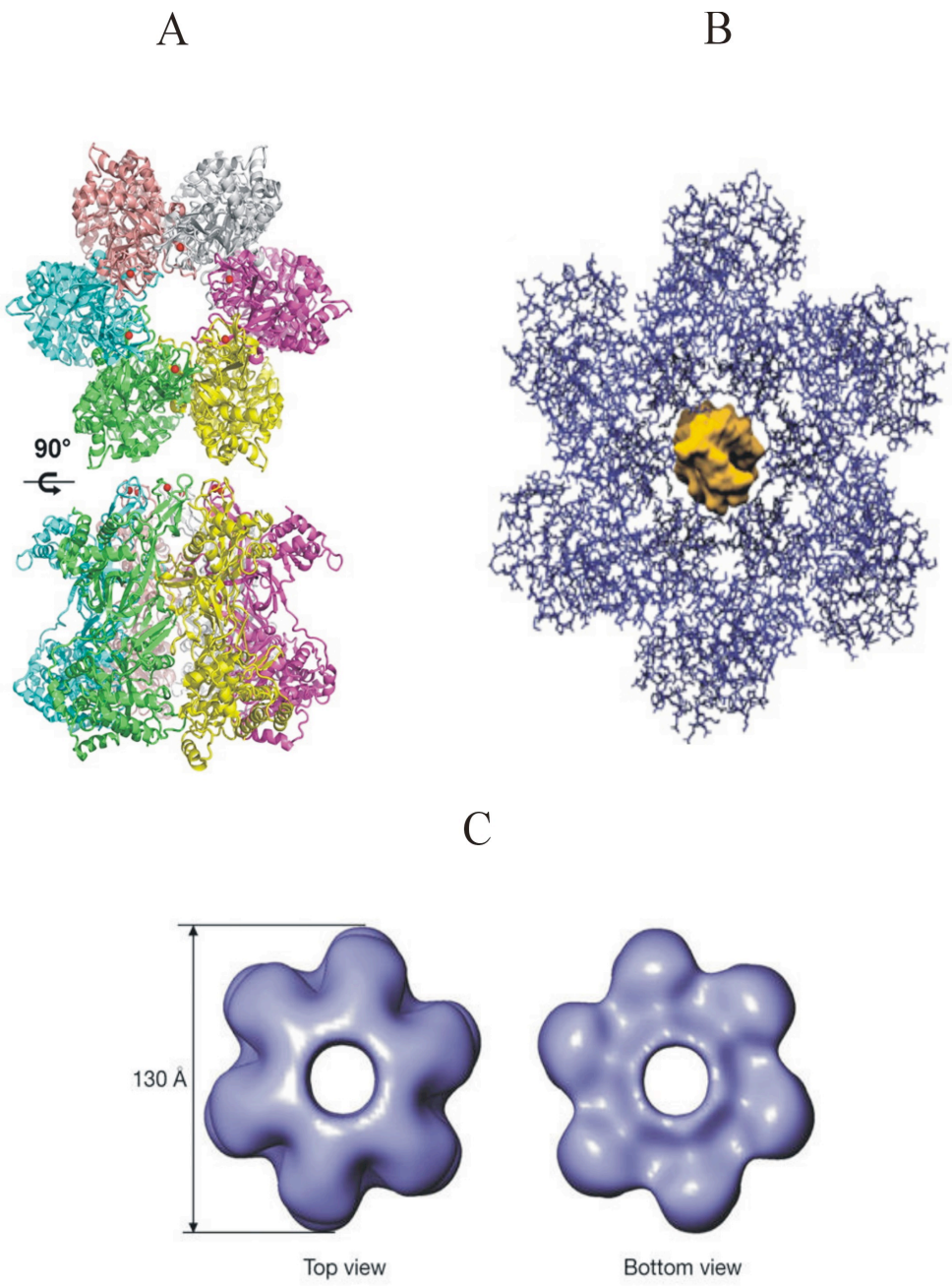


Figure 1.6 Structure of the Mcm complex: (A) Crystal structure (ribbon diagram) showing the top and side views of a hexamer model of *S. solfataricus* Mcm (Brewster et al., 2008), (B) The channel of the double hexamer with modelled dsDNA (yellow) at the narrowest point (Fletcher et al., 2003) (C) Cryo-EM structure of top and bottom half of the molecule shows asymmetry, the top half is dome shaped (Pape et al., 2003).

1.6.1 Interactions between Mcm 2-7

Purification of Mcm2-7 proteins as individual subunits of *S. cerevisiae* Mcm proteins as recombinant proteins in *E. coli* can be achieved and it was found that the Mcm heterohexamer could be reconstituted from these individual subunits (Davey et al., 2003). The Mcm subunits were studied individually and in many combinations to see if any ATPase activity was present. The study showed that individual Mcm subunits on their own could not hydrolyse ATP effectively (even though they have an ATP binding site); to achieve ATP hydrolysis two or more Mcm subunits are needed. Analysis of the Mcm2-7 active sites, both as isolated dimers and within the context of the Mcm2-7 heterohexamer indicated that they contribute unequally to ATP hydrolysis and viability (Bochman et al., 2008). ATPase activity requires residues from both subunits; a catalytic arginine on one subunit and an ATP binding site on the other subunit (Davey et al., 2003). There are fifteen possible combinations to pair up the six subunits and it was found that in the absence of DNA that only three of these combinations showed any ATPase activity. The three pairs that showed ATPase activity are; Mcm3/7 proteins (which showed the most ATPase activity), Mcm4/7 proteins (showed weaker ATPase activity) and Mcm2/6 proteins. These three pairs all showed significantly more ATPase activity than the heterohexamer of all six Mcm proteins (Davey et al., 2003). The suggestion that ATPase activity requires residues from two subunits was tested by analysing the activity of paired subunits with mutations. The arginine residues in catalytic SRF motif in the Mcm3 and Mcm7 were changed to alanine and then these mutated proteins were mixed with wild type proteins. It was shown that pairs containing mutant Mcm3 and wild type Mcm7 did not show any ATPase activity. Whereas mutated Mcm7 and wild type Mcm3 showed as much ATPase activity as the wild type Mcm3/7 pairing. This suggests that only one subunit provides the catalytic arginine (in this case it is Mcm3) whilst the other provides the ATP binding site (here it is Mcm7).

Stable interactions between Mcm proteins were tested by mixing pairs of Mcm subunits and analysing their gel elution volume (the oligomeric state of

each Mcm protein can be determined by elution volume by comparing elution volumes to already known elution volumes of each Mcm subunit). Interactions have been shown between Mcm3/5 and Mcm4/6 and ambiguous results were found for Mcm pairs; Mcm2/3, Mcm2/4, Mcm3/4, Mcm3/6, Mcm6/7 so it was concluded that in the absence of any convincing data that these pairs of proteins do not interact with each other. From these observed interactions along with ATPase activities a proposed structure of a Mcm2-7 heterohexamer is the ring shown in Figure 1.7 (Davey et al., 2003). The same organisation of subunits within the Mcm2-7 complex has been predicted in several independent studies (Crevel et al., 2001; Davey et al., 2003; Yu et al., 2004; Bochman et al., 2008;). It has been suggested that a discontinuity between the Mcm2 and Mcm5 subunits forms a "gate" that allows loading of the ring-shaped complex onto DNA (Bochman et al., 2008; Bochman and Schwacha, 2008). Although there are only three Mcm proteins that show ATPase activity (Mcm3, 4 and 7) this does not mean that the other subunits are not catalytic, they may need extra factors to activate them *in vivo* that were not available in tests done *in vitro.* This may be a mechanism to make sure that the ATPase activity of the Mcm proteins is not on unless the Mcm hexamer is under the control of some other factors.

1.6.2 Interaction of Mcm2-7 with other proteins

Interactions have been described between Mcms and components of the origin recognition complex ORC, which was identified by its ability to bind to the conserved ARS consensus sequence of replication origins in *S. cerevisiae* (Bell and Stillman, 1992). ORC binds replication origins *in vivo* (Diffley and Cocker, 1992) and six components of the complex encoded by the ORC1 – 6 genes have been identified (Bell et al., 1993; Foss et al., 1993; Li and Herskowitz, 1993; Micklem et al., 1993; Bell et al., 1995; Loo et al., 1995). Mutations in ORC genes cause defects in DNA replication, and reduce the efficiency of initiation at origins (Bell et al., 1993; Foss et al., 1993; Micklem et al., 1993; Loo et al., 1995). The *S. pombe* homologue of Orc1 has been shown to interact with Mcm4 both genetically and biochemically (Grallert and Nurse, 1996) and genetic interactions have also

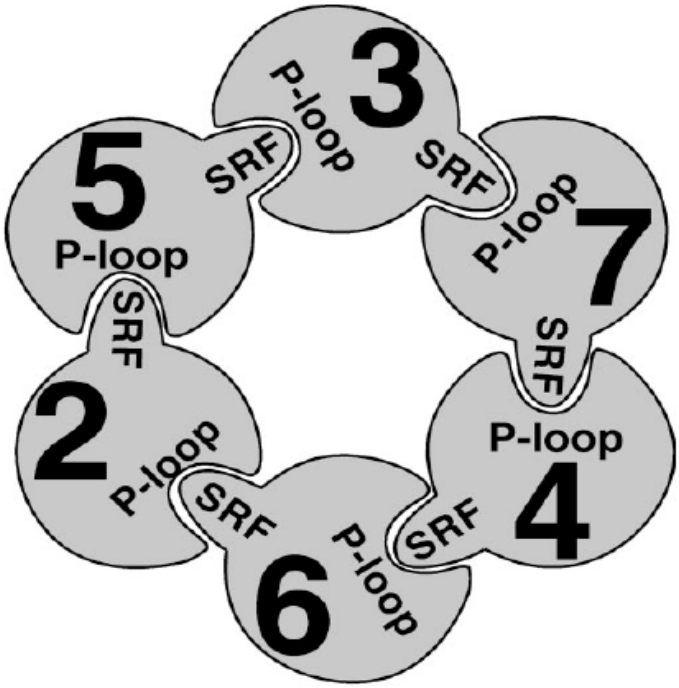


Figure 1.7 Proposed structure of the Mcm2-7 heterohexamer: This structure is based on the interactions observed in the gel filtration analysis and ATPase activity. This ring is viewed from the C-terminal face. SRF represents the catalytic SRF motif (arginine finger) and P-loop (Walker A motif) represents the ATP binding site (Image and legend, Davey et al., 2003).

been shown between the Mcm5 and ORC6 genes (Li and Herskowitz, 1993).

In addition to this ORC interaction, both Mcm5 and Mcm7 were originally identified as extragenic suppressors of a mutation in the CDC45 gene (Moir et al., 1982), which encodes another essential replication factor (Hennessy et al., 1991) and this interaction was both dominant and allele specific. This interaction is most easily explained if Cdc45 physically associates with Mcm proteins and indeed the proteins have subsequently been shown to co-immunoprecipitate (Hopwood and Dalton, 1996). The role of Cdc45 in DNA replication is not known, but it also appears to interact at least genetically with ORC (Hardy, 1997; Zou et al., 1997). CDC45 is the eukaryotic orthologue of the bacterial and archaeal RecJ family nucleases (Makarova et al., 2012). *In vitro* and *in vivo* experiments have demonstrated that the Mcm2-7 helicase complex requires the association with GINS and the CDC45 protein and thought to be the active replicative helicase unit *in vivo* (Moyer et al., 2006; Pacek et al., 2006; Labib and Gambus, 2007).

In addition to interacting with ORC and other initiation factors, which would potentially allow origin-specific association with chromatin, Mcms may also associate less specifically with nucleosomes via an interaction with histones. The tetrameric Mcm complex Mcm2, Mcm4, Mcm6, Mcm7 interacts with histone H3, which could be important for stabilizing the interaction between Mcm proteins and chromatin, or changing the stability of nucleosomes in the vicinity of replication origins (Ishimi et al., 1996). The Mcm2, Mcm4, Mcm6, Mcm7 tetramer does not interact with histone H2A or H2B and the dimeric Mcm subcomplex Mcm3, 5 appears to show no interaction with histones. The stability of the interaction between Mcm3 and Mcm7 and chromatin in extracted cells has been shown to be increased in the presence of ATP or non-hydrolysable analogues of ATP (Fujita et al., 1997) suggesting that ATP binding by Mcms or other chromatin proteins e.g., components of ORC, may in some way enhance the interaction with histones. Human Mcm3 was first described as a protein associated with DNA polymerase α primase (Thommes et al., 1992), although subsequent analysis failed to demonstrate a direct interaction between the two proteins

(Burkhart et al., 1995). However, genetic interactions have been found between subunits of DNA polymerase δ and Mcm2 in fission yeast (Forsburg et al., 1997). Archaeal GINS complex was identified and showed that it directly physically interacts with the archaeal Mcm (Marinsek et al., 2006). Very recently interaction of McmA from *M. maripaludis* has been found with other hypothetical proteins namely MMP1213, MMP0004 and MMP0457 (Chong lab unpublished data).

1.7 Regulation of Mcm helicase activity

Two key classes of regulatory molecules, cyclins and cyclin dependent kinases (CDKs), determine a cell's progress through the cell cycle (Nigg et al., 1995). Cyclins form the regulatory subunits and CDKs the catalytic subunits of an activated heterodimer cyclins have no catalytic activity and CDKs are inactive in the absence of a partner cyclin. When activated by a bound cyclin, CDKs perform a common biochemical reaction called phosphorylation that activates or inactivates target proteins to orchestrate coordinated entry into the next phase of the cell cycle. Different cyclin-CDK combinations determine the downstream proteins targeted. CDKs are constitutively expressed in cells whereas cyclins are synthesised at specific stages of the cell cycle, in response to various molecular signals.

CDK/cyclin (CDK) and Cdc7/Dbf4 (DDK) activity are required for the initiation of DNA replication in all eukaryotic cells (Stillman, 1996; Dutta and Bell, 1997) Phosphorylation plays a key role in regulating Mcm activity (Coue et al., 1996).

In *Xenopus*, levels of Cdk2-cyclin E appear to be important for regulating Mcm chromatin binding. Although overall levels of Cdk2-cyclin E levels are similar throughout the embryonic cell cycle, compartmentalization of the kinase in newly formed nuclei gives rise to local concentrations around chromatin that are much higher than those in the cytosol (Hua et al., 1997).

Low levels of Cdk2-cyclin E or cyclin A are essential to allow formation of replication-competent chromatin, apparently because Mcm3 can only associate with chromatin under these conditions (Li et al., 2011). *S. cerevisiae* Cdks also seem to have a role in controlling the association of Mcms with chromatin. The ability of Mcm7 to associate with origins (Tanaka et al., 1997), or the ability of origins to form pre-RCs as defined by genomic foot-printing requires expression of Cdc6 during a period of low Cdk Cdc28-Clb activity. Cdk activity could in principle cause this effect by destabilizing Cdc6 and thus preventing both its association with chromatin and the subsequent binding of Mcms in *S. cerevisiae* (Piatti et al., 1997).

In a related study in *Xenopus* egg extracts, Cdc2-cyclin B has also been shown to be an inhibitor of a component of replication licensing (Mahbubani et al., 1997), and in *Drosophila*, mitotic degradation of cyclin A but not cyclin B is required for rebinding of Mcms to chromatin (Su and O'Farrell, 1997). Mcm4 may be phosphorylated by Cdk2/cylin A, since it contains consensus phosphorylation sites in the N-terminal region of the protein that are conserved amongst eukaryotes, and there is direct experimental evidence that it is a substrate *in vitro* (Kudoh et al., 2006). Examination of Mcm4 during the *Xenopus* cell cycle provides evidence for Cdk phosphorylation. Mcm4 is hypo- or de-phosphorylated in G1 phase but chromatin-bound Mcm4 is phosphorylated in S phase before it is displaced by a replication fork (Coue et al., 1996). Recently it has been proved that the Dbf4–Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4 in *S. cerevisiae* (Sheu and Stillman, 2010).

At least some Mcm proteins are phosphorylated in a cell cycle specific manner. Mcm2 from *S. cerevisiae* (Todorov et al., 1995; Young and Tye, 1997) Mcm3 from *S. cerevisiae* (Young and Tye, 1997) Mcm4 from *Xenopus* (Hendrickson et al., 1996) and Mcm7 from human sources (Tsuruga et al., 1997) have been shown to be phosphoproteins. Mcm2, Mcm3 and Mcm4 proteins show a similar cell cycle pattern of phosphorylation, becoming dephosphorylated or hypophosphorylated on exit from mitosis, and being phosphorylated as cells enter S phase (Hendrickson

et al., 1994; Todorov et al., 1996; Young and Tye, 1997). Mcm proteins bound to chromatin are hypophosphorylated compared to the displaced proteins suggesting that phosphorylation triggers or shortly precedes the displacement of Mcms, or alternatively that phosphorylation can only occur on displaced Mcm proteins (Kimura et al., 1994; Musahl et al., 1995; Todorov et al., 1995; Hendrickson et al., 1996). For at least two Mcms, human Mcm2 and *Xenopus* Mcm4, the proteins are phosphorylated first in S phase, but become hyperphosphorylated during mitosis (Todorov et al., 1995; Coue et al., 1996).

Not much is known about the regulation of the archaeal Mcm helicase *in vivo*. *In vitro* studies on the regulation of the helicase yield conflicting results. To date, two proteins affect helicase activity *in vitro*. The *M. thermautotrophicus* and *S. solfataricus* Cdc6 protein (a homologue of the eukaryotic initiation protein Cdc6; also referred to as Orc or Orc/Cdc6) inhibits helicase activity *in vitro* (De Felice et al*.,* 2003; Shin et al*.,* 2003b; Kasiviswanathan et al*.,* 2005). Cdc6 protein dissociates the hexameric helicase (Shin et al*.,* 2008), which could be the mechanism by which helicase activity is eliminated. This dissociation of Mcm may also suggest a mechanism for helicase loading at the origin. The Mcm protein forms stable hexameric rings and these rings have to open at one interface, or to dissociate into monomers, prior to assembly at the replication bubble. This function may be mediated by Cdc6-mediated Mcm dissociation (Shin et al*.,* 2008).

However, not all archaeal Cdc6 proteins inhibit their respective Mcm helicases. The Mcm protein from *T. acidophilum* exhibits very poor helicase activity on its own (Haugland et al., 2006; Haugland et al., 2009) but its activity is enhanced by the *T. acidophilum* Cdc6 protein (Haugland et al*.,* 2006). Binding of Cdc6 to Mcm protein stimulates the helicase ATPase activity and enhances helicase activity (Haugland et al., 2006; Haugland et al., 2008). Binding of Cdc6 to Mcm protein might induce conformational changes in the helicase that stimulate ATP hydrolysis.

Another archaeal protein that regulates helicase activity is the GINS complex from *Pyrococcus furiosus* (Yoshimochi et al., 2008). Like the Mcm helicase from *T. acidophilum* (Haugland et al*.,* 2006), the *in vitro* helicase activity of the *P. furiosus* Mcm is very weak. However, it is stimulated upon interaction with the GINS complex (Yoshimochi et al*.,* 2008*)* and, like the *T. acidophilum* Cdc6 protein (Haugland et al., 2008), GINS stimulates *P. furiosus* Mcm helicase activity by stimulating its ATPase activity (Yoshimochi et al*.,* 2008). However, the effect of GINS on Mcm activity might not be a generalized phenomenon, because although Mcm and GINS proteins from *S. solfataricus* interact, no effect on Mcm helicase activity *in vitro* was observed (Marinsek et al*.,* 2006).

To date, in archaea, only the *A. pernix* Mcm protein was shown to be phosphorylated by a Cdc6 homologue (Atanassova and Grainge, 2008). Thus, phosphorylation or other modification might be involved in the regulation of the archaeal Mcm protein *in vivo*.

1.8 Cancer and Mcm

Association of the Mcm2-7 complex to ORC is a crucial moment in initiating the replication fork. Mcm proteins play a role in maintaining genome integrity and prevent re-replication once per cell cycle. Proliferating cells have high levels of Mcm, whereas they are not detected in quiescent, differentiated or senescent cells (Hiraiwa et al., 1997; Freeman et al., 1999; Meng et al., 2001; Ramnath et al., 2001; Going et al., 2002; Rodins et al., 2002; Ishimi et al., 2003). They are also potentially useful markers of cell proliferation. Recent studies suggested that Mcms are good markers of proliferation activity, because they are highly expressed in a variety of tumours. Since molecular studies showed that increased levels of Mcms mark proliferative malignant cells (Ishimi et al., 2003) and may prove to be effective diagnostic markers for tumourigenesis (Todorov et al., 1998; Freeman et al., 1999; Stoeber et al., 1999; Lei, 2005). Studies on human Mcms are therefore potentially useful in providing targets for medicine or therapies.

Recently, several groups have reported that Mcm proteins are more frequently detected in cells from malignant tissues than those from normal tissues (Hiraiwa et al., 1997; Hiraiwa et al., 1998; Todorov et al., 1998; Williams et al., 1998; Freeman et al., 1999; Stoeber et al., 1999; Coleman et al., 2001; Endl et al., 2001; Stoeber et al., 2001; Tan et al., 2001; Wharton et al., 2001). This phenomenon was also observed in dysplastic cells (Freeman et al., 1999). Thus the presence of Mcms in precancerous cells and the potential reoccurrences (Alison et al., 2001; Hunt et al., 2002) make them effective markers.

The fact that Mcms are over expressed in a cell cycle specific manner in a wide range of cancers means that they are not only potentially useful markers for cancers but that they also have the potential to be anti-cancer drug targets. There are several reports showing that Mcm helicase activity can be inhibited *in vivo* due to modifying Mcms (Ishimi et al., 2000; Takei et al, 2002). Cyclin A/Cdk2 phosphorylates human Mcm4 and inactivates the human Mcm4/6/7 complex helicase activity (Ishimi et al., 2000). Human Mcm2, 3 and 7 are phosphorylated by the ATM/ATR checkpoint kinases in response to DNA damage (Cortez et al., 2004; Shi et al., 2007). Over expression of human Mcm3 acetylating protein (Mcm3AP) inhibits the initiation of DNA replication (Takei et al, 2002).

In vitro studies showed that Mcm2 could be efficiently silenced in human cells by RNAi or antisense oligodeoxynucleotides (ODNs) and this led to the inhibition of DNA replication. This study analyzes the transforming activity and signalling of Mcm7, the oncogenic function of the miRNA cluster that is embedded in the Mcm7 genome, and the potential of gene therapy that targets Mcm7 (Luo, 2011). Furthermore RNAi mediated knockdown of Mcm2-7 in human cells also increases in sensitivity to DNA damage (Ibarra et al., 2008). However, the tests have so far only been conducted *in vitro* and although several RNAi based therapies targeting other proteins are currently in clinical trials, delivering the antisense ODN/RNAi molecules to cancerous cells *in vivo* remains a significant problem (deFougerolles, 2008).

1.9 Model organisms to study Mcm complex

Eukaryotes need cell cycle regulation control by DDK and CDK activity, the homologues of which are not present in archaea. It is thought that fewer proteins are needed in archaeal replication (Edgell and Doolittle, 1997). Methanogens were the first organisms to be identified as archaea and classified as a separate domain (Woese and Fox, 1977). Methanogens are divided into five orders: Methanobacteriales, Methanococcales, Methanomicrobiales, Methanosarcinales and Methanopyrales (Thauer, 1998). Complete genome sequences for representatives of all of these orders are available. The observation that archaeal information processing systems are similar to but simpler than those in eukaryotes, along with the thermostable nature of proteins from many of these organisms, has led to a number of archaeal species being used as models for a variety of eukaryotic processes (Tye, 2000). DNA replication is a good example of how a eukaryotic process is simplified in archaea. In archaea that have been characterized to date, a single Mcm protein forms a homohexameric complex that possesses ATP-dependent DNA helicase activity (Chong et al., 2000). Mcm complexes from a number of species have been characterized and have provided insight into the mechanisms that govern helicase activity in these proteins. Motifs essential for Mcm function have been identified using ATP hydrolysis, DNA binding and DNA helicase activity assays (Kelman et al., 1999; Chong et al., 2000; Shechter et al., 2000; Poplawski et al., 2001; Carpentieri et al., 2002; Kasiviswanathan et al., 2004; Jenkinson and Chong, 2006; Barry et al., 2007; Bochman and Schawacha, 2008). The genome sequence of *Methanocaldococcus jannaschii* (Mj) has revealed that in contrast to the single Mcm found in other archaea, this organism has multiple, apparently functional Mcm homologues (Bult et al., 1996).

Methanococcus maripaludis is a mesophilic methanogen (Bult et al., 1996) that has one of the best-developed genetic systems of any organism in the archaea domain (Hendrickson et al., 2004). Because the genome of *M. maripaludis* S2 has been completely sequenced, this strain is the strain of choice for genetic manipulation (Jones, 1983; Hendrickson et al., 2004).

Analysis of the genome sequence of *M. maripaludis* S2 has revealed that it also contains multiple Mcms (Hendrickson et al., 2004; Walters and Chong, 2009). A well-defined system for genetic manipulution of *M. maripaludis*, greatly increases its usefulness as a model system (Tumbula and Whitman, 1999; Allers and Mevarech, 2005). The four *M. maripaludis* S2 Mcm homologues are distributed across the genome (Hendrickson et al, 2004) and are assigned ORF numbers that have further supplemented with gene names (indicated in parentheses) as MMP0030 (McmA), MMP0470 (McmB), MMP0748 (McmC) and MMP1024 (McmD) (Walters and Chong, 2010).

It has been shown that Methanococcales have multiple functional Mcms (Walters and Chong, 2010) and may therefore be a good model for understanding eukaryotic Mcm interactions.

1.10 Issues with eukaryotic Mcm studies

In eukaryotes, *in vivo* observations implicate the Mcm2-7 complex as the replicative helicase (Bell and Dutta review, 2002). Despite this *in vivo* observation, Mcm2-7 has been reported to lack *in vitro* helicase activity (Schwacha and Bell, 2001; Davey et al., 2003; Bochman and Schwacha, 2007). Interestingly, both an archaeal Mcm complex (Kelman and White, 2005) and an alternative hexameric Mcm complex containing only three of the six eukaryotic Mcm subunits (the Mcm4/6/7 complex) have DNA unwinding activity (Ishimi, 1997; Lee and Hurwitz, 2001; Kaplan et al., 2003). The helicase activity was only recently identified in *S. cerevisiae* Mcm2-7 complex expressed in baculovirus infected insect cells (Bochman and Schwacha, 2008).

Although a lot of progress has been made in understanding eukaryotic Mcm there are still gaps that needed to be filled like how Mcms are loaded on to DNA? How and where exactly are Mcms modified to be regulated? Compared to archaeal systems, lots of replication origins and a requirement for other additional proteins partners makes it very difficult to dissect the exact process in eukaryotes. Also the post translation modifications that can

inhibit hexameric Mcm2-7 complex activity in eukaryotes make eukaryotic Mcm a difficult choice to work with. For this reason most of the research concerning Mcms are either on archaeal system or on simple eukaryotic system of yeast, *Xenopus* or *Drosophila.*

1.11 Project Aims and Objectives

An increased understanding of molecular Mcm function may aid in filling the gaps in DNA replication and regulation of DNA replication mechanisms and in identifying possible new targets for inhibition of cancer. Further understanding may also help in identifying new proteins involved in the replication process.

Although *S. cerevisiae* Mcm2 - 7 have already been shown to possess helicase activity (Bochman and Schwacha, 2008), to date no one has performed a study on human Mcm and *M. maripaludis* for helicase activity.

1.11.1 Aims

The aim of this study is to investigate the human and *M. maripaludis* Mcm helicase activity by expressing the recombinant Human and *M. maripaludis* Mcm proteins in *E.coli* so that the inhibitory post translational modifications can be omitted. A further aim is to test the purified complexes for helicase activity and effect of phosphorylation on active Mcm2-7 helicase so that some insight into the regulation by inhibiting helicase activity by some kinases can be obtained.

1.11.2 Objectives

The objectives of this project were:

- To express and purify recombinant individual human Mcms and co-express and co-purify human Mcms complex.

- To test human Mcm2-7 complex for helicase activity.

- To test human Mcm2-7 complex as a target for kinase activity and whether phosphorylation modulates helicase activity?

- To express and purify recombinant *M. maripaludis* Mcms

- To test the *M. maripaludis* Mcm individual and in complex for helicase activity.

- To determine the optimum conditions for *M. maripaludis* Mcms helicase activity.

- To find out the role of hypothetical proteins as a possible kinase that targets for *M. maripaludis* Mcm activity.

Chapter 2: Materials and Methods

2.1 Chemicals, reagents and other materials

All chemicals and reagents were supplied by Sigma-Aldrich (UK) unless stated otherwise.

2.1.1 Primer design

Primers used during PCR for amplification of DNA fragments were designed using Seqbuilder (Lasergene) software programme and ordered in lyophilized form from MWG Biotech. Primers were diluted to a final concentration of 100 μM in sterile dH_2O (See Appendix 2 for detail).

2.1.2 Growth media

E.coli cells were cultured in Luria-Bertani or lysogeny broth (LB) liquid media containing 10 g tryptone, 10 g NaCl and 5 g yeast extract per litre. LB solid medium contained the same components as liquid LB, with 15 g/litre of agar added. LB media was sent to autoclave at 121 °C.

2.1.3 Antibiotics

All antibiotic stock solutions were filter sterilized before being added to growth media. Antibiotics were added to liquid and solid growth media to give the following final concentrations; 100 μg/ml ampicillin, 30 μg/ml kanamycin and 20 μg/ml chloramphenicol and 25 μg/ml spectinomycin unless stated otherwise, were used in this study.

2.2 Sub-cloning and expression of hMcm2-7

2.2.1 Plasmid construction

Human Mcm2-7 cDNAs cloned into pBluescript SK^+ vector were provided by the Mendez lab (CNIO Spain).

2.2.2 Transformation

pBluescript SK^+ vectors containing hMcm genes were transformed into the appropriate cells (*E.coli* Novablue). 1 µl of DNA (pBluescript SK^+ vector) was added to 50 µl of cells on ice for 20 minutes and then heat shocked at 42 ^oC for 45 seconds. Cells were placed back on ice for 2 minutes followed by the addition of 250 µl of SOC (0.5 g yeast extract, 2 g tryptone, 50 mg NaCl, 2.5 mM KCl, 100 ml of sterile milliQ H_2O, pH adjusted to 7.0 and solution autoclaved, then added 0.2 µm filter sterilized sterilized 10 mM $MgCl_2$, 20 mM glycerol). Transformed cells were incubated at 37 ^oC for 1 hour with moderate agitation (approximately 150 rpm). After 1 hour incubation, 100 µl of this culture was spread on to the appropriate antibiotic plate and incubated overnight at 37 ^oC. To extract DNA from colonies, over night colonies were grown in 5 ml of LB with appropriate antibiotics.

2.2.3 Restriction digestion

Reactions were carried out in a volume of 30 µl. 5 µg of extracted DNA was digested over the period of 1 hour at 37 ^oC by using 10 units of each of NEB restriction enzyme and 1x reaction buffer adjusted with water.

2.2.4 Agarose gel electrophoresis

Digested products were resolved on 1% Seaplaque agarose gel (Scientific solutions) at 100 v for 1 hour. The samples were loaded on the gel in the respective wells after mixing with 5 µl of 6x loading dye (30% glycerol, 0.1% bromophenol blue). Gels were visualized on UV transilluminator and photographed with gel photo documentation system (Fuji). Q-Step IV DNA ladder (York Biosciences) was used to estimate the size of DNA bands unless otherwise stated.

2.2.5 Purification of restriction digest

After running on gel, individual DNA bands were excised from the gel and purified using Wizard SV PCR clean-up kit (Promega) using the standard Manufacturer's protocol. DNA concentration was measured on a nanodrop ND-1000 spectrophotometer (Thermo Scientific) using A_{260}. For more concentrated DNA, ethanol precipitation was carried out.

2.2.6 Ethanol precipitation

1/10th volume of 3 M sodium acetate (pH 5.2) was added to the purified DNA followed by 3x volume of 100% chilled ethanol and incubated on ice for 30 minutes. DNA was pelleted by centrifugation at 13,000 x *g* (Eppendorf, rotor F34-6-38) for 15 minutes at room temperature. Supernatant was removed and DNA was rinsed with 70% ethanol and centrifuged again at 13,000 x *g* (Eppendorf, rotor F34-6-38) for 15 minutes at room temperature. Supernatant was removed and pellet was either dried at room temperature for 20 minutes or vacuum dried for 10 minute. Pellet was resuspended in 100 µl of TE pH 8.0. DNA concentration was measured using a on nanodrop ND-1000 spectrophotometer (Thermo Scientific).

2.2.7 DNA ligation

Total reaction volume for ligation was 10 µl that contained 50 ng of vector with either 1 or 3-fold molar excess of insert. T4 ligase and 1x reaction ligation buffer adjusted with water were added and centrifuged briefly and incubated either at room temperature (22 °C) for 1 hour or at 4 °C overnight.

2.2.8 Sequencing

After ligation the new vector containing hMcm gene was transformed into the appropriate cells (*E.coli* Novablue) as described in section 2.2.2.

Over night colony growth showed the hMcm gene had been ligated into the vector. To confirm this, extracted DNA was digested with the same restriction enzyme used previously and compared to the expected band size.

Extracted DNA was sent for sequencing to the sequencing laboratory of York University (Technology Facility) or to MWG Biotech. Sequencing results were confirmed using SeqMan Pro DNA analysis software (Lasergene) for the presence of individual hMcm genes in respective vectors. Sequencing results showed that all of the genes were ligated correctly in vector pET47b but were out of frame, so site directed mutagenesis for all Mcms was performed to get them in the correct reading frame for expression. As the pET47b was used on direction of the Mendez lab (CNIO Spain) who provided the original clones but actually it should be pET47c vector that would have given the right frame.

2.2.9 Site directed mutagenesis of hMcm2-7 genes

To correct the reading frame of the constructs described above, hMcm2-7 genes were subjected to site directed mutagenesis (Stratagene, Quickchange SDM kit) following the manufacturers protocols and using the mutagenic primers of hMcm2-7 listed in Appendix 2. Amplification conditions included 90 ^{o}C for 30 seconds followed by the 12 cycles of PCR (95 ^{o}C for 30 sec, 55 ^{o}C for 1 minute, with a final 1 minute extension at 68 ^{o}C)

PCR products were kept on ice for 1 minute and then 1 µl of DpnI added. These samples were incubated at 37 ^{o}C for 1 hour. The resultant constructs were use to transform 50 µl of DH5α cells with 1 µl of SDM-hMcm2-7. Obtained colonies were grown overnight at 37 ^{o}C in 5 ml of LB + antibiotics with 150 rpm shaking.

2.2.10 Plasmid Preparation

DNA was isolated from overnight cultures either by using Wizard Plus SV Miniprep DNA Purification Kit from Promega or the Miniprep DNA Purification System from Qiagen throughout this study. After mini preparation to recover DNA, all the genes were sequenced to ensure that they were in frame and there were no mutations introduced by PCR.

2.2.11 Protein expression

2.2.11.1 Autoinduction

Autoinduction medium was made in a volume of 1 Litre. For 1 L volume, added 1 ml of 1 M $MgSO_4$ (24.65g $MgSO_4.7H_2O$, water to make up 100 ml, autoclaved separately), 1 ml of 1 M 10000X Metals (0.1 M $FeCl_3$. $6H_2O$ (in 0.1 M HCl), 1 M $CaCl_2$, 1 M $MnCl_2$. $4H_2O$, 1 M $ZnSO_4$. $7H_2O$, 0.2 M $CoCl_2$. $6H_2O$, 0.1 M $CuCl_2$. $2H_2O$, 0.2 M $NiCl_2$. $6H_2O$, 0.1 M Na_2MoO_4. $5H_2O$, 0.1 M Na_2SeO_3. $5H_2O$, 0.1 M H_3BO_3. $4H_2O$, autoclaved), 20 ml of 50x 5052 (25 g glycerol, 73 ml water, 2.5 g glucose, 10 g α-lactose, autoclaved), 50 ml of 20x NPS (90 ml water, 6.6 g $(NH_4)_2SO_4$, 13.6 g KH_2PO_4, 14.2 g Na_2HPO_4 pH, 6.75, autoclaved), appropriate antibiotic in ZY (10 g tryptone, 5 g yeast extract, 925 g water, autoclaved separately before addition of antibiotic). Autoinduction medium was inoculated with 1 ml of a 5 ml of culture grown at 37 ^oC in LB medium for 8 hours with 180 rpm shaking. After that culture was incubated at 30 ^oC for about 18 hours with 180 rpm shaking. 1 ml of culture was pelleted and resuspended in 100 µl of water. 20 µl benzonase (Novagen), 6x loading buffer (30% 2-mercaptoethanol, 12% SDS, 10% glycerol, 440 mM Tris pH 6.8, 0.1% bromophenol blue) was added and samples were boiled for 5 minutes at 100 ^oC before loading on a 10% SDS-polyacrylamide gel.

2.2.12 Cell culture

5 ml of culture was grown overnight by the addition of a single colony to LB medium containing appropriate antibiotics. This was sub-cultured into 750 ml of LB containing appropriate antibiotics and grown to optical density OD_{595nm} 0.5-0.7. At this point 0.5 mM IPTG was added to the culture. Cultures were then incubated overnight with 180 rpm shaking at 37 ^oC.

2.2.13 Protein purification

2.2.13.1 Sample preparation

Cells were harvested from overnight culture by centrifugation at 4000 rpm (Eppendorf, model 5810 using swing out rotors A- 4-81) for 15 minutes at 4 ^{o}C. The cells (4-7 g) were resuspended in 5 ml/g of lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 5% Glycerol, 0.1 mM PMSF, 1 μM Pepstatin, 5 mM Imidazole, MilliQ water) followed by the addition of 0.75 mg/ml of lyzozyme and incubating at room temperature for 30 minutes to facilitate lysis. Samples were then sonicated for 3x 30 seconds at 100% power on ice. Insoluble material was removed by centrifugation at 30,000 x *g* (High Speed Sorvall Evolution, rotor F34) for 20 minutes at 4 ^{o}C.

2.2.13.2 Binding protein to the column

After preparing the Talon metal affinity resin bound with cobalt (Clontech) for purification by washing three times with wash buffer (50 mM Tris pH 8.0, 300 mM NaCl, 5% Glycerol, 0.1 mM PMSF, 1 μM Pepstatin, 10 mM Imidazole, MilliQ water), soluble protein was added. The beads were incubated for 30 minutes at room temperature with agitation followed by a spin at 4,600 x *g* (Eppendorf, model 5810 using swing out rotors A- 4-81) for 5 minutes. After washing with wash buffer the protein was eluted from the beads with elution buffer (50 mM Tris pH 8.0, 300 mM NaCl, 5% Glycerol, 0.1 mM PMSF, 1 μM Pepstatin, 150 mM Imidazole, MilliQ water) and samples were run on 10% SDS-PAGE gels and eluant fractions were stored at -80 ^{o}C. Purified proteins were sent for mass spectrometry to confirm protein ID using trypsin digestion in proteomics laboratory, York University (Technology Facility).

2.2.14 SDS-PAGE

Mini gel system (CBS Scientific) was used for all SDS PAGE. 8% to 15% gels were used throughout this study. For hMcm protein expression and

purification 12.5% gel was used. For resolving gel the correct percentage of Protogel pre mix (30% solution of 37.5:1 acrylamide:bis acrylamide, (National Diagnostics) was mixed with 375 mM Tris pH 8.7 and 0.1 % SDS. Mixed gently 5 µl TEMED and 16 µl 10% APS to 5 ml of gel mixture and poured between plates. After the gel polymerized (40 min), poured off the overlay and added stacking gel. 5% final concentration of acrylamide was mixed with 50 mM Tris pH 6.9 and 0.1% SDS for stacking gel. 5 µl TEMED and 16 µl of 10% APS were added to 2 ml of stacking gel mixture. After 5 minutes when stacking gel polymerized, samples were prepared by mixing 6x Laemmli buffer (30% 2-mercaptoethanol, 12% SDS, 10% glycerol, 440 mM Tris pH 6.8, 0.1% bromophenol blue) to the final concentration of 1x followed by 5 minutes boiling on hot block at 100 ^{o}C. Gels were run at 20 mA constant current per gel in 1x running buffer (25 mM Tris base, 192 mM glycine and 0.1% SDS).

Gels were stained with Coomassie blue R250 (40% MeOH, 10% Acetic acid, 0.1% coomassie blue (Fisher) and destained in 40% methanol and 10% acetic acid. Gels were then stored in gel preservative (18% ethanol, 5% glycerol). Gels were either photographed using gel documentation system (Fuji) or scanned after drying overnight at 37 ^{o}C between dH_2O wet acetate sheets.

2.2.15 Silver staining of protein gels

Gels were first fixed in 50% methanol for 2 x 15 minutes and then rinsed with dH_2O. Gels were then washed for 20 minutes in 32.5 µM DTT followed by 20 minutes stained in 0.1% silver nitrate. After washing with dH_2O, the gels were quickly rinsed with developer (3% sodium carbonate, 0.02% formaldehyde) till the bands became visible. To stop developing further, 5 ml of 2.3 M citric acid was added for every 100 ml of developer and incubated for 10 minutes at room temperature. The gels were washed in dH_2O for 30 minutes and then stored in gel preservative (18% ethanol, 5% glycerol) for photography or scanning.

2.2.16 Protein quantification

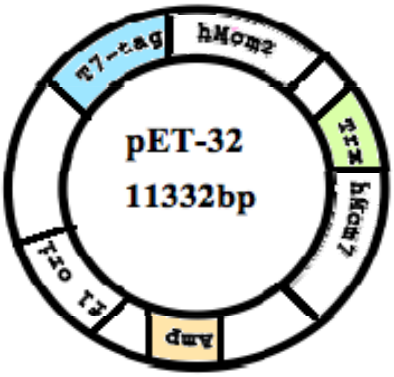
Protein concentrations were determined by the Bradford reagent (Bio-Rad) using the Bradford method (Bradford, 1976) and plotting the values against a BSA standard curve and/or on the nanodrop ND-1000 spectrophotometer at A_{260}. Concentrated proteins were flash frozen in liquid nitrogen and stored at -80 ^{o}C in small aliquots.

2.2.17 hMcm 4/6/7 complex assembly

As high imidazole concentrations were present in each of the protein fractions (hMcm4, 6 and 7) prior to assembling a complex, it was necessary to remove all salt. A HiTrap desalting column was used for this purpose. The column was connected to an AKTA system and washed with MilliQ water. The bound protein was then eluted in elution buffer (20mM Tris HCl, 1mM EDTA, 1mM EGTA, 1M KCl, 5% Glycerol, freshly added 0.1% β-mercaptoethanol and 1 μg/ml Pepstatin). Equimolar amounts of each protein (Mcm 4, 6, 7) were pooled and applied to a Q-sepharose column. The sample was applied using a 2.5 ml syringe and eluted with a linear gradient of 0-100% of KCl from 50 to 1000 mM in 0.5 ml fractions. The fractions containing protein were pooled, samples of which were run on a 10% SDS-PAGE gel.

2.3 hMcm complex co-purification by co-expression

This work has been carried out in collaboration with Dr Richard Parker, a post-doctoral researcher in the Chong group. Using pET Duet LIC vectors Mcm 4, 6 was ligated into pCDF, Mcm 3, 5 in pRSF, Mcm 2, 7 in pET (Figure 2.1). All three vectors were co-transformed into BL21 Rosetta cells. A single colony was picked to express all six hMcm2-7 in 4.5 L of LB medium using kanamycin (15 μg/ml), ampicillin (25 μg/ml) and spectinomycin (25 μg/ml). The cultures were incubated at 37 ºC with shaking until an OD_{600} of 0.887 was reached. At that point, the cultures were incubated in an iced water bath for 30 min. IPTG was then added to a A



B

C

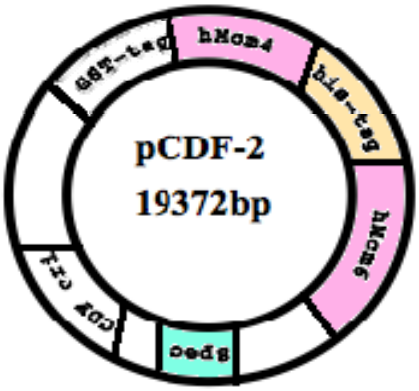


Figure 2.1 Constructs used for co-expression of human Mcms: (A) pET-32 was generated by ligating Mcm2 and Mcm7. (B) pRSF-2 was generated by ligating Mcm3 and Mcm5 (C) pCDF-2 was generated by ligating Mcm4 and Mcm6. Each of the Mcms has a different N-terminal affinity tag for protein.

final concentration of 0.1 mM. The cultures were then incubated for 24 hours at 12 ºC with 180 rpm shaking. The cells were centrifuged for 10 min

at 4,600 x *g* (Eppendorf, model 5810 using swing out rotors A- 4-81) at 4 ºC, yielding a single pellet of 18 g, which was stored at -80 ºC until further use. A process flow diagram (PFD) is shown for protein purification steps of hMcm complex (Figure 2.2)

2.3.1 Sample Preparation

The pellet was immersed in a beaker of water at room temperature to thaw and once thawed kept on ice. The pellet of cells was re-suspended in 90 ml nickel binding buffer (25 mM Tris pH 7.6, 150 mM NaCl, 3 mM KCl, 40 mM imidazole 10% glycerol, and freshly added 1 µg/ml Pepstatin A, 1 µg/ml Aprotinin, 1 µg/ml Leupeptin, 0.5 µg/ml PMSF, 1 mM 2-mercaptoethanol). The cell suspension was then passed through an automated cell disruptor (20 Kpsi) once and pooled. Next, the following reagents were added: 6 mM $MgCl_2$, 0.75 mg/ml lysozyme, 10 µg/ml DNase I and 5 µg/ml RNase A and the lysate was incubated for 10 min at room temperature with inversion. Next, the lysate was clarified by centrifugation for 30 min at 50,000 x g (High Speed Sorvall Evolution, rotor F34) and 4 ºC.

2.3.2 Ammonium sulphate precipitation

Saturated ammonium sulphate was prepared by adding 110 g of ammonium sulphate to 200 ml of nickel binding buffer (minus protease inhibitors and 2-ME), heating the solution to ~80 ºC and then cooling it to 4 ºC overnight while stirring. An equal volume of saturated ammonium sulphate solution was added to the clarified lysate and then incubated on ice for 30 min (with gentle mixing by inversion every 10 min) to precipitate the Mcm complex. The protein precipitate was centrifuged for 20 min for 10,000 x *g* (Eppendorf, rotor F34-6-38) at 4 ºC. The supernatant was discarded and the pellets were stored at -80 ºC until required for further purification.

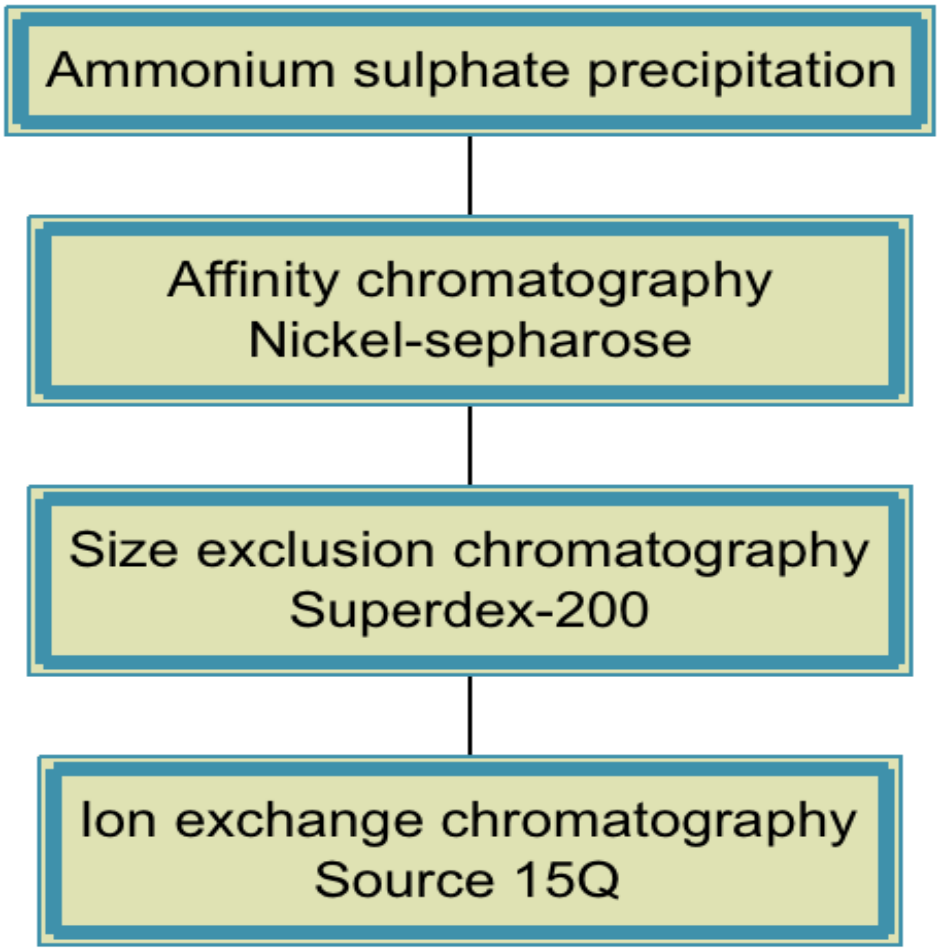


Figure 2.2 A schematic representation of protein purification steps: The hMcm2-7 complex was purified over the range of chromatographic purification after ammonium sulphate precipitation as a first step of purification.

2.3.3 Nickel column

The precipitated protein pellet was re-dissolved in nickel binding buffer to a volume 0.5x that of the original clarified lysate and then passed through a 0.45 μm filter. The first chromatographic step of the purification was a 1 ml nickel sepharose HP column with an imidazole step elution. A flow rate of 0.5 ml/min was maintained throughout the nickel column step. First, the column was equilibrated with 5 column volumes (CV) of nickel binding buffer. Next, the sample was loaded onto the column, which was then washed with 10 column volumes of nickel binding buffer. Finally, the bound proteins were eluted using 100% nickel elution buffer (25 mM Tris pH 7.6, 150 mM NaCl, 3 mM KCl, 500 mM imidazole 10% glycerol, and freshly added 1 μg/ml Pepstatin A, 1 μg/ml Aprotinin, 1 μg/ml Leupeptin, 0.5 μg/ml PMSF, 1 mM 2-mercaptoethanol). All 0.5 ml fractions were pooled together and further purified by size exclusion chromatography.

2.3.4 Size exclusion chromatography

A HiLoad 200 Superdex 16/60 column (GE Healthcare) was connected to an Äkta FPLC system at 4 ºC. The column was first equilibrated with > 1.5 column volumes of GF buffer (25 mM tris pH 7.6, 150 mM NaCl, 3 mM KCl, 10% glycerol, 1 mM EDTA, 1 mM EGTA and freshly added 1 μg/ml Pepstatin A, 1 μg/ml Aprotinin, 1 μg/ml Leupeptin, 0.5 μg/ml PMSF, 1 mM DTT). The nickel elution pool sample was loaded into a 2 ml loop and injected onto the column. Proteins were eluted isocratically with GF buffer at 1 ml/min and collected in a series of 2 ml fractions. The fractions containing high molecular mass proteins as indicated by the first peak on the A_{280} chromatogram (i.e. A3-A6 8 ml) were pooled ready for the next chromatography step.

2.3.5 Anion exchange

The final chromatographic step utilised a 1 ml Source 15Q anion exchange column. 1ml/min flow rate was maintained through out this purification step. 2 volumes (16 ml) of sample dilution buffer were added to 1 volume of size exclusion pool (8 ml) in order to reduce the NaCl concentration to 50 mM. This sample was then loaded into a 50 ml super-loop. The column was

rinsed with water, charged with Cl^- ions with 5 column volumes of AIEX elution buffer (25 mM tris pH 7.6, 1000 mM NaCl, 3 mM KCl, 10% glycerol, 1 mM EDTA, 1 mM EGTA and freshly added 1 µg/ml Pepstatin A, 1 µg/ml Aprotinin, 1 µg/ml Leupeptin, 0.5 µg/ml PMSF, 1 mM DTT) and equilibrated with 5 CV of AIEX start buffer (25 mM tris pH 7.6, 50 mM NaCl, 3 mM KCl, 10% glycerol, 1 mM EDTA, 1 mM EGTA and freshly added 1 µg/ml Pepstatin A, 1 µg/ml Aprotinin, 1 µg/ml Leupeptin, 0.5 µg/ml PMSF, 1 mM DTT). The sample was then injected onto the column. Unbound material was then washed out of the column with 5 column volumes of AIEX start buffer. The proteins were then eluted with a linear gradient of NaCl from 50 to 1000 mM over 20 CV and collected in 1 ml fractions.

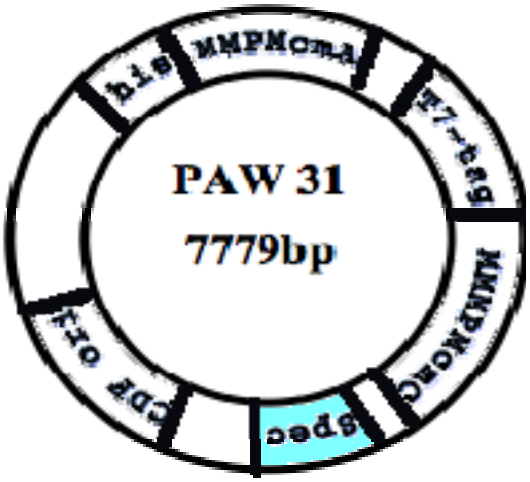
2.4 Purification of *M. maripaludis* Mcms

Dr Alison Walters in the Chong group provided two constructs used in this section. One construct pAW 31 contained His-tagged MMP McmA and T7 tagged MMP McmC in pCDF-2 while the other construct pAW 32 contained Trx-tagged MMP McmB and GST tagged MMP McmD in pET32 EK/LIC (Figure 2.3).

2.5 Cell culture

5 ml of culture was grown overnight by the addition of a single colony to LB medium containing appropriate antibiotics. This was used to inoculate 750 ml of LB containing appropriate antibiotics and grown to optical density of OD_{595nm} = 0.5-0.7. At this point 0.5 mM IPTG was added to the culture. Cultures were incubated for a further 24 hours with shaking at 180 rpm at 12 ^oC.

A



B

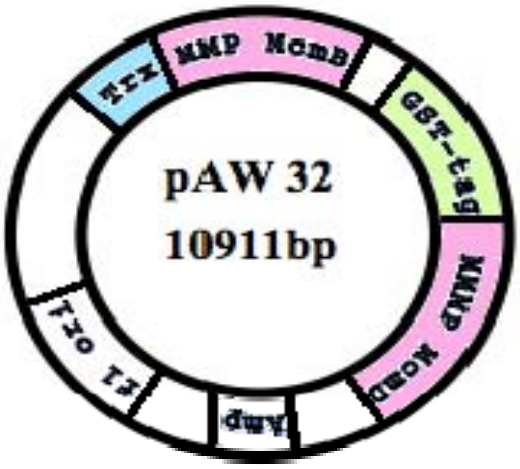


Figure 2.3 Constructs used for purification of *M. maripaludis* Mcms: (A) pAW 31 containing His-tagged MMP McmA and T7 tagged MMP McmC in pCDF-2 (B) pAW 32 containing Trx-tagged MMP McmB and GST-tagged MMP McmD in pET32 EK/LIC.

2.5.1 Protein purification

2.5.1.1 His-tagged protein purification

A 4 g pellet was resuspended in 5 ml/g of lysis buffer (50 mM KPi pH 7.8, 200 mM NaCl, 20% Glycerol, 10 mM Imidazole, 1x mini Roche protease inhibitor tablet, MilliQ water). 0.75mg/ml of lyzozyme was added to the mixture and was incubated at room temperature for 30 minutes to facilitate lysis. Samples were then sonicated for 3x 30 seconds at 100% power on ice. Insoluble material was removed by centrifugation at 30,000 x *g* (High Speed Sorvall Evolution, rotor F34) for 20 minutes at 4 ^{o}C.

Ni-NTA resins were used to purify his-tagged MMP Mcms. Nickel beads were prepared for purification by washing with 5 resin volumes of wash buffer (50 mM KPi pH 7.8, 200 mM NaCl, 20% Glycerol, 50 mM Imidazole, 1x mini Roche protease inhibitor tablet, MilliQ water). These resin was transferred into 6 x 2 ml disposable gravity columns (Pierce) along with an equal volume of clarified cell lysate in each column and left for 2 hour on the rotor at 4 ^{o}C. After washing with 20 CV of wash buffer, the protein was eluted with 0.5 CV of elution buffer (50 mM KPi pH 7.8, 200 mM NaCl, 20% Glycerol, 500 mM Imidazole, 1x mini Roche protease inhibitor tablet, MilliQ water) at first and then 2 CV afterwards. Samples were analysed on 10% SDS-PAGE gels and by western blot. Protein concentrations were determined as described in section 2.2.16. Eluant fractions were pooled, snap frozen in liquid nitrogen and stored immediately at -80 ^{o}C

2.5.1.2 GST-tagged protein purification

A 3 g pellet was resuspended in 5 ml/g of lysis buffer (50 mM Tris-Cl pH 8.5, 250 mM NaCl, 1 mM EDTA, fresh 1 mM DTT). 0.75 mg/ml lyzozyme and 1mM PMSF were added and the mixture was incubated on ice for 30 minutes to facilitate lysis. Samples were then sonicated for 5 pulses (50 W, 20 kHz) on ice after adding 1% sarcosyl. Insoluble material was removed by centrifugation at 30,000 x *g* (High Speed Sorvall Evolution, rotor F34) for 20 minutes at 4 ^{o}C. 12 ml of clarified cell lysate was transfered into 6 disposable gravity columns (Pierce) containing 200 µl of glutathione-sepharose to purify GST-tagged MMP Mcms proteins. Cell lysate was

incubated with slurry for 1 hour at 4 ^{o}C on a shaker. After washing with 0.5 % Triton X-100, the protein was eluted with 0.5 CV of elution buffer (50 mM glutathione, 50 mM Tris-Cl pH 8.5, 250 mM NaCl, 1 mM EDTA,) at first and then 2 CV afterwards. Samples were analysed on 10% SDS-PAGE gels and by western blot using antibiotics as detailed below. Protein concentrations were determined as described in section 2.9.7 Eluant fractions were pooled to subject to anion exchange (section 2.3.5)

2.6 Western blotting

Western blot were carried out using the transblot-SD semi dry transfer cell (BioRad). Proteins were separated using SDS polyacrylamide gel electrophoresis (SDS-PAGE) providing information about molecular weight using pre-stained protein molecular markers (BioRad) so that efficient transfer could be easily confirmed. One sponge and 2 pieces of 3 mm filter paper (Whatman) soaked in anode buffer 1 (0.3 M Tris pH 10.4, 10% methanol) were placed on the anode plate. Another piece of filter paper soaked in anode buffer 2 (25 mM Tris pH 10.4, 10% mathanol) was placed and above that pre-soaked PVDF membrane (GE, Heathcare) in methanol was added. After placing the SDS polyacrylamide gel on top of PVDF membrane, 3 whatman filter paper and sponge soaked in cathode buffer (25 mM Tris pH 9.4, 40 mM glycine, 10% methanol) were added. This stack was placed in the transfer apparatus and the transfer was carried out at 70 mA for an hour. The membrane was separated from the rest of the stack after transfer and stained in 1x Ponceau S (BioRad). At this stage lane positions were marked. The membrane was then washed 2x for 20 minutes in wash buffer TBS (10 mM Tris pH 8.0, 150 mM NaCl) in case of Anti-TetraHis (Qiagen) and Anti-GST (Novagen) or TBS-T wash buffer (TBS, 0.1% Tween 20) for Anti-S (Novagen) and Anti-T7 (Novagen). The membrane was blocked for an hour in 20-30 ml of blocking buffer. For Anti-TetraHis and Anti-GST 3% and fat-free dried 10% milk powder in TBS was used respectively and for Anti-S and Anti-T7 5% fat-free dried milk powder in TBS-T was used. Membranes were then washed 3x for 10 minutes and incubated in primary antibody overnight at room temperature. For Anti-TetraHis and Anti-GST 1:5000 and 1:3000 with 5% fat-free dried milk

powder in TBS was used and for Anti-S and Anti-T7 1:5000 and 1:2000 TBS-T was used. Membranes were then washed 6x 5 minutes with wash buffer (Table 2.1) and then incubated in secondary antibody for 1 hour. Anti-rabbit-HRP 1:5000 in 10% fat-free dried milk powder TBS-T for Anti-TetraHis and anti-mouse HRP (Zymed) 1:10000 TBS-T for Anti-S and Anti-T7 and Anti-GST was used. Membranes were washed 3x 10 minutes in 50 ml of wash buffer accordingly (Table 2.1). Protein was then detected using ECL Kit (Pierce) and wrapped in plastic cling film before exposing to photographic film (GE Healthcare).

Antibody	Wash buffer	Blocking buffer	Primary antibody dilution	Secondary antibody
Anti-TetraHis (Qiagen)	TBS	3% BSA in TBS	1:5000, 3% BSA in TBS	Anti-rabbit-HRP, 1:5000 in 10%milk TBS-T
Anti-S (Novagen)	TBS-T	5% milk in TBS-T	1:5000 TBS-T	Anti-mouse-HRP (Zymed), 1:10000 TBS-T
Anti-T7 (Novagen)	TBS-T	3% BSA in TBS-T	1:20000 TBS-T	Anti-mouse-HRP (Zymed), 1:10000 TBS-T
Anti-GST (Novagen)	TBS	10% milk in TBS	1:4000 5% milk in TBS	Anti-mouse-HRP (Zymed), 1:20000 TBS-T

Table 2.1 Antibody and buffer solutions used for western blots:
TBS: 10 mM Tris pH 8.0, 150 mM NaCl; TBS-T: TBS plus 0.1% Tween 20.

2.7 Substrate preparation for helicase assays

A forked substrate was generated by annealing a ^{32}P-labelled 57 base oligonucleotide (HS2) and a non-labelled 74 base oligonuleotide (HS1) to produce a forked DNA substrate with a 25 bp double-stranded region. (Figure 2.4). 50 pmol of 57 base substrate was radiolabelled using 3 µl of

Chapter 3: Biochemical studies on human Mcms

3.1 Introduction

Initiation of eukaryotic DNA replication requires the assembly of multiprotein complexes at the replication origin (Stillman, 1996). The minichromosome maintenance proteins form a hexameric complex required at the time of DNA replication initiation and provide the replicative DNA helicase activity that includes DNA dependent ATPase activity (Chong et al., 2000). Mcm helicase complexes are conserved in all eukaryotic organisms and are composed of six different but highly related proteins. Mcm complexes can be isolated from eukaryotic cells as several stable subassemblies (Thommes et al 1997). Weak DNA helicase activity was identified in a human complex of a dimer of trimers containing Mcm4/6/7 that was only sufficient to displace some 30bp of short oligonucleotides from complementary DNA templates (Ishimi, 1997). Recently it has been demonstrated that *S. cerevisiae* Mcm2-7 complex expressed in baculovirus-infected insect cells has helicase activity *in vitro* (Bochman and Schwacha, 2008). However to date no studies from human Mcms expressed in *E. coli* have shown helicase activity. In order to achieve robust activity of hMcm, we expressed and purified hMcms from *E. coli* so that proteins would not contain any post-translational modifications that are present in protein purified from eukaryotic sources. We tested this complex for helicase activity. Furthermore we investigated the affect of phosphorylation on helicase activity of hMcm2-7 complex.

3.2 Cloning, expression and purification of individual human Mcms

3.2.1 Plasmid construction

Human Mcms 2-7 cDNAs cloned into pBluescript SK^+ vector were provided by the Mendez lab (CNIO Spain). The hMcm genes (2-7) were subcloned into the expression vector pET 47b (Novagen) using following restriction sites. hMcm2 was digested with *BamH1* and *Xho1* and inserted into *BamH1*

and *Xho1* sites of pET 47b to yield a pET 47b-Mcm2 construct. Mcms 3, 4, 5, and 7 were digested with *EcoR1* and *Xho1* and were inserted into *EcoR1* and *Xho1* of pET 47b to yield pET 47b-Mcm3, pET 47b-Mcm4, pET 47b-Mcm5 and pET 47b-Mcm7 constructs. Mcm6 was digested with *Xho1*and *Sal1* and inserted into *Xho1*and *Sal1* sites of pET 47b to yield pET 47b-Mcm6 construct. Sequencing results showed that all of the genes were ligated correctly in vector pET 47b but were out of frame, so site directed mutagenesis for all Mcms was performed (See section 2.2.9 in materials and methods) to place them in the correct reading frame for expression. A summary of human Mcms proteins properties is provided in table 3.1.

3.2.2 Expression

3.2.2.1 Small-scale expression

Initial expression and solubility studies were carried out by growing 50 ml cultures of individual proteins. Mcms were found to over express well in small-scale cultures but the majority of proteins were packaged into inclusion bodies (IBs). For increasing solubility, different temperature trials were carried out after induction. Amongst 12 ^{o}C, 16 ^{o}C, 25 ^{o}C and 37 ^{o}C, it was found out that for Mcm4, 12 ^{o}C was best and for rest of the Mcms 37 ^{o}C was working.

To increase the solubility of recombinant proteins different strains were used in trials. These were Rosetta BL21, Rosetta BL21 (DE3), Arctic Express (DE3) RIL and Arctic Express (DE3) pLysS. Rosetta host strains are designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli* (Brinkmann et al., 1989; Seidel et al., 1992; Kane, 1995; Kurland and Gallant, 1996). While DE3 indicates that the host is a lysogen of λDE3, and therefore carries a chromosomal copy of the T7 RNA polymerase gene under control of the *lacUV5* promoter (Studier et al., 1990; Dubendorff and Studier, 1991a; Dubendorff and Studier, 1991a). Arctic Express (DE3) RIL cells are optimised for expression of recombinant proteins at low temperature because they contain a plasmid expressing the

hMcm	Gene length (bp)	Residues (aa)	Mol. weight (Da)	pI
hMcm2	2715	904	101891	5.34
hMcm3	2427	808	90977	5.53
hMcm4	2592	863	96552	6.28
hMcm5	2209	734	82283	8.64
hMcm6	2466	821	92883	5.29
hMcm7	2160	719	81303	6.08

Table 3.1 Summary of human Mcm2-7 proteins properties: Properties described are for untagged proteins.

chaperons Cpn10 and Cpn60 from the psychrophillic bacterium *Oleispira antactica*. These chaperones are homologues of the *E.coli* chaperones GroEL and GroES, and are adapted to function optimally at 4-12 ^{o}C and therefore allow protein folding at low temperature (Lund, 1994). pLysS strains express T7 lysozyme, which further suppresses basal expression of T7 RNA polymerase prior to induction (Moffatt and Studier, 1987; Studier, 1991). There was not any significant difference in solubility of Mcms between these strains therefore Rosetta BL21 (DE3) was widely used in these studies.

Different buffers were used throughout this study for protein purification. A buffer system that worked for one Mcm did not necessarily worked for other Mcms, so each of the buffer systems is discussed with reference of individual purified Mcm in the respective sections.

Furthermore expression vectors were also considered to change. For this purpose Mcm5 and Mcm6 were individually ligated into pQE-31 vector but only Mcm6 was purified this way (discussed later in section 3.2.5).

For SDS-polyacrylamide gel, samples were prepared by addition of 20 µl benzonase nuclease (Novagen) 6x loading buffer into 100 µl total protein. Samples were boiled for 5 minutes at 100 ^{o}C before loading on a 10% SDS-polyacrylamide gel (Figure 3.1).

3.2.2.2 Large-scale expression

For large-scale expression, 750 ml of media with appropriate antibiotics was used to grow cultures in a 2 litre flask.

3.2.3 Protein purification

Proteins were purified either manually by using disposable gravity columns (Pierce) or a Tricorn 5/50 column (GE Healthcare) using an AKTA prime automated purification system. Mcm6 and Mcm7 were purified from 750 ml of culture using the same set of buffers while Mcm4 was purified from 4.5 L

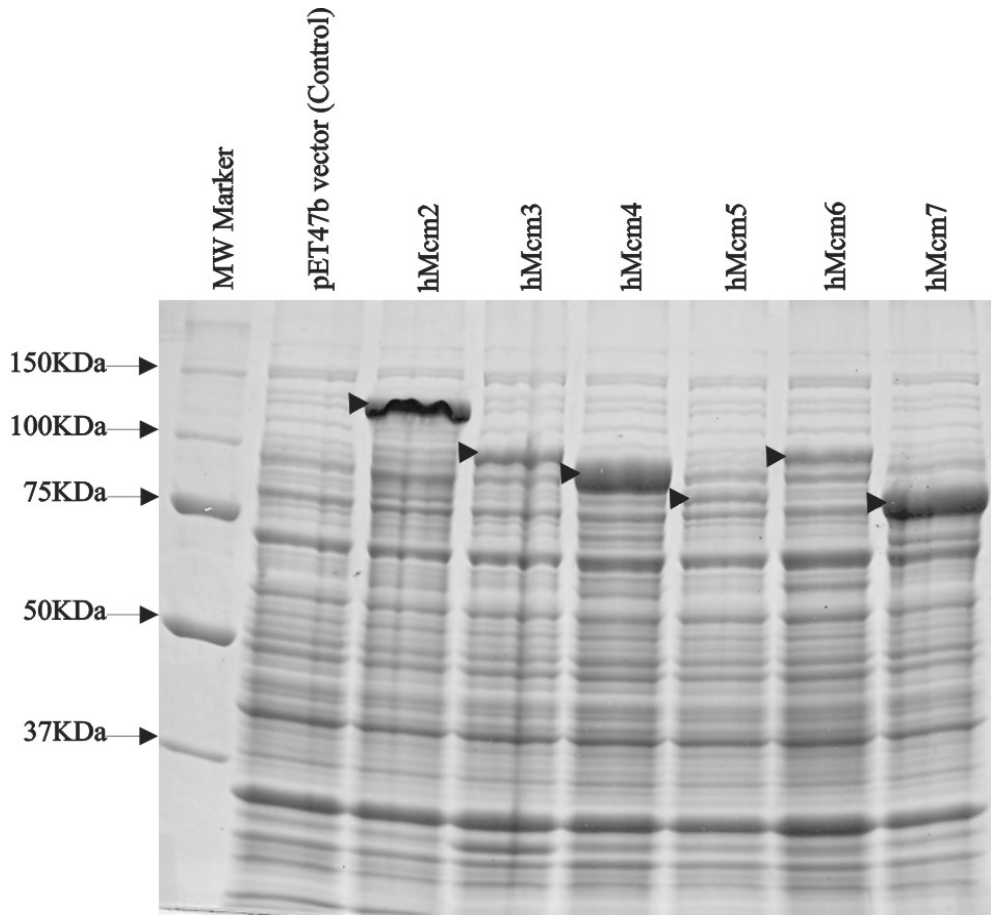


Figure 3.1 Expression of hMcm2-7: Expression trials of individually $(His)_6$-tagged hMcms 2-7 in autoinduction medium from total protein fractions showed that all individual proteins were expressed in *E. coli.* Individual proteins were expressed in Rosetta *E.coli* BL21 (DE3) cells at 37 ^{o}C. Small arrowheads near each protein indicate the individual protein band corresponding to their sizes (See table 3.1).

of culture with different set of buffers (Section 3.2.4).

Differences in each of the hMcm purifications are discussed under each section of the respective protein purifications.

3.2.4 Expression and purification of hMcm4

Initial expression and purification trials indicated that hMcm4 had low solubility reducing the amount of protein that could be purified from cell extract. Therefore this protein was expressed in 4.5 litre of culture rather than 750 ml as for the other Mcms. Recombinant Mcm4 was expressed at 12 ºC in Rosetta DE3 cells and purified from a 4.5 litre culture that yielded a cell pellet weighing 21 grams.

As nickel gives different binding characteristics from cobalt. i.e. Ni^{2+} is the metal ion considered to have the strongest affinity to histidine-tagged proteins and may therefore allow binding and washing under more stringent conditions (higher imidazole) and therefore removal of more contaminants that could have purified due to expressing protein at 12 ºC in Rosetta DE3 cells. It was hope that an imidazole gradient will elute non-specifically bound proteins before the hMcm4 was eluted.

The buffers used for the hMcm4 purification were different to those used for the other Mcm purifications. The pellet was resuspended in 40 ml of nickel binding buffer (50 mM sodium phosphate pH 7.4, 0.5 M NaCl, 10% glycerol, 30 mM Imidazole, 1 mM 2-mercaptoethanol and 1x Roche complete EDTA free protease inhibitor mixture). The cells were lysed by one pass through a manual French press rather then sonication.

A Tricorn 5/50 column (GE Healthcare) column was used for washing and elution. The bound protein was eluted with a linear gradient of 0-100% elution buffer H (50 mM sodium phosphate pH 7.4, 0.5 M NaCl, 10% glycerol and 500 mM imidazole) over a period of 40 min at 5 ml/min. The eluate was collected in a series of 2 ml fractions, samples of which were run

on 10% SDS-PAGE gels. The ~ 57 KDa *E.coli* chaperone GroEL protein was present in large amounts (Fraction containing GroEL protein not shown in the gel) although it appeared to be washed from the column before hMcm4 had finished eluting. The SDS-PAGE gel showed a very clean band at the expected size of hMcm4 (96 KDa) just near the 100 KDa protein marker (Figure 3.2). Samples containing only hMcm4 were pooled together and concentration of which was measured. The final volume of protein was 10 ml and the concentration was calculated to be 15. 1 µg/ml by Bradford assay (See section 2.2.16 in Materials and Methods).

3.2.5 Expression and purification of hMcm 6

The gene for hMcm6 (which was initially ligated into pET47b) was digested with *BamH1* and *Sal1* and inserted into *BamH1* and *Sal1* sites of pQE-31 to yield pQE-31- hMcm6 construct.

It was noted that no induction was required for pQE-31- hMcm6. The pQE-31 vector uses a T5 promoter as opposed to a T7 promoter used commonly in other expression systems. The T5 promoter is recognized by the *E. coli* RNA polymerase and the lac operator sequence, which binds lac repressor and ensures efficient repression of the powerful T5 promoter in *E. coli*. pQE-31-hMcm6 construct was sequenced before expressing protein and was confirmed that Mcm6 gene was full length.

hMcm6 was expressed from 750 ml of culture, yielding 1.5 gram of pellet and purified under the conditions described in section 2.2.13 using the same set of buffers. Elution fraction samples 1 and 2 were run on 10% SDS-PAGE gels (Figure 3.3) and remaining fractions were flash frozen in liquid nitrogen and stored at -80 ^{o}C. Protein bands were visible in elution fraction 1 and elution fraction 2, just below the expected size for hMcm6 (93 KDa). Purified hMcm6 protein ID was confirmed by mass spectrometry (MS). The possible explanation of hMcm6 mobility being smaller then expected size in SDS-PAGE gel is its degradation during expression and purification steps.

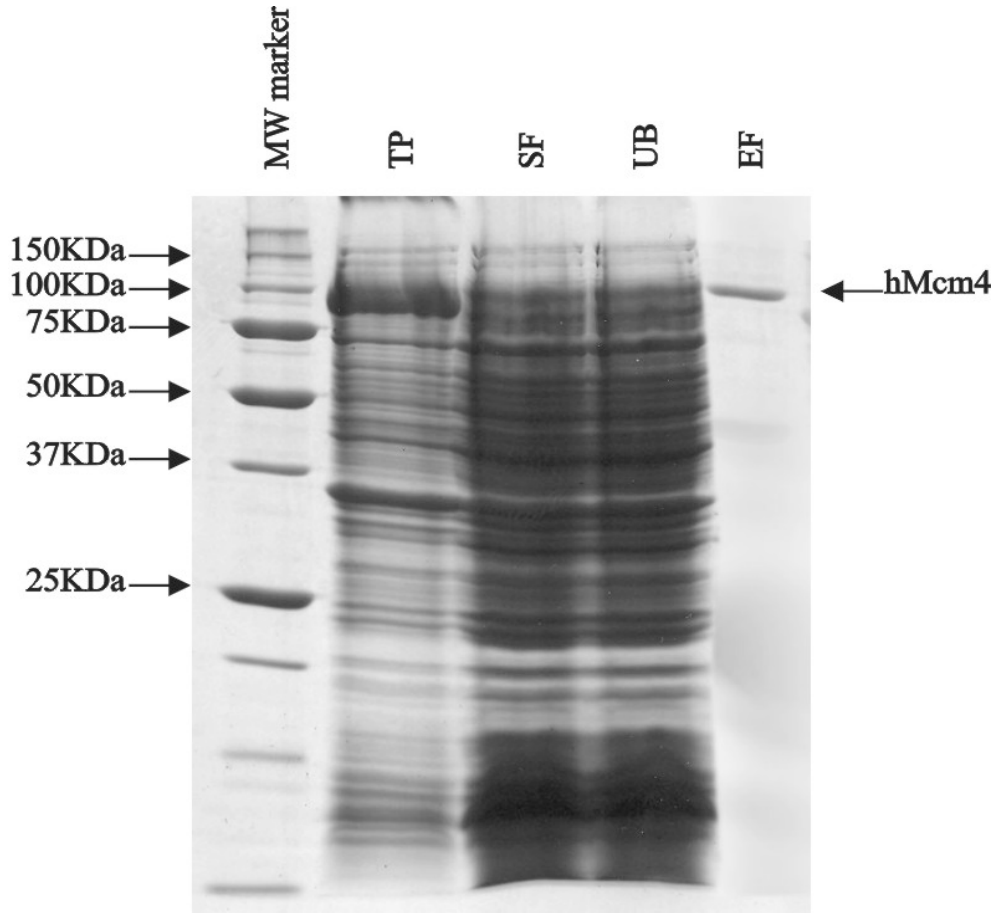


Figure 3.2 Purification of hMcm4: SDS-PAGE analysis of fractions from purification trials visualised by Coomassie blue staining. Protein was expressed in *E.coli* BL21 (DE3) Rosetta cells at 12 ^{o}C overnight. Pellets from 4.5 L of culture was resuspended, and the soluble protein fraction was bound, washed and eluted from nickel beads using 500 mM imidazole. hMcm4 protein band of 96 KDa in the elution fraction is visible. Protein ID was confirmed by mass spectrometry (MS).

TP, total protein; SF, soluble fraction; UB, unbound fraction; EF, elution fraction

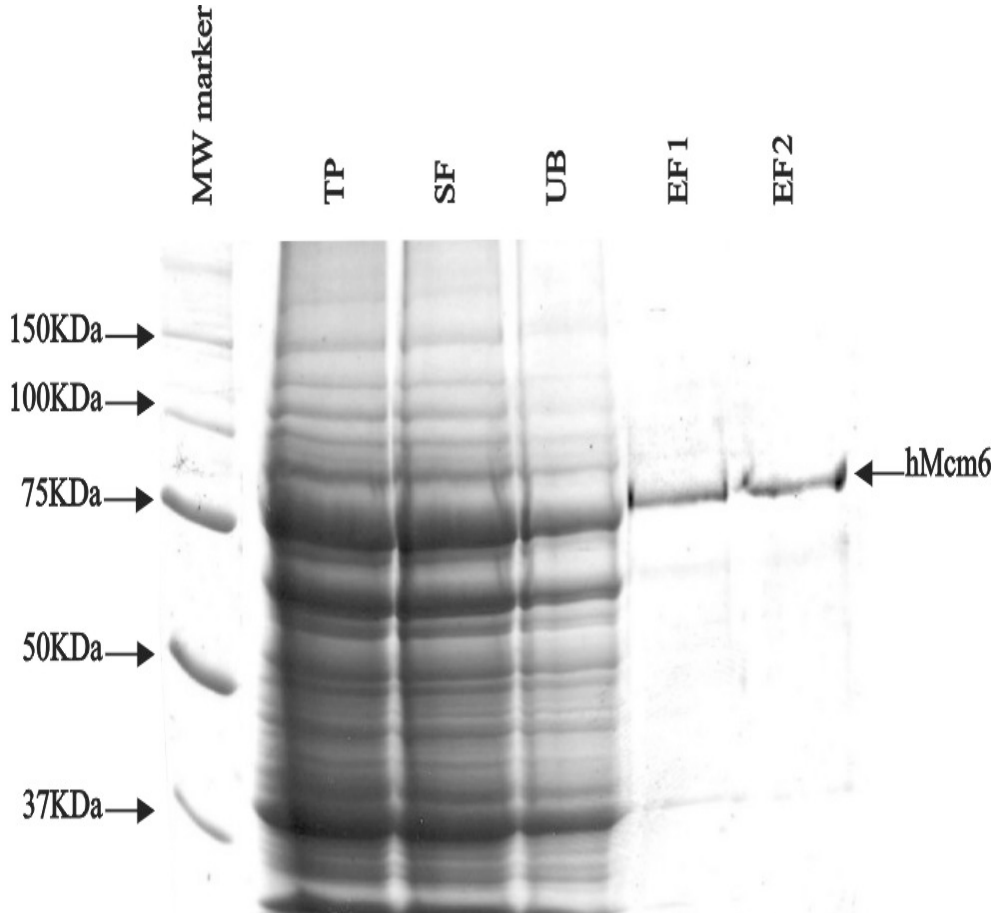


Figure 3.3 Purification of hMcm6: SDS-PAGE analysis of fractions from expression trials after Coomassie blue staining. Protein was expressed in *E.coli* BL21 (DE3) Rosetta cells at 37 ^{o}C overnight. 750 ml culture pellet was resuspended and the soluble fraction bound, washed and eluted from Talon beads using 150mM imidazole. hMcm6 protein band is visible around 75 KDa in the elution fraction below its expected size of 93 KDa . The protein ID was confirmed by mass spectrometry. This protein might have degraded during purification steps thus the mobility is smaller then expected size.

TP, total protein; SF, soluble fraction; UB, unbound fraction; EF, elution fraction

The final volume of protein was 5 ml and the protein concentration was calculated to be the 15.8 µg/ml by Bradford assay (See section 2.2.16).

3.2.6 Expression and purification of hMcm7

hMcm7 was expressed in 500 ml culture which yielded 2 gram of pellet. hMcm7 protein was purified under the conditions described in section 2.2.13 using the same set of buffers. Samples of the protein eluted with elution buffer were run on 10% SDS-PAGE gels and eluant fractions were stored at -80 ^{o}C. The SDS-PAGE gel showed a clean band at the expected size for hMcm7 (81 KDa) near the 75 KDa protein marker (Figure 3.4). The final volume of protein was 6 ml and the concentration was calculated to be 15.5 µg/ml by Bradford assay (See section 2.2.16).

3.2.7 Expression and purification of hMcm2, 3, 5

Despite trying several trials that included changes in buffers for purification, expression conditions, different temperatures for purification of hMcm2, 3 and 5, all attempts were largely unsuccessful. hMcm2 showed a very small amount of solubility in *E.coli* Arctic Express (DE3) RIL cells but this was not enough to carry out further assays. Also for further assays all 6 proteins needed to be purified.

3.3 hMcm 4/6/7 complex

An attempt to assemble a first complex of hMcm4/6/7 was not successful. The apparent reason seemed to be proteolysis, but the addition of protease inhibitors did not make any difference.

The concentration of the proteins purified was too low (hMcm4 was 15.1 µg/ml, hMcm6 was 15.8 µg/ml and hMcm7 was 15.5 µg/ml) to perform further assays on them. As all six Mcm proteins were needed for helicase and further assays, there was no point to concentrate these proteins.

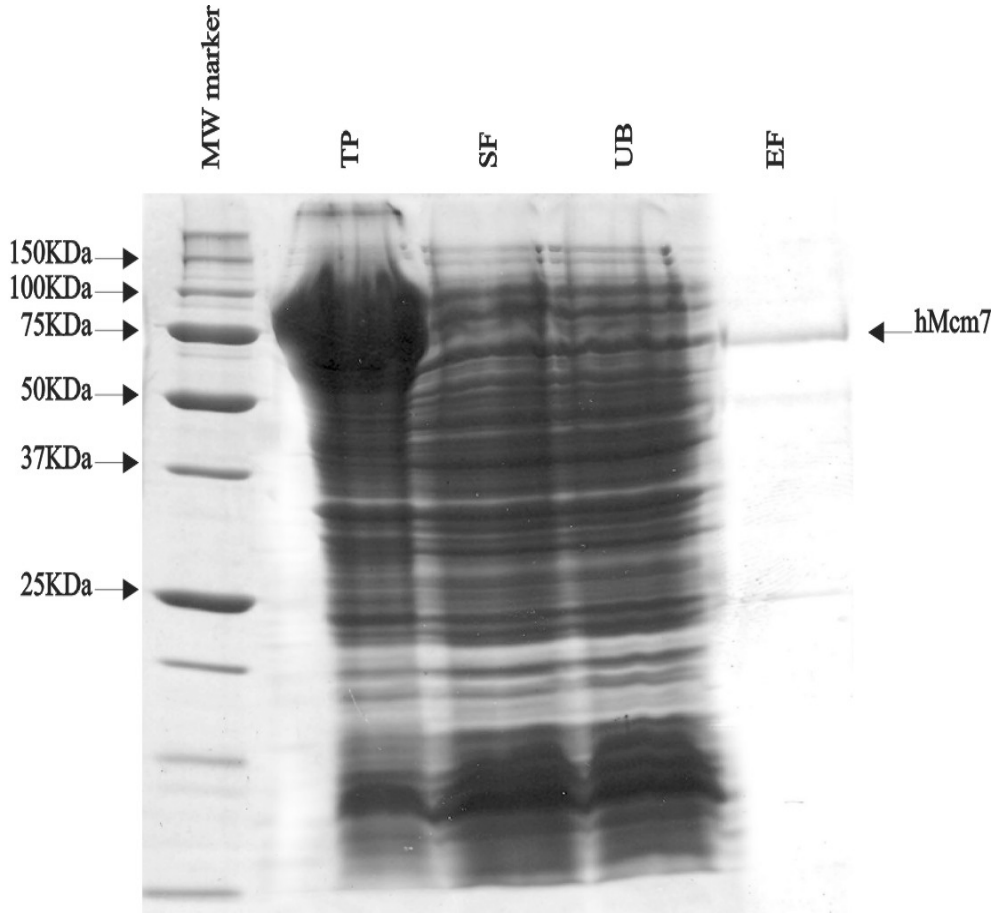


Figure 3.4 Purification of hMcm7: SDS-PAGE analysis of fractions from expression trials after Coomassie blue staining. Protein was expressed in *E.coli* BL21 (DE3) Rosetta cells at 37 ^{o}C overnight. The pellet from a 500 ml culture was resuspended and eluted from Talon beads using 150 mM imidazole. A protein band of 81 KDa in the elution fraction was visible at the expected size for hMcm7. This band was later confirmed by mass spectrometry.

TP, total protein; SF, soluble fraction; UB, unbound fraction; EF, elution fraction

Results of expression and purification trials showed that hMcm7 expressed at 37 ^{o}C. However under these conditions other hMcms were not present in the soluble fraction of cell extracts. hMcm6 could be purified only after expressing in pQE-31- hMcm6 construct under the same buffers as described for hMcm7 but the same approach for hMcm5 did not yield purified protein. Purification of hMcm4 was possible by expressing it at low temperature in *E.coli* BL21 (DE3) Rosetta cells in 4.5 L of culture.

Even after all these attempts the level of purified protein was very low. Eventually Mcm2 was also purified using low temperature after induction. As we hypothesized that all six were necessary for robust helicase activity therefore without the purification of hMcm3 and 5, further assays could not be performed. It was then decided that the individual protein purification approach should be discontinued.

3.4 Cloning, co-expression and co-purification of human Mcms

The expression and purification of individual hMcms showed that hMcm 2, 3 and 5 were insoluble and difficult to purify. As all of the effort to get them in soluble form remained largely unsuccessful. Also the concentration obtained from hMcm 4, 6 and 7 proteins was small in amount and varying in purity making it very difficult to proceed towards reconstituting a complex. Furthermore the attempts to make a hMcm4/6/7 complex failed so it was decided to conduct a trial to co-express and co-purify all these proteins for further assays.

Co-expression of interacting proteins can increase the solubility of recombinant proteins. As Mcm2-7 interacts with each other (Davey et al., 2003), therefore co-expression and co-purification could provide an increase in solubility of insoluble Mcms. Often the separate components of a complex are not soluble mostly due to hydrophobic patches that are exposed to the solvent (Damodaran, 1986). These patches are usually involved in and protected by the binding to the other component of the complex (Kurochkina and Mesyanzhinov, 1999). Therefore, co-expression of the different

components could result in a higher expression level of soluble protein complex.

3.4.1 Plasmids construction for co-expression of hMcms

This work has been carried out in collaboration with Dr Richard Parker, a post-doctoral researcher in the Chong group. For hMcm2-7 subunit cloning, LIC vectors (Novagen) were used. cDNAs encoding the six hMcm subunits (hMcm2-7) were cloned in pairs into three pET Duet vectors (Novagen) to allow co-expression of all six subunits in *E. coli*. First, each cDNA was PCR amplified from the constructs provided by Mendez lab (CNIO Spain) using primers with 5' sequences designed to be compatible with the ligation independent cloning (LIC) system employed by the Duet vectors. The amplicons were gel purified and treated with T4 DNA polymerase in the presence of dATP in order to recess the 3' ends. Each pair of hMcm cDNAs was annealed with a Duet vector backbone and an adaptor DNA segment according to the Novagen LIC kit directions. hMcm2 and hMcm7 cDNAs were annealed with the pET32 Ek/LIC vector and the T7 Ek/LIC adaptor to give: pET32-hMcm2 and hMcm7. hMcm6 and hMcm4 cDNAs were annealed with the pCDF-2 Ek/LIC vector and the GST Ek/LIC adaptor to give: pCDF-2-hMcm4 and hMcm6. hMcm5 and hMcm3 cDNAs were annealed with the pRSF Ek/LIC vector and the minimal Ek/LIC adaptor to give: pRSF- hMcm3 and hMcm5 (See Appendix 2 for primers sequences). A detailed modified diagram of the predicted interactions of the six recombinant Mcm proteins in a heterohexameric complex from Davey et al., (2003) with different N-terminal tags that were used in affinity purification was created (Figure 3.5). Overall three vectors encoding six different proteins were generated. One vector contained hMcm2 and hMcm7 with an N-terminal T7-tag on hMcm2 and an N-terminal S/his/trx-tag on hMcm7. A second vector contained hMcm4 and hMcm6 with an N-terminal GST-tag on hMcm4 and an N-terminal his-tag on hMcm6. The third vector contained hMcm5 and hMcm3 with an-N terminal his-tag on hMcm5.

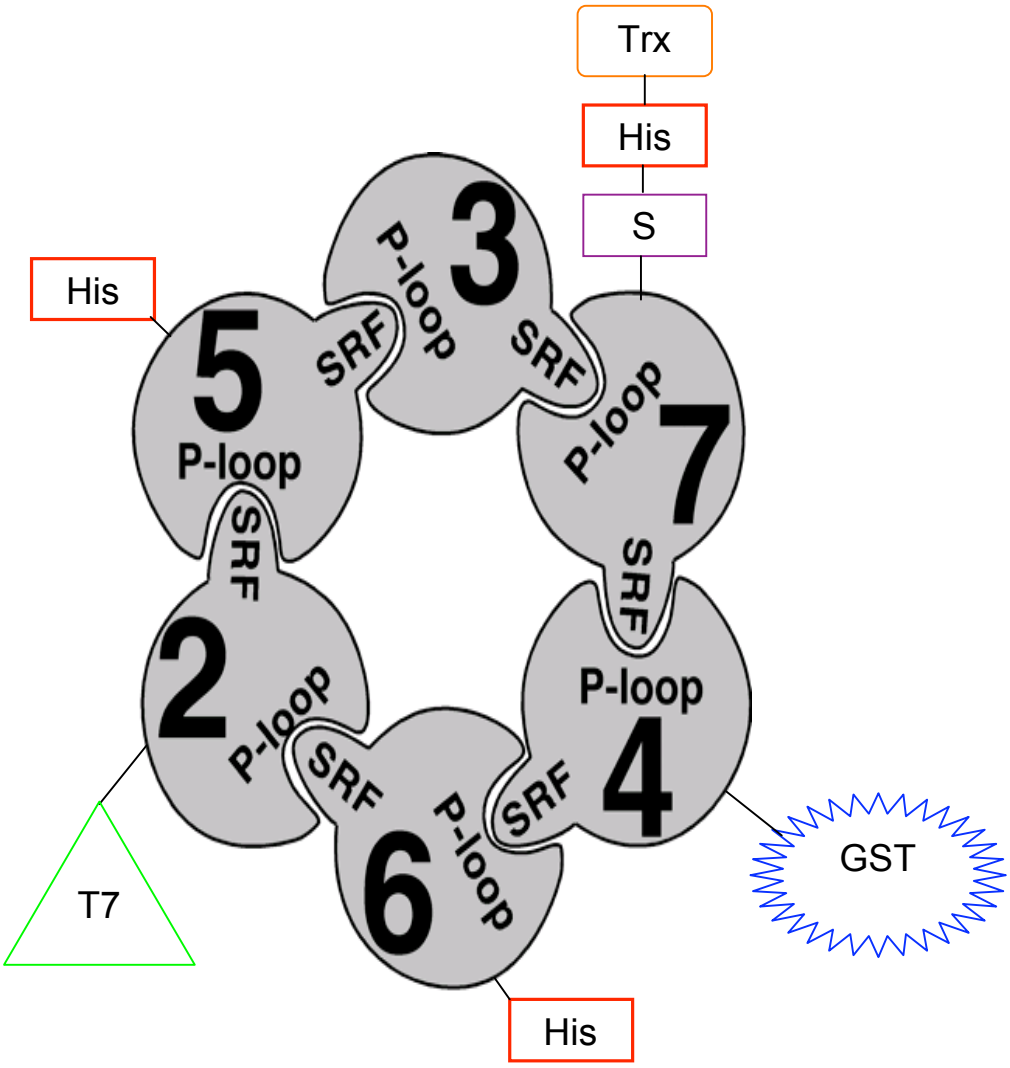


Figure 3.5 Diagram of the predicted interaction of the six recombinant Mcm proteins, in a heterohexameric complex with different affinity tags used for multistep protein purification (adapted from Davey et al., 2003).

Extra molecular weight was added due to tags and linkers on the proteins. A linker was placed between the proteins (hMcm2-7) and the tag. Linkers (or adaptors) were created from synthetic oligonucleotides, providing sequence overlaps with each of two sequences that were to be joined.

The number of residues and molecular weight of each hMcm2-7 was calculated with different tags on them and are given in table 3.2.

3.4.2 Co-expression

Three plasmids containing six Mcm proteins in pairs were used to co-transform *E.coli* BL21 (DE3) Rosetta cells (See section 2.3). As from the knowledge of individual protein purification trials, it was most difficult to purify Mcm4 compared to Mcm6 and Mcm7, therefore it was decided to follow the steps that were used to express Mcm4. Like Mcm4 was expressed in *E.coli* BL21 (DE3) Rosetta cells with low temperature, it was decided to follow the same approach. Rosetta cells containing all three vectors were used to inoculate 6 x 750 ml of LB containing appropriate antibiotics. These cultures were incubated at 37 ºC with 200 rpm shaking until an OD_{600} of 0.887 was reached and then chilled for 30 min. IPTG was then added to a final concentration of 0.1 mM. The cultures were then incubated at 12 ºC with 180 rpm shaking for a further 24 hours. The cells were harvested by centrifugation, and were stored at -80 ºC prior to protein purification.

3.4.3 Multi-step purification

A number of chromatographic steps were carried out to co-purify the co-expressed hMcms. The steps included in the purification scheme were ammonium sulphate precipitation, affinity chromatography, size exclusion chromatography and anion exchange chromatography (Figure 2.2).

The 18 grams of pellet (which was from 4.5 litre of culture) was purified first through ammonium sulphate purification. Ammonium sulphate is used

hMcm	Molecular mass of tagged protein (Da)	Length of protein (residues)	Tags	pI of protein
hMcm2	104,285	928	T7	**5.29**
hMcm3	91,680	815	-	**5.53**
hMcm4	123,946	1102	GST	**6.05**
hMcm5	83,995	748	His	**8.36**
hMcm6	98,360	835	His	**5.35**
hMcm7	94,5969	877	S, His, Trx	**5.86**
hMcm2/3/4/5/6/7 complex	1. 596,862	-	All the above	**-**

Table 3.2 Properties of the six recombinant human Mcm proteins after adding tags and linker molecular weight.

to purify proteins by altering their solubility, as its solubility is so high that salt solutions with high ionic strength are allowed (Chick and Martin, 1913). At low salt concentrations, the solubility of the protein increases with increasing salt concentration (salting in). As the salt concentration (ionic strength) is increased further, the solubility of the protein begins to decrease. At sufficiently high ionic strength, the proteins almost completely precipitate from the solution (salting out). After salting out, from 50% saturated ammonium sulphate solution, the pellet was stored at -80 ºC until further purification.

The first chromatographic step of the purification after ammonium sulphate purification was a nickel sepharose HP column with 500 mM imidazole step elution (See section 2.3.3). Elution was carried out using an AKTA 100 at 4 ^{o}C. The SDS-PAGE analysis showed that the complex eluted across four 0.5 ml fractions (2 ml total), which were pooled together and stored at 4 ^{o}C for the next purification step by size exclusion chromatography.

Size exclusion chromatography was carried out by using Superdex-200 16/60 column (120 ml bed volume). Affinity purified protein (2 ml) was loaded onto the column that was then connected to an ÄKTA FPLC system (4ºC). Proteins were eluted with gel filtration buffer and collected in a series of 2 ml fractions (Materials and methods section 2.3.4). The fractions containing high molecular mass proteins as indicated by the first peak on the A280 chromatogram (i.e. A3-A6) were pooled (4x 2 ml) and subjected to a final purification step.

For anion exchange step an automated ÄKTA 100 system was used. The sample was injected onto the Tricorn 5/50 column that was packed with Source 15Q resin (GE Healthcare) after diluting in anion exchange start buffer (Materials and methods section 2.3.5) in order to reduce the NaCl concentration to 50 mM so that the protein could be bound to the column and eluted from 50 to 500 mM NaCl gradient by changing the ionic strength. The proteins were eluted with a linear gradient of NaCl from 50 to 500 mM and collected in 1 ml fractions. The fractions from clarified lysate,

ammonium sulphate supernatant, resuspended ammonium sulphate pellet, nickel column pool, nickel column flow through, and size exclusion elution were analysed by the SDS-PAGE (Figure 3.6). The Coomassie blue stained gel showed several bands were present below the expected molecular weight (See table 3.2 for sizes of hMcm2-7) suggesting protein degradation during different steps or it may be contamination with other non-specific proteins. A few bands were present above the expected molecular weights in the ammonium sulphate precipitation and nickel column chromatography steps but were not present in size exclusion. There is a dark band in the nickel column fraction that could be degradation of proteins through nickel purification. A lot of protein is going in to nickel column flow through. The numbers of non-specific higher and lower bands other than expected molecular weight bands were significantly reduced after size exclusion chromatography (Figure 3.7). However, the hMcm2-7 as visualised on a Coomassie blue stained SDS-PAGE, individual subunits as well as the presence of some putative degradation products but it is clear that if all the hMcms were present in final chromatographic step, all six proteins were not in equal amount.

The presence of all six hMcm subunits needed to be confirmed at the end of the purification by western blotting using specific antibodies. Western blotting was carried out on fraction A13 of anion exchange chromatography elution to confirm the presence of each hMcm in final step of purification. Using antibodies specific to each of the hMcm western blot was performed. hMcms 5, 6 and 7 were detected at around the 100 KDa marker in anti-His-tag western blot using anti-tetra-His antibody (Qiagen). The western blot for anti-T7-tag was performed in presence of anti-T7-tag monoclonal antibody (Novagen) and detected the presence of hMcm2, which migrated just above the 100 KDa markers. Anti-S-tag western blot of anion exchange fraction A13 against anti-S-tag monoclonal antibody (Novagen) resulted in the detection of hMcm7 that migrated just above the 100 KDa markers. In anti-hMcm3 western blot of anion exchange fraction A13, membrane was probed with anti-hMcm3 monoclonal antibody (MBL) confirmed the presence of hMcm3 that migrated to just above the 75 KDa markers. This antibody also

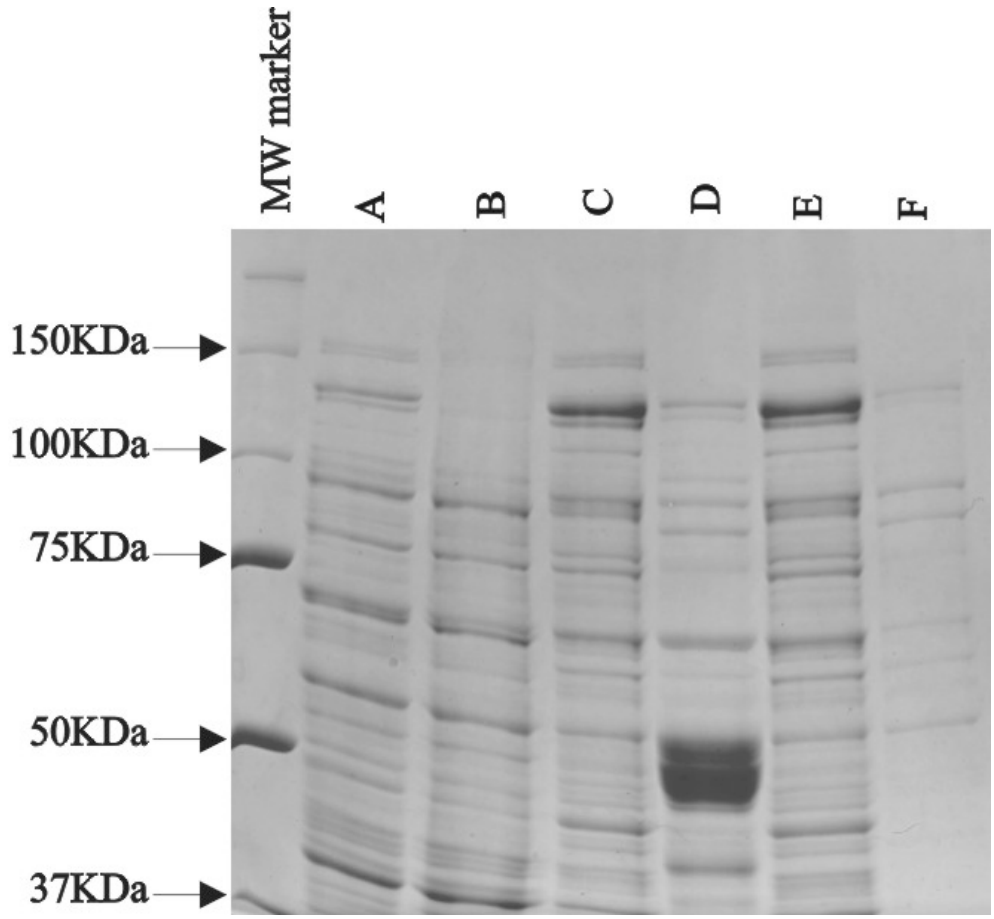


Figure 3.6 Co-purification of hMcm2-7: 10% SDS-PAGE analyses of fractions from multi-step purification of co-expressed proteins. A Coomassie blue stained gel showed several bands were present above and below the expected molecular weight, along with expected sizes of hMcm2-7 proteins that could due to contamination or protein degradation. The 20 µl of samples from various stages of purification were loaded as follows: A) clarified lysate, B) ammonium sulphate supernatant, C) resuspended ammonium sulphate pellet, D) Nickel column fractions pool, E) Nickel column flow through, F) Size exclusion column fractions pool. A thick band near 50 KDa marker in nickel column fraction seems to indicate the presence of some putative degradation products.

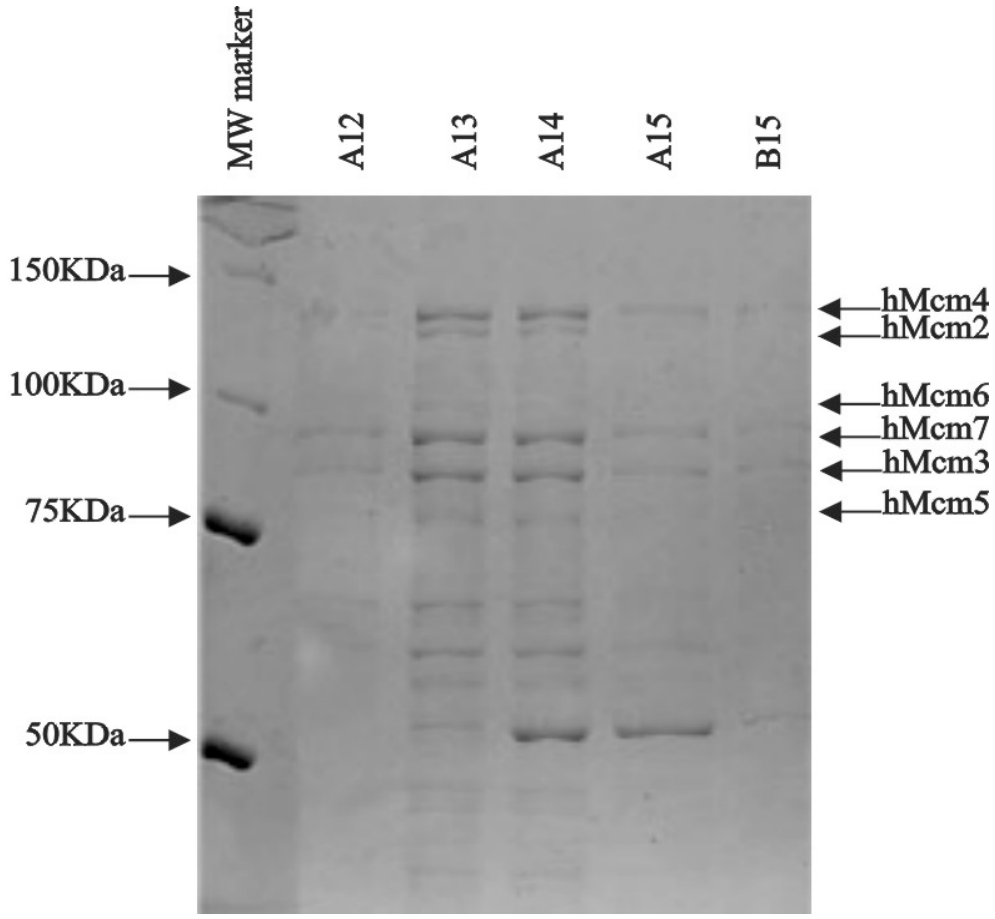


Figure 3.7 Anion exchange chromatography for purification of hMcm2-7: SDS-PAGE analysis of fractions from the anion exchange column stage. A Coomassie blue stained gel showed six individual bands were present at the expected sizes for hMcm2-7 in several fractions from anion exchange chromatography. A few low molecular weight bands were also visible which could be the degradation of hMcm2-7 proteins due to multi-step purification. Arrowheads indicate expected bands of individual hMcm2-7.

cross reacted with some lower molecular mass bands that might be the result of proteolysis products. Anti-hMcm4 western blot was carried out in the presence of anti-hMcm4 (Santa Cruz Biotechnology) and gel showed the expected band just above the 100 KDa markers with some other non-specific bands which could be protein degradation or some contamination or even non specific binding to antibody. The presence of hMcm6 protein in elution fraction was detected by using anti-hMcm6 (BD Biosciences) in anti-hMcm6 western blot and showed that the hMcm6 migrated between the 100 and 150 KDa markers (Figure 3.8).

SDS PAGE analysis and western blots indicated the presence of all six hMcms, after various co-purification steps. The intensity of different bands of each hMcm varied in the elution fractions of anion exchange chromatography suggesting that all six hMcms were not present in equal amounts. Several bands were present below the expected molecular weight, suggesting protein degradation during different steps. An alternative explanation for the bands could be contamination from other proteins.

To date the exact stoichiometry of hMcm complex has not been accurately measured. If the approach of purifying individual Mcms would have been successful for all of the Mcms, it might have been possible to form homogeneous complexes. By using analytical ultra centrifugation, the molecular weight of individual complexes could have been measured and stoichiometry of hMcm complex would have been measured by quantitative mass spectrometry. Quantification of the hMcm complex was possible by using MALDI-MS fingerprinting by running six analyses of each homohexamer complex. The abundance of the six proteins in a co-purified fraction varied significantly and in the presence of low molecular weight contaminating proteins or protein degradation in the final purification step made it difficult to use this complex to estimate stochiometry.

Similarly the use of different tags on different Mcms had given the choice to co-purify these Mcms with different tags using pull down assays and it could have suggested that Mcms could make heterohexamer complex. As the

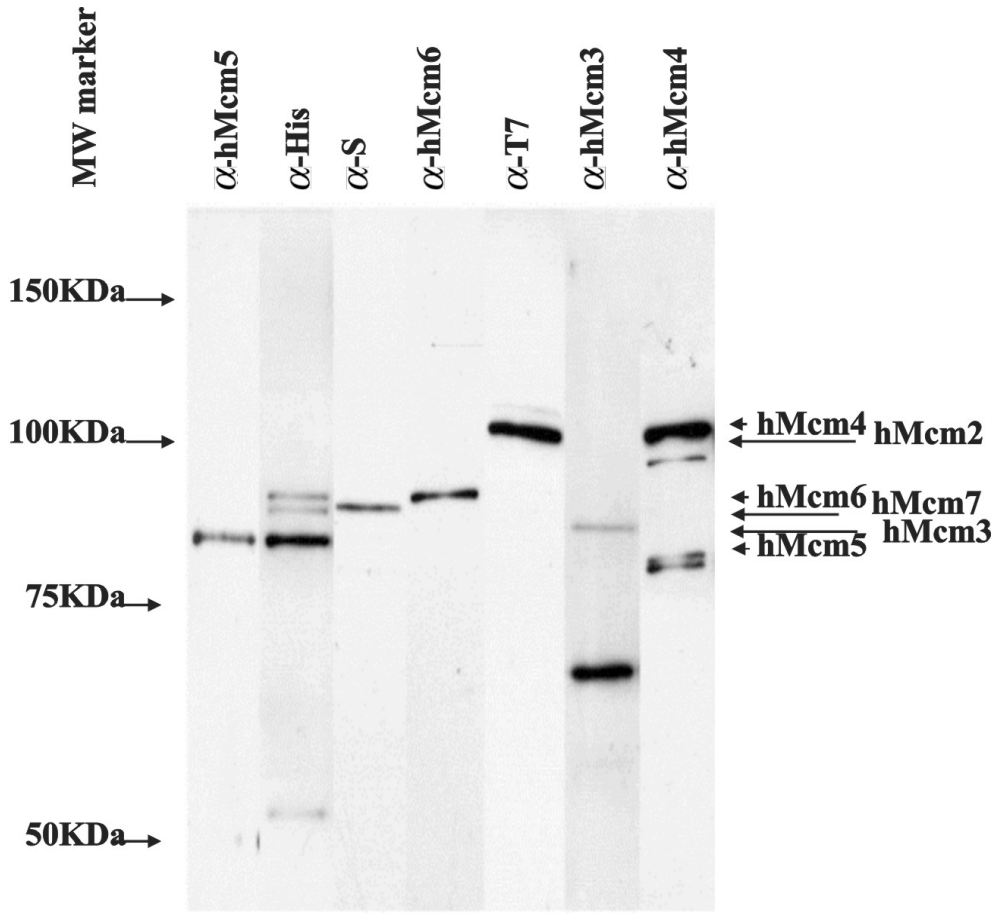


Figure 3.8 Western blot of co-purified hMcm2-7: Western blot analysis of elution fraction A 13 from anion exchange chromatography of co-expressed hMcms. Using antibodies against each of the hMcm2-7 showed that all six hMcms were present in elution fraction of anion exchange chromatography. Anti-T7 for hMcm2, anti-hMcm3 monoclonal antibody for hMcm3, anti-hMcm4 for hMcm4, anti-His-tag for hMcm5, anti-hMcm6 for hMcm6 and anti-S-tag monoclonal antibody for hMcm7 were used. Arrowheads indicate positions of individual hMcm2-7. This blot was provided by Dr. Richard Parker.

Mcms have similar biochemical properties, it is also possible that because of similar properties of these six proteins, homohexamers of each protein are co-purifying or making a combination of homohexamers and heterohexamers.

The final concentration of this complex was measured by Bradford assays and the concentration of this complex was found to be 0.28 mg/ml. As the amount of purified protein was very low and a further protein concentration step might result in loss of protein, therefore it was decided that this protein would not be concentrated. These purified proteins were further characterised using helicase assays and kinase assays.

3.5 Helicase assays for human Mcms

The main purpose of purifying hMcms was to carry out helicase assays to determine whether hMcm 2-7 has the ability to separate two annealed oligonucleotides efficiently *in vitro.* A strand displacement or helicase assay was carried out using a forked substrate that was generated by annealing a radiolabelled 57 bases $\gamma[^{32}P]$ATP oligonucleotide and non-radiolabelled 74 bases oligonucleotide (See appendix 2 for sequences) resulting in a 25 bp double stranded complementary region. The helicase assay was performed in the presence of helicase buffer (See section 2.7.2). Increasing protein concentrations (50 nM to 500 nM) were used while the substrate concentration remained 1 nM for each reaction. Reactions were carried out for 1 hour at 37 ^{o}C and then stopped by the addition of helicase stop buffer that contained EDTA to inhibit ATP-dependent enzyme activity. Four negative controls were generated by incubating reaction on ice, adding EDTA at the start of reaction, omitting protein from the reaction and just running substrate without buffer, along with one positive control that was prepared by boiling the substrate at 100 ^{o}C for 5 minutes in the helicase buffer with subsequent addition of helicase stop buffer. Reactions were electrophoresed on a pre-run 12% (polyacrylamide) TBE gel to visualize the separated labelled oligonucleotide. After drying, the gel was exposed to a phosphorimager screen.

Protein dependent helicase activity was shown by the hMcm complex. Unwinding started significantly at protein concentration of 5 nM and peaked at 200 nM (Figure 3.9A and B). After 200 nM, there were very faint signals just below the expected bands, suggesting nuclease activity due to contamination.

The helicase activity was measured using densitometry on Quantity One software (BioRad) and a graph was plotted between DNA unwound and protein concentrations (Figure 3.9B) where the boiled substrate was taken as to displace DNA 100% in graphic values.

3.6 Assays for the activity of kinases on human Mcms

Kinase activity assays were carried out to test the phosphorylation of purified hMcm2-7. Cyclin dependent kinases (CDK) phosphorylate Mcms on serine and threonine amino acids with the help of cyclin (Masai et al., 2000). Without cyclin, CDK has a little or no kinase activity, so the active form is a CDK/cyclin. Two complexes, CDK 2/cyclin A and CDK 2/cyclin E (provided by Dr Cyril Sanders, The University of Sheffield) were used separately to test their activity on human Mcms *in vitro*. Assays were carried out in 20 µl of kinase buffer (See section 2.8) with subsequent addition of $\gamma[^{32}P]$ATP and kinase complexes. Kinase activity was measured by the $\gamma[^{32}P]$ATP incorporated into hMcm complex over the period of 30 minutes at 37 ^{o}C. Reactions were stopped by the addition of stop buffer and the samples were analysed on 4-12% polyacrylamide gradient gel. Gels were then exposed to film or phosphorimager screen overnight to detect phosphorylation.

WT human Mcm2-7 complex was strongly phosphorylated by both complexes, CDK 2/cyclin A and CDK 2/cyclin E. Mutant hMcm2-7 complex (each hMcm was truncated from C terminal) was provided by Dr Richard Parker, a post-doctoral researcher in the Chong group to test if the mutant hMcm2-7 complex could also act as a substrate for phophorylation.

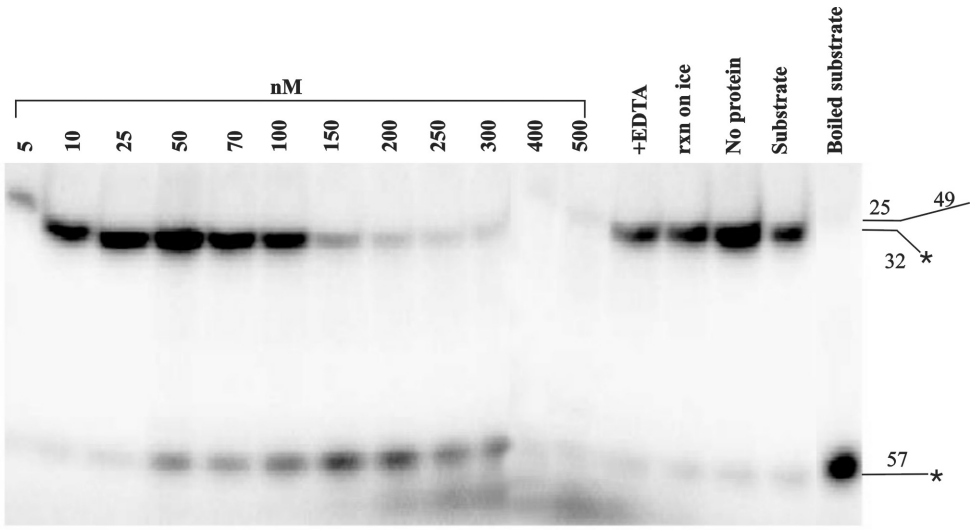


Figure 3.9 (A) Helicase activity of human Mcm2-7: PAGE analysis of strand displacement assays of hMcm complex with various protein concentrations. This activity was measured by the amount of displaced single stranded DNA from with double stranded substrate over the period of 1 hour at 37^{o} C. Each helicase assay was performed at least three times. The possible explanation of faint signals just below the expected bands could be contamination due to nuclease activity.

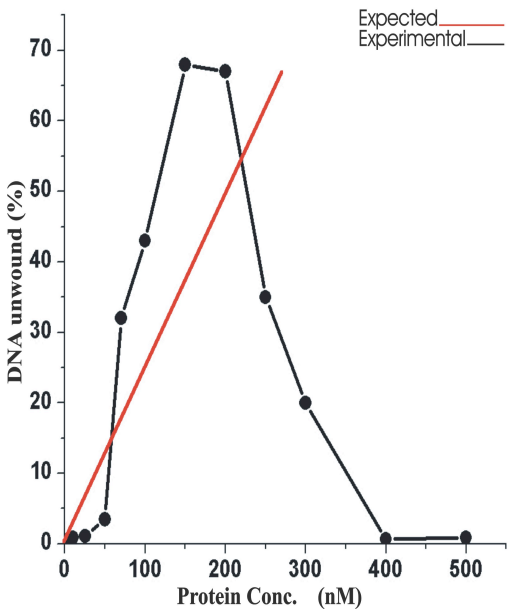


Figure 3.9 (B) Quantification of helicase activity of human Mcm2-7: Densitometry was used for quantification of strand displacement activity compared with boiled substrate and indicated as a percentage and plotted on a graph. Helicase activity increased with increase in protein concentration to 200 nM and then gradually decreased.

Deletion of a C-terminal domain from the *M. thermautotrophicus* Mcm protein resulted in more processive helicase activity (Jenkinson and Chong, 2006), therefore each human Mcm subunit was truncated from C-terminal and purified over range of chromatographic steps as explained for full length wild type human Mcm complex. CDK 2/cyclin E showed more robust activity compared to CDK 2/cyclin A. Wild type human Mcm2-7 complex showed stronger signals of phosphorylation (Figure 3.10 A and B). There was another strong band below the expected band size, above 36 KDa marker in CDK 2/cyclin E gel (Figure 3.10 B), consistent with the autophosphorylation activity of cyclin E (47 KDa) as reported previously (Won and Reed, 1996).

Mass spectrometry is a useful technique to identify the phosphorylated residues. Phosphorylation sites of human Mcm complex could be quantified by using mass spectrometry but due to PhD time frame as a limiting factor, samples were not sent for mass spectrometry. This could be done in future to identify known and novel phosphorylated sites on Mcms.

3.7 Helicase activity of phosphorylated human Mcm2-7 complex

hMcm4,6,7 complex helicase activity was shown to be inhibited when Mcm4 was phosphorylated by CDK 2/cyclin A (Ishimi et al., 2000). To test if phosphorylated human Mcm2-7 complex still had helicase activity, strand displacement assays were carried out on phosphorylated human Mcm2-7 complex as explained in section 2.7.2. Reactions were analysed on a 12%TBE gel that was exposed to a phosphorimaging screen overnight.

As the phosphorylation activity was better with CDK 2/cyclin E, thus this complex was used for further assays of kinase activity to see the effect of phosphorylation of the hMcm complex on its helicase activity. As the protein concentration of 50 nM to 200 nM showed robust helicase activity so the increasing amount of phosphorylated hMcm2-7 protein concentration (50 nM to 200 nM) was used to perform the helicase assays on phosphorylated hMcm2-7 complex. Boiled substrate and two different concentrations of

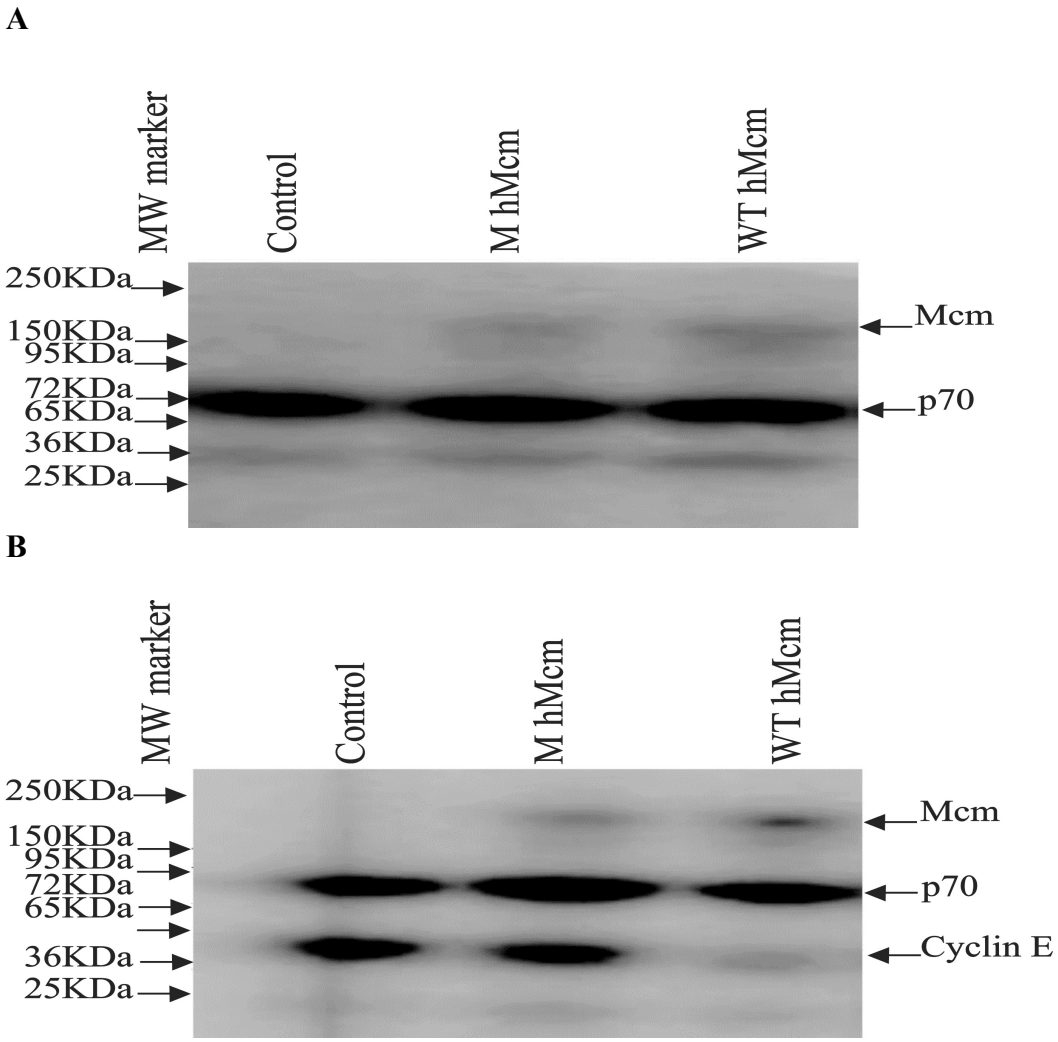


Figure 3.10 Phosphorylation of human Mcm2-7 WT and mutant complex: (A) with CDK 2/cyclin A (B) with CDK 2/cyclin E. Protein samples (human Mcm complex wild type and human Mcm complex mutant) were incubated at RT for 30 minutes in a kinase buffer in presence of CDK 2/cyclin A or CDK 2/cyclin E before running on a 4 to 12% SDS gradient polyacrylamide gel. The gels were exposed overnight to X-ray film to detect phosphorylation. Control reactions were performed without adding any human Mcm protein complex. Arrowhead near 150 KDa protein markers showing the phosphorylation on hMcm complexes, while another arrowhead near 36 KDa protein markers on figure B showed the autophophorylation of cyclin E. Another arrow indicates the presence of p70, which is common contamination of Cdk protein purification.

dephophorylated hMcm2-7 complex generated three positive controls. In the control reaction 1 and 2, dephophorylated hMcm2-7 protein showed good helicase activity while the phosphorylated hMcm2-7 did not show any helicase activity (Figure 3.11). These results showed that when the complex was phosphorylated by CDK 2/cyclin E, its helicase activity was significantly reduced. A super-shift was observed in all phosphorylated hMcm2-7, which is consistent with published (Jenkinson and Chong, 2006), and unpublished data where a super-shift was found in negative controls or reactions where no helicase activity was observed. This result confirms that CDK 2/cyclin E negatively regulates the helicase activity of human Mcm complex.

3.7 Discussion

Despite trying various protocols, buffers and vectors, different conditions, individual human Mcms remained largely insoluble. hMcm4, 6 and 7 were purified eventually with low solubility. hMcm6 was purified using the pQE vector under T5 promotor to increase the expression levels. hMcm4 solubility increased when induced at the lower temperature of 12 ^{o}C after addition of IPTG. To get more protein, hMcm4 was purified from 4.5 L of culture. Even after all these efforts, not enough proteins were purified to carry out further assays on them. It was then decided to try to co-purify these proteins, which could help the solubility of these proteins by interacting with each other during the expression and purification steps. All six proteins were co-purified successfully over multiple purification steps, although there were various low molecular weight bands visible in the gel, suggesting protein degradation in these steps. Protein co-purification over several chromatographic steps suggests protein/protein interactions of hMcm *in vitro.* Helicase activity was tested on co-purified hMcms successfully. hMcm showed robust protein-dependent helicase activity. This confirms that hMcm possesses a reproducible DNA helicase activity. Furthermore the hMcm complexes were tested as a target for kinase activity in the presence of CDK 2/Cyclin A and CDK 2/Cyclin E complex. CDK 2/Cyclin A and CDK2/Cyclin E phosphorylated the hMcm complex. As the phosphorylation

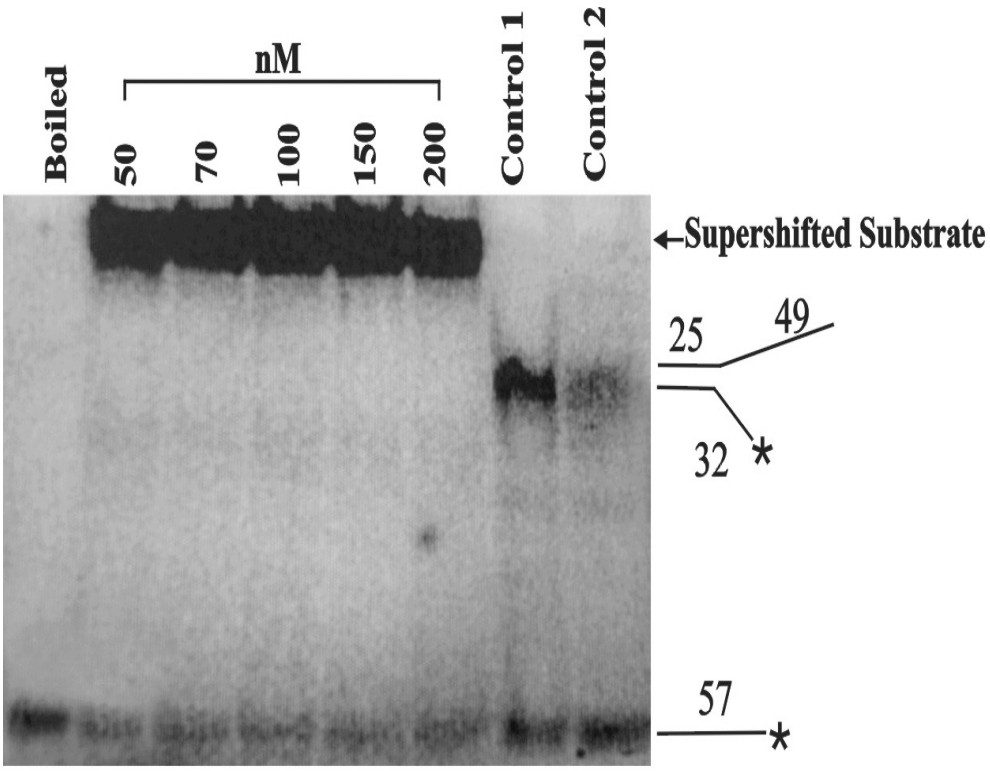


Figure 3.11 Helicase activity of phophorylated human Mcm2-7: PAGE analysis of strand displacement assays of phosphorylated hMcm complex with various protein concentrations. Control 1 and control 2 were generated by performing the helicase assays on two different concentrations of dephosphorylated hMcm2-7 complex (50 nM and 70 nM). Phosphorylated hMcm2-7 complex did not show any helicase activity and a supershift of the substrate was observed with all concentrations of phosphorylated hMcm2-7 complex.

by CDK 2/cyclin E complex was better then CDK 2/Cyclin A therefore phosphorylated hMcm complex with CDK 2/cyclin E was further tested for helicase activity. Phosphorylated hMcm2-7 complex did not show any helicase activity and a super-shift was observed in helicase assays of phosphorylated hMcm2-7 complex. These results suggest that phosphorylation of hMcm2-4 complex by CDK 2/cyclin E down-regulates hMcm helicase activity.

Chapter 4: Biochemical studies of *M. maripaludis* Mcms

4.1 Introduction

Replication in archaea appears to be a simplified version of the eukaryotic process as fewer polypeptides participate in each phase of replication (Edgell and Doolittle, 1997). Thus, archaea represent the simplest model to study the molecular mechanisms of eukaryotic DNA replication (Tye, 2000). Extensive research has been carried out on archaeal Mcms. In all eukaryotes six Mcm helicase proteins are found but this is not true for archaea. The genome of *Methanococcus maripaludis* S2 has been sequenced and revealed that this archaeon contains four Mcm homologues (Hendrickson et al., 2004) while *Methanothermobacter thermautotrophicus* has just a single Mcm that forms a double hexamer complex and has an ATP-independent DNA-binding activity which can distinguish between single stranded and double stranded DNA and 3'-5' helicase activity that requires of ATP hydrolysis (Kelman et al., 1999; Chong et al., 2000; Shechter et al., 2000). Archaeal Mcms provide the best examples of biochemical activity in Mcms todate.

Purified recombinant MMP (*Methanococcus maripaludis*) Mcms both individually and in combination were provided by Dr Alison Walters in the Chong group (Walters and Chong, 2010). As the concentration of MMP McmD and MMP Mcm complex were very low and poor quality, trials were conducted to co-purify proteins from the two constructs pAW 30 and pAW 31 provided by Dr Alison Walters (Walters and Chong, 2010). Furthermore *M. maripaludis* individual MMP Mcms designated (MMP McmA, MMP McmB, MMP McmC and MMP McmD) and complex were tested for helicase activity. A number of different experiments were performed to determine the optimum conditions for helicase activity of these proteins.

4.2 Co-purification of *M. maripaludis* Mcms

M. maripaludis Mcms were co-purified from two constructs. One construct contained MMP McmA and T7 tag adaptor upstream of MMP McmC into

pCDF-2 designated as pAW 31 while the other construct contained MMP McmB and GST tag adaptor on MMP McmD in pET32 EK/LIC designated as pAW 32 (Figure 2.3).

As the hMcm4 seemed to be the most difficult protein to purify amongst the human purified proteins, therefore it was decided to use the similar approach of hMCM4 to purify these MMP Mcms, that is to use Rosetta *E. coli* BL21 (DE3) cells at low temperature after induction but with different sets of buffers (Materials and methods section 2.5). Two pellets from two different expression constructs were harvested by centrifugation at 4000 rpm for 15 minutes at 4 ^{o}C.

4.2.1 Purification of his-tagged protein

pAW 31 contained a his tag on MMP McmA and T7 tag on MMP McmC. It was decided to purify this protein using the his-tag.

Harvested pellets were resuspended in resuspension buffer (See section 2.5.1.1 in Materials and Methods) before storing at -80 ^{o}C until further use. Proteins were purified from 750 ml of culture using Talon beads. Elution fractions were separated by 10% SDS-PAGE (Figure 4.1). Coomassie stained gels showed a band of protein just above the 75 KDa marker. As the molecular weights of MMP McmB and MMP McmC were 77 KDa and 78 KDa respectively, this band could be the result of one or two proteins being purified over Talon beads. To confirm whether this thick band was actually from two different proteins, western blots were carried out on three elution fractions. Western blot using anti-his for MMP McmA and anti-T7 for MMP McmC antibodies indicated the presence of both proteins in all three elution fractions (Figure 4.3 A and C). The large number of bands in the anti-his western blot of elution fraction 1 could be contamination or protein degradation.

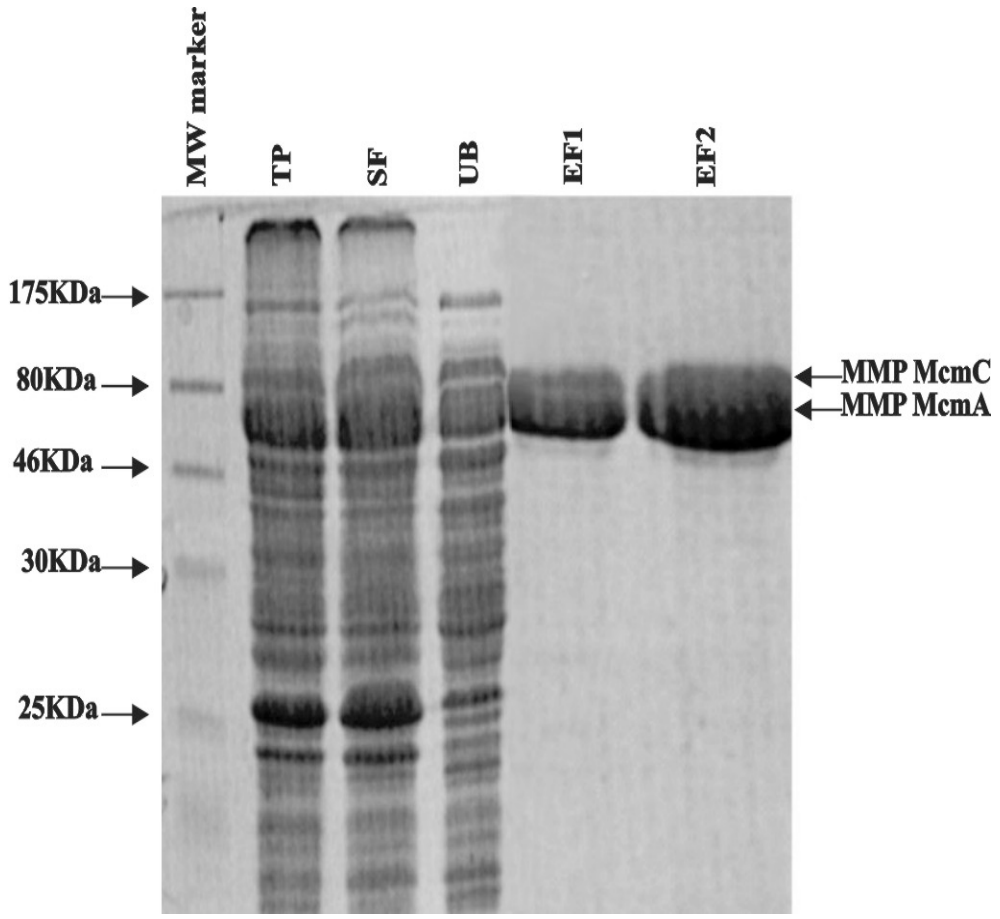


Figure 4.1 Purification of co-expressed MMP McmA and MMP McmC: SDS-PAGE analyses of fractions from purification trials. Proteins were expressed in *E.coli* BL21 (DE3) Rosetta cells at 12 ^{o}C overnight. 750 ml culture pellet was resuspended and eluted from nickel beads using 500 mM imidazole. A band of protein is visible at the expected sizes of the proteins, this could be due to a single protein as the sizes of the two proteins are very close to each other 77 KDa (MMP McmA) and 78 KDa (MMP McmC). Western blots were carried out to confirm the presence of two proteins in elution fractions (See figure 4.3).

TP, total protein; SF, soluble fraction; UB, unbound fraction; EF, elution fraction

4.2.2 Purification of GST-tagged protein

pAW 32 expression construct contained S-tag on MMP McmB and GST tag on MMP McmD. It was decided to purify this protein using GST-tag.

Harvested pellets were resuspended in resuspension buffer (See section 2.5.1.2) before storing at -80 ^{o}C until further use. Proteins were purified from 750 ml of culture using glutathione sepharose beads. Elution fractions were separated by 10% SDS-PAGE (Figure 4.2). Coomassie stained gel showed a protein band just below 100 KDa protein marker which was expected for MMP McmB band of 93 KDa and a faint band just over 100 KDa protein marker which could be MMP McmD (108 KDa). There were non-specific bands below the expected bands; therefore, one further step of anion exchange was carried out as explained in section 2.3.5. After anion exchange chromatography step, MMP McmD, which was largely insoluble, was not present (Figure 4.2).

Western blot (See section 2.6) using anti-S for MMP McmB and anti-GST for MMP McmD antibodies indicated that two proteins were being purified in elution fractions 1, 2 and 3 (Figure 4.3 B and D).

4.2.3 Quantification of purified proteins

Purified proteins were quantified using Bradford assays. The concentration of proteins from the pAW31 construct was 5 mg/ml in a volume of 18 ml. The concentration of proteins from pAW32 construct was 2 mg/ml and the volume was 10 ml.

Although the concentrations of MMP Mcm proteins were greatly enhanced by this method of purification, the MMP McmD was mainly in the insoluble fraction. It was then decided to use the MMP Mcm complex (MMP McmA, MMP McmB, MMP McmC and MMP McmD) provided by Dr Alison Walters for further helicase assays although the concentration of provided MMP Mcm complex was 0.256 mg/ml.

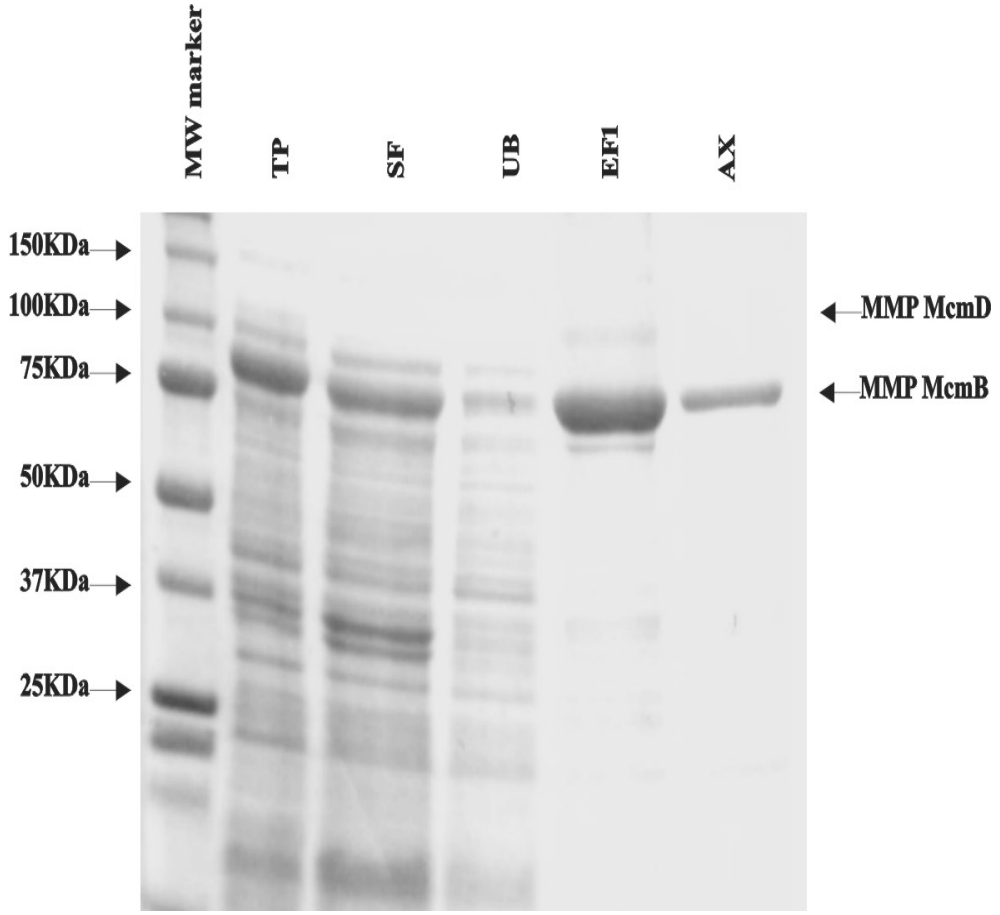


Figure 4.2 Purification of co-expressed MMP McmB and MMP McmD: SDS-PAGE analyses of fractions from muti-step purification trials. Proteins were expressed in *E.coli* BL21 (DE3) Rosetta cells at 37 ^{o}C overnight. Cells from 500 ml culture were lysed and protein from soluble fraction was purified using 50 mM glutathione and further purified using anion exchange chromatography. A protein band of 93 KDa for MMP McmB was visible in EF1 (Nickel chromatography) and EF2 (anion exchange chromatography). A faint band of 108 KDa (MMP McmD) was present in EF1 but not in further purification steps.

TP, total protein; SF, soluble fraction; UB, unbound fraction; EF1, Nickel chromatography; AX, elution fraction from anion exchange chromatography

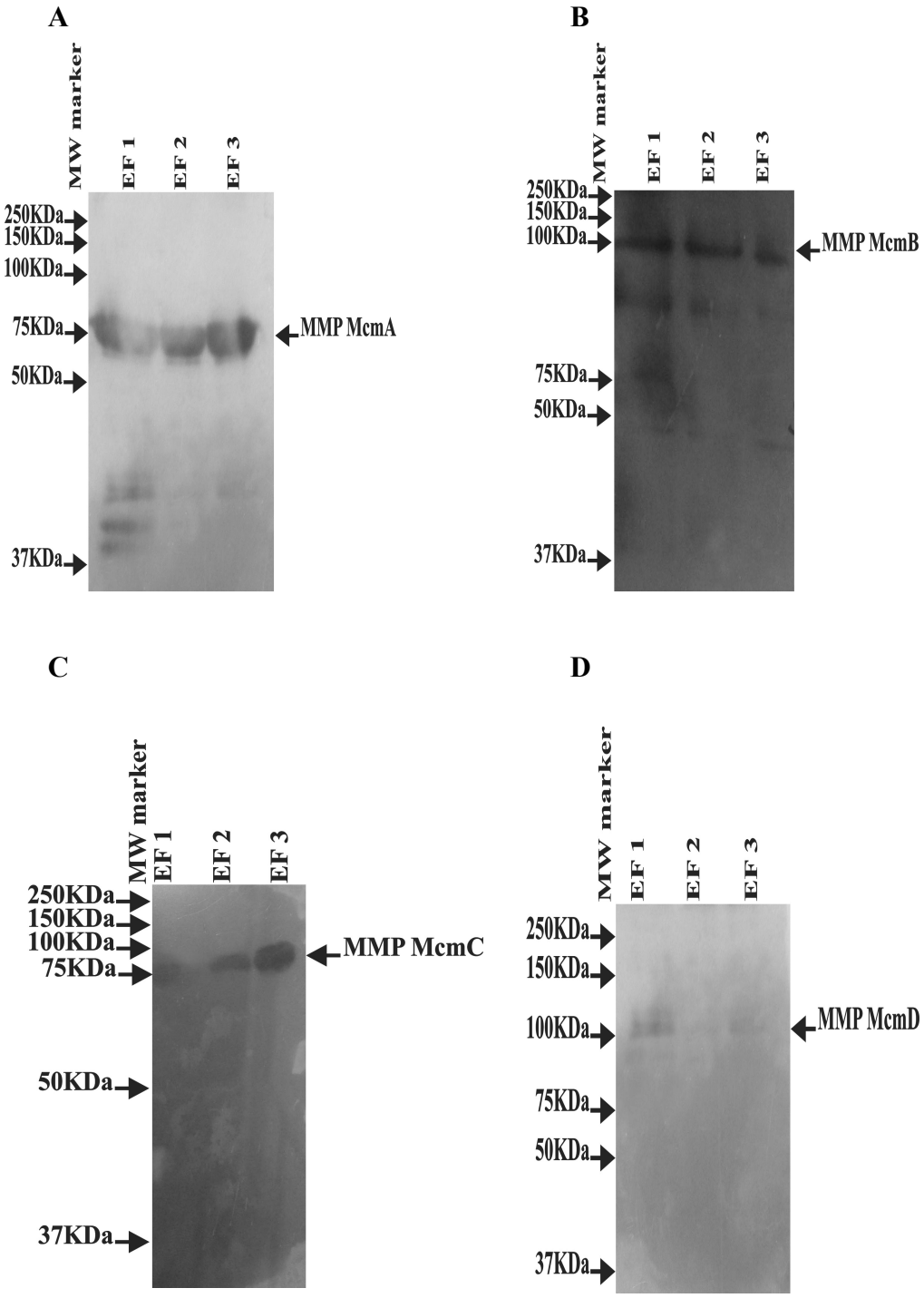


Figure 4.3 Western blot analysis of elution fractions of co-purified protein: Western blot analysis of differently tagged MMP Mcms showed that all the proteins were co-purified using (A) Anti-his tag for MMP McmA, (B) Anti-S for MMP McmB, (C) Anti-T7 for MMP McmC, (D) Anti-GST for MMP McmD. Although the signals from western blot confirmed the presence of each of Mcm but the use of positive and negative control would have provided further evidence to impact on results.

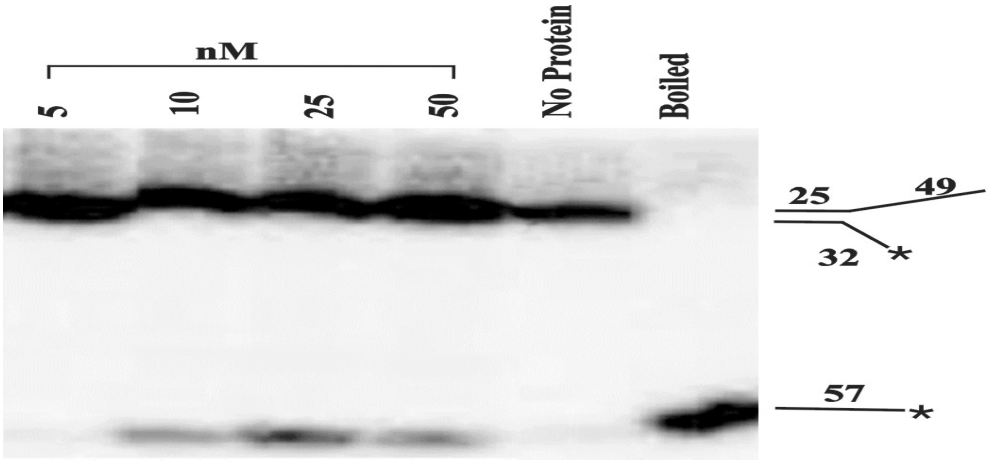
4.3 Helicase assays for *M. maripaludis* Mcms

Protein sequence analysis showed that all four Mcms in *M. maripaludis* might be functional, as they all possess highly conserved motifs in their sequences (Walters and Chong 2010), which are helix-2 insert, pre-sensor 1 β-hairpin and arginine finger. On that basis helicase activity assays were performed for all four Mcms individually. Strand displacement assays were carried out using a forked substrate that was generated by 57 bp ^{32}P-labelled oligonucleotide and non-labelled 74 bp oligonucleotide that annealed to produce a 25 bp double-stranded complementary region. The helicase assay was performed in the presence of helicase buffer (Section 2.7.2). Increasing protein concentrations (5 nM to 50 nM) were tested while the substrate concentration remained 1 nM for each reaction. Reactions were carried out for 1 hour at 37 ^{o}C and then stopped by the addition of helicase stop buffer (Section 2.7.2) that contained EDTA to inhibit unwinding. A negative control was generated by performing the reaction without adding any protein along with one positive control that was prepared by boiling the substrate at 100 ^{o}C for 5 minutes in helicase buffer with subsequent addition of helicase stop buffer. Reactions were electrophoresed on a pre-run 12% TBE gel to visualize the separated labelled oligonucleotide. After drying, the gel was exposed to a phosphoimager screen overnight and imaged using a phosphoimager (BioRad). Each experiment was performed at least three times with representative gels being shown.

MMP McmA and B showed protein dependent helicase activity (Figure 4.4 and 4.5) while MMP McmC did not display any activity or if there was any, it was a very low level of activity (Figure 4.6). MMP McmD was also tested for helicase activity and it did not show any helicase activity. MMP Mcm complex also showed robust helicase activity (Figure 4.7). However, the stoichiometry of the MMP Mcm complex was not determined, and it was possible that several heteromeric complexes composed of different Mcm subunits could be formed (Walters and Chong, 2010). On that basis it might be possible that in MMP Mcm complex, only MMP McmA and B were playing a major role in helicase activity. The results from McmD helicase

MMP McmA

A



B

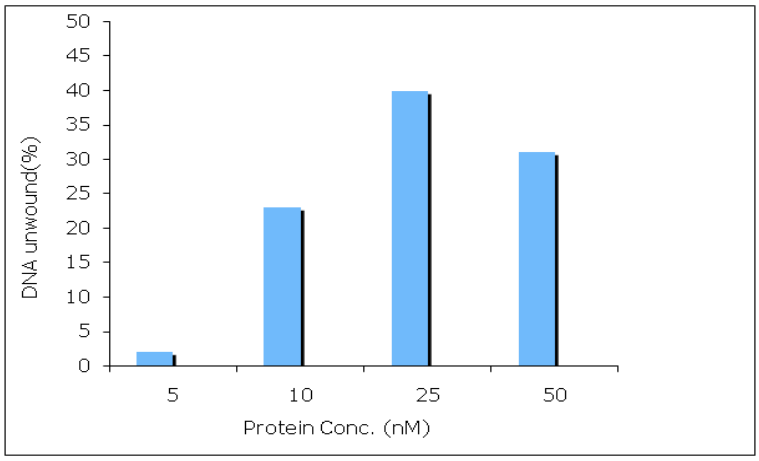
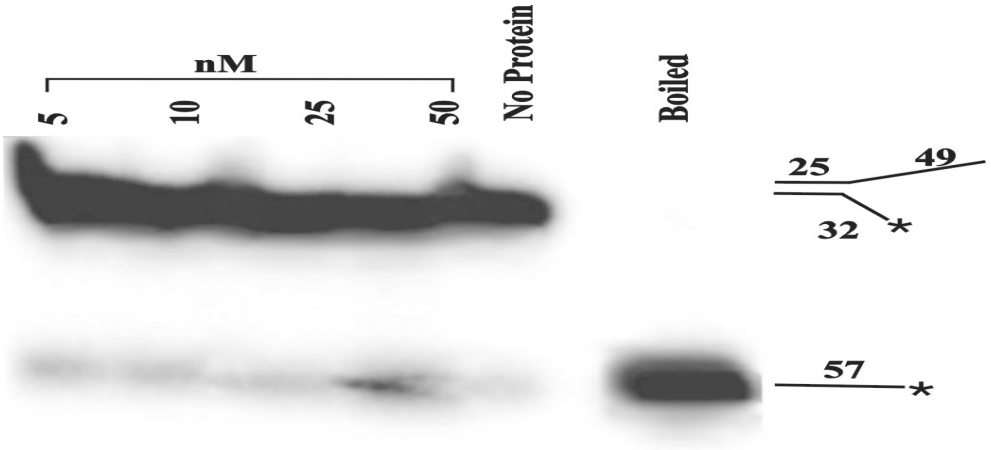


Figure 4.4 Helicase activity of MMP McmA: Increasing amounts of purified MMP McmA protein was added to ^{32}P-labelled DNA substrate with a 25 bp double-stranded region and incubated at 37^{o}C for 1 hour. Reactions were analysed by separating the dsDNA substrate and unwound single stranded ^{32}P-labelled oligonucleotide using PAGE and exposing the gel to a phosphorimaging screen overnight. MMP McmA (A) showed robust protein dependent helicase activity. (B) The helicase activity was quantified using densitometry. Boiled substrate was considered to displace DNA 100% in graphic values.

No protein was used as a negative control while boiled substrate was a positive control.

MMP McmB

A



B

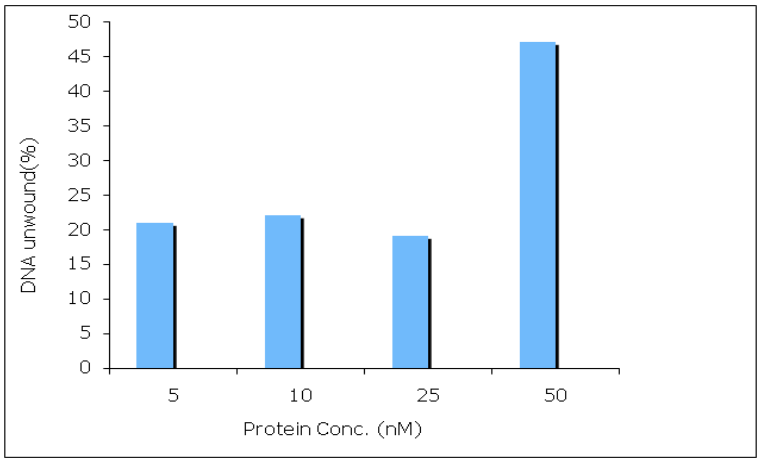
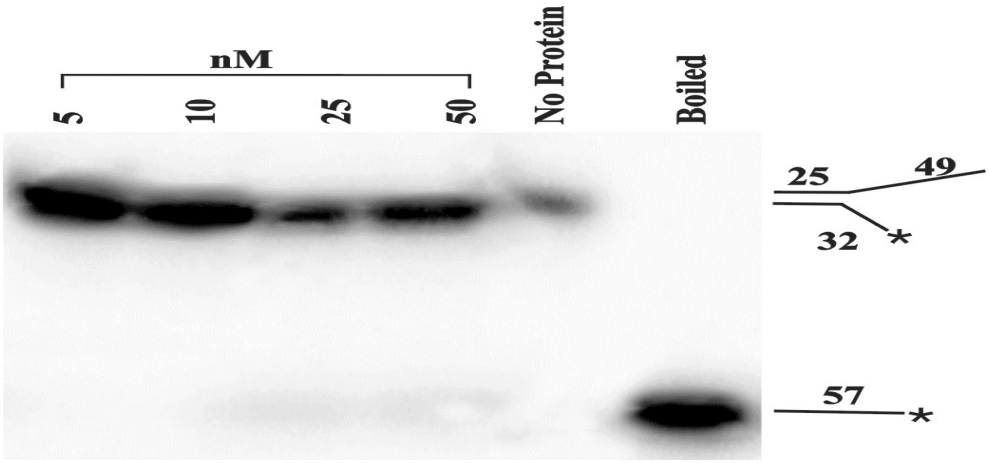


Figure 4.5 Helicase activity of MMP McmB: Increasing amounts of purified MMP McmB protein was added to ^{32}P-labelled DNA substrate with a 25 bp double-stranded region and incubated at 37^{o}C for 1 hour. Reactions were analysed by separating the dsDNA substrate and unwound single stranded ^{32}P-labelled oligonucleotide using PAGE and exposing the gel to a phosphorimaging screen overnight. MMP McmB (A) showed robust protein dependent helicase activity. (B) The helicase activity was quantified using densitometry. Boiled substrate was considered to displace DNA 100% in graphic values.

No protein was used as a negative control while boiled substrate was a positive control.

MMP McmC

A



B

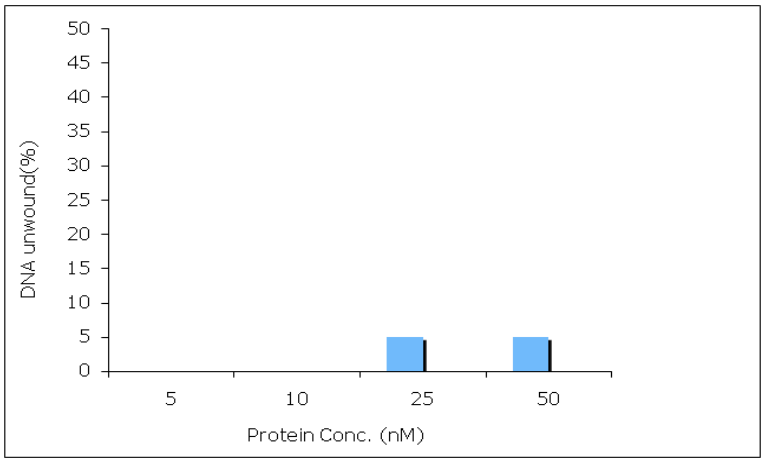
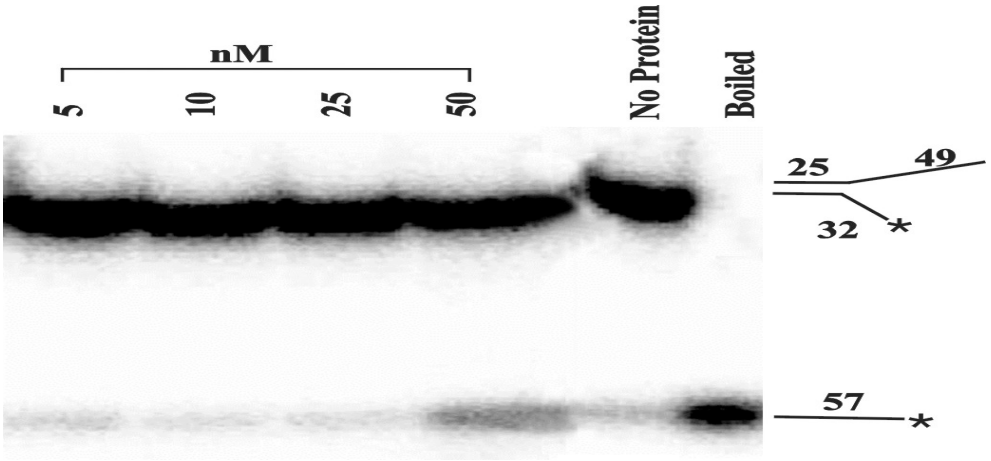


Figure 4.6 Helicase activity of MMP McmC: Increasing amount of MMP McmC purified protein was added to a ^{32}P-labelled DNA substrate with a 25 bp double-stranded region and incubated at 37^{o}C for 1 hour. Reactions were analysed by separating the dsDNA substrate and unwound single stranded a ^{32}P-labelled oligonucleotides using PAGE and exposing the gel to the phosphorimaging screen overnight. MMP McmC (A) did not show any helicase activity or there may be very level of activity. (B) The helicase activity was quantified using densitometry. Boiled substrate was considered to displace DNA 100% in graphic values.

The negative control was a reaction that did not contain protein while boiled substrate was a positive control.

MMP Mcm complex

A



B

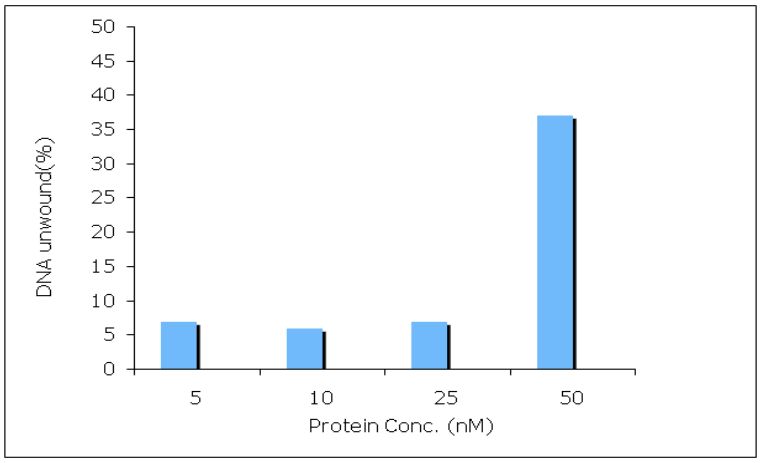


Figure 4.7 Helicase activity of MMP Mcm complex: Increasing amount of purified Mcm proteins were added to a ^{32}P-labelled DNA substrate with a 25 bp double-stranded region and incubated at 37^{o}C for 1 hour. Reactions were analysed by separating the dsDNA substrate and unwound single stranded a ^{32}P-labelled oligonucleotides using PAGE and exposing the gel to the phosphorimaging screen overnight. MMP Mcm complex (A) showed protein dependent helicase activity. (B) The helicase activity was quantified using densitometry. Boiled substrate was considered to displace DNA 100% in graphic values.

The negative control was a reaction that did not contain protein while boiled substrate was a positive control.

could not be interpreted very well, as the concentration of MMP McmD was very low and it was also contaminated with a protein, which was probably the Cpn60 chaperone. As MMP McmD was expressed and purified from Arctic Express (DE3) RIL cells and these cells co-express the cold-adapted chaperonin Cpn60 at low temperature (Lund, 1994). These chaperonins confer improved protein processing at lower temperatures, potentially increasing the yield of active, soluble recombinant protein.

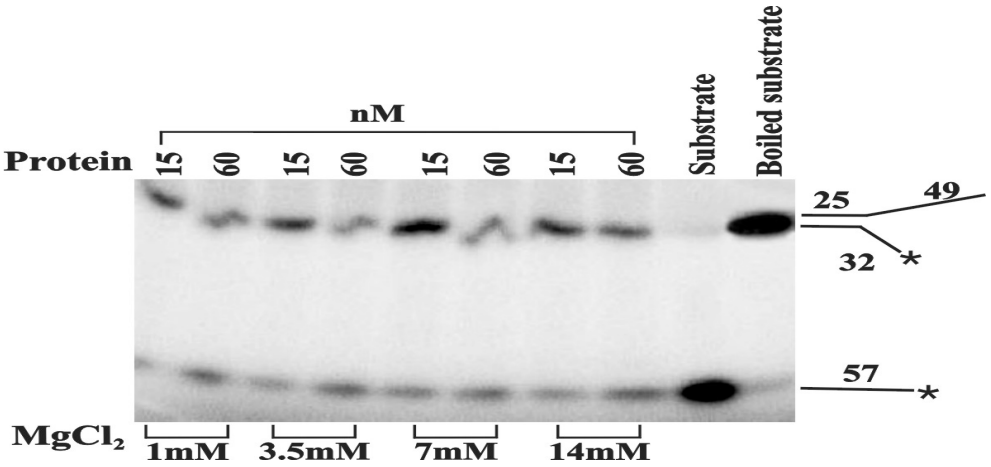
4.4 Reaction conditions of helicase assays for *M. maripaludis* Mcms

Co-purified MMP Mcm complex indicates that it forms a complex *in vitro* but it may be possible that a combination of homohexamers and heterohexamers were present within the purified MMP Mcm complex (Walters and Chong, 2010). As helicase activity of MMP McmA was quite strong (as discussed in section 4.3), it might be possible that it plays a major role in the helicase activity of MMP Mcm complex. Also the concentration of purified MMP McmA was 20 mg/ml and relatively purer then other MMP Mcms. For this reason, in further helicase assays MMP McmA was used. A number of different experiments were carried out to determine the optimum conditions for measuring helicase activity of this protein.

4.4.1 Magnesium chloride and helicase assays

Magnesium is an important component of helicase assays and without magnesium, unwinding is not possible (Costa et al., 1999). Helicase assays were performed in the presence of different magnesium chloride concentrations while the rest of the reaction conditions were the same as described in section 2.7.2. Unwinding was then quantified using densitometry on Quantity One software (Bio Rad). After quantification, the percentage values of DNA displaced were shown graphically (Figure 4.8) taking boiled substrate as 100% DNA displaced. 7 mM magnesium chloride gave the best results for helicase assays. This result is consistent with Costa et al., 1999 where they have shown in HDH VIII (human DNA helicase) 10 mM is an optimal concentration for helicase assays.

A



B

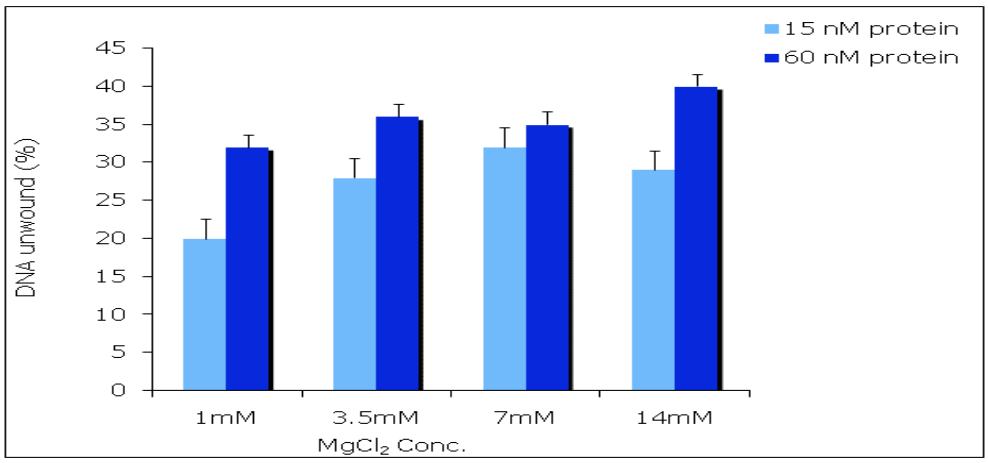


Figure 4.8 Effect of magnesium chloride concentrations on helicase activity of MMP McmA: Representative gel (A) under increasing magnesium chloride concentrations shown protein dependent helicase activity. (B) The helicase activity was quantified using densitometry. Boiled substrate was considered to displace DNA 100% in graphic values. Error bars indicate standard error of the mean. The helicase activity of MMP McmA was measured under increasing magnesium chloride concentrations. The x-axis shows magnesium chloride concentrations while the y-axis shows the percentage of DNA unwound. 7 mM magnesium chloride was found to produce optimal unwinding activity over two protein concentrations.

4.4.2 Incubation period of helicase assays

Helicase assays were carried out at the optimum temperature of 37 ^{o}C over 20 minutes to 180 minutes for 15 nM and 60 nM protein concentrations. Maximum helicase activity was measured by densitometry of displaced band to occur in the reaction of 60 minutes. After quantification, the percentage values of DNA displaced were shown graphically. Error bars indicate standard error of the mean (Figure 4.9).

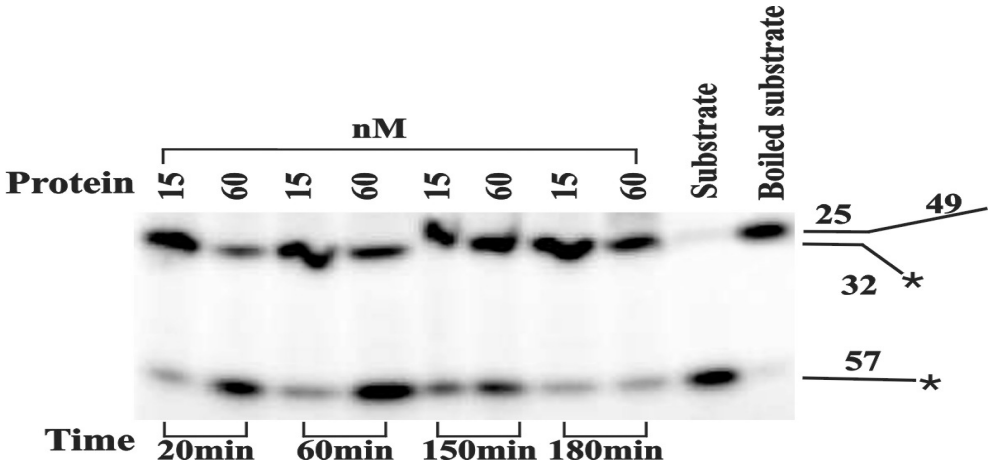
4.4.3 pH range for helicase assays

Helicase assays were performed at the temperature of 37 ^{o}C at different pHs. Three different concentrations of protein were used in these assays, which were 15 nM, 60 nM and 120 nM. The pH of helicase assay was maintained in the helicase buffer (Section 2.7.2). At pH 6 and pH 6.5, MMP McmA showed minimal unwinding. At pH 7.6 and pH 8, there was good improvement in unwinding efficiency. There was a maximum of 12 fold increase in helicase activity due to increasing pH of reaction mixture. Maximum helicase activity was measured by densitometry for the reaction at pH 8. Error bars indicate standard error of the mean in the graph (Figure 4.10).

4.4.4 ATP concentration for helicase assays

Helicase activity depends on the concentration of ATP. Without ATP, no unwinding can be detected (Costa et al., 1999). These assays were performed at two protein concentrations, which were 15 nM and 60 nM. The rest of the reaction conditions were the same as described in Materials and Methods section 2.7.2. ATP concentrations used in these assays were 2 mM, 4 mM, 6 mM, and 8 mM. Maximum unwinding activity was measured with the 6 mM ATP (Figure 4.11). Above 6 mM ATP, there was a decrease in the efficiency of unwinding. There is precedence for the increased ATP concentrations having an inhibitory effect on unwinding (Costa et al., 1999). Although from 2 mM to 6 mM maximum increase in helicase activity was

A



B

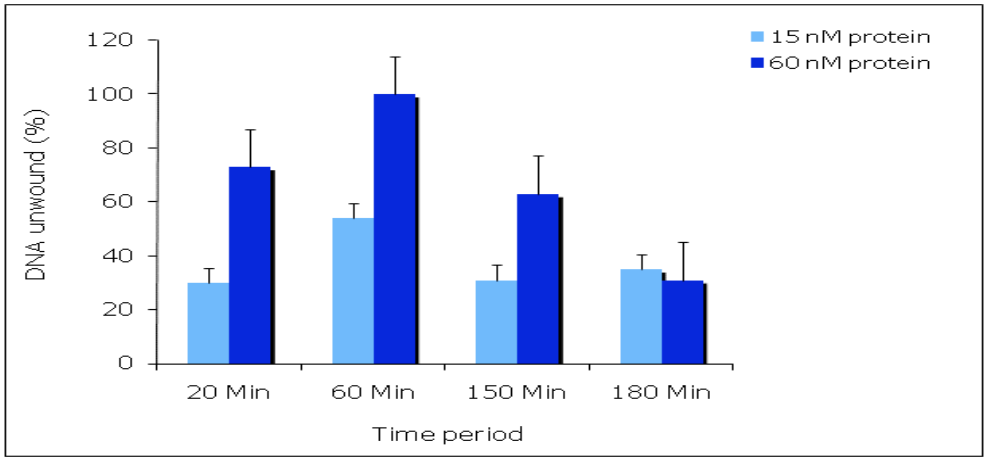
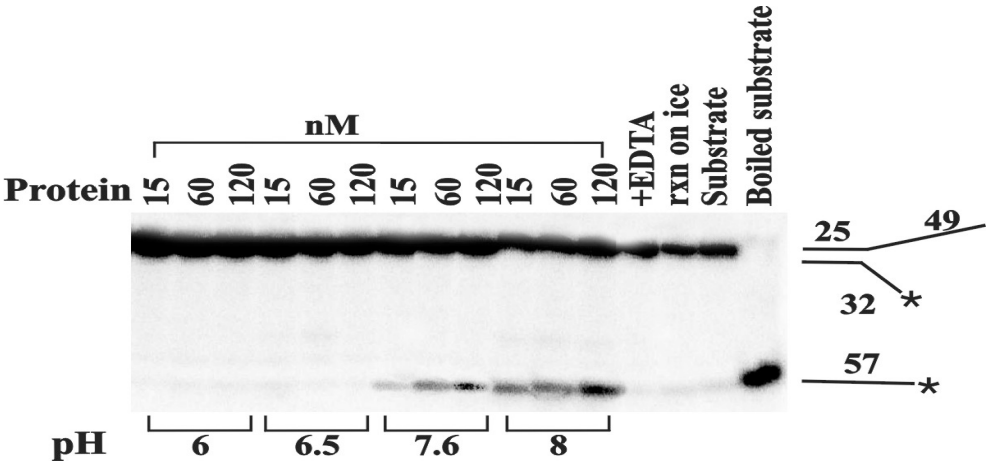


Figure 4.9 Effect of incubation duration on helicase activity of MMP McmA: The helicase activity of MMP McmA was measured over increasing time periods from 20 minutes to 180 minutes (A). The x-axis showed the increasing time period while the y-axis shows the percentage of DNA unwound (B). 60 minutes incubation shows maximum helicase activity for MMP McmA. There is a possibility that over a time after 60 minutes, DNA could be re-annealed in a reaction mixture.

A



B

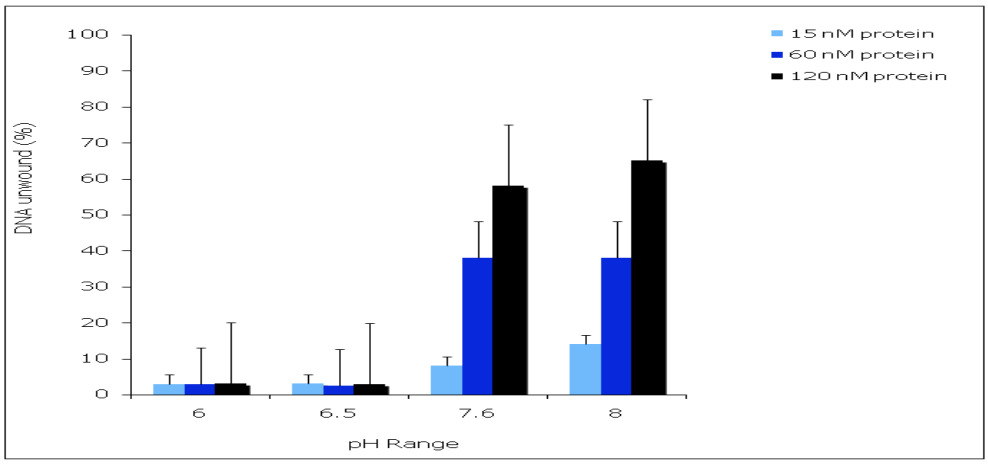
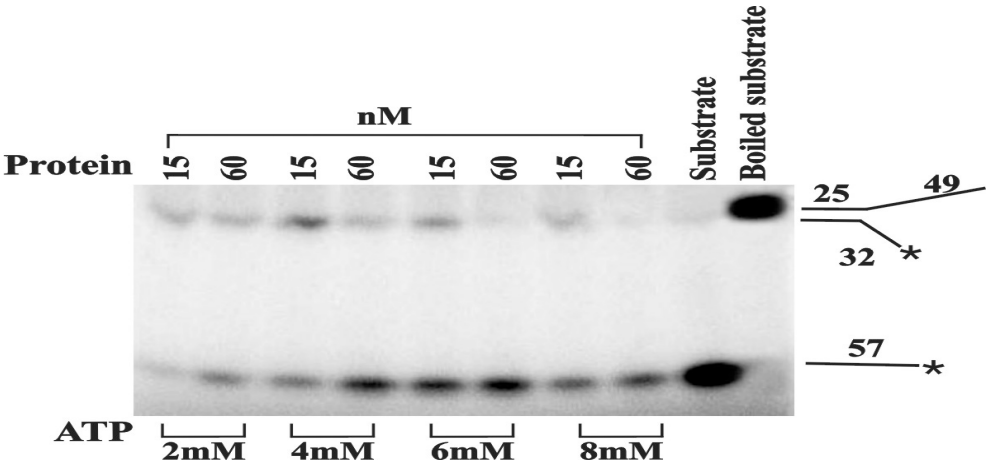


Figure 4.10 Effect of pH on helicase activity of MMP McmA: The helicase activity of MMP McmA was measured over pH range from 6 to 8. Representative gel (A) under increasing pH range showed increasing helicase activity from 6 to 8. (B) The helicase activity was quantified using densitometry. The x-axis shows the increasing pH range while the y-axis shows the percentage of DNA unwound. Mcm McmA at pH 8 displays the highest level of DNA helicase activity for MMP McmA.

A



B

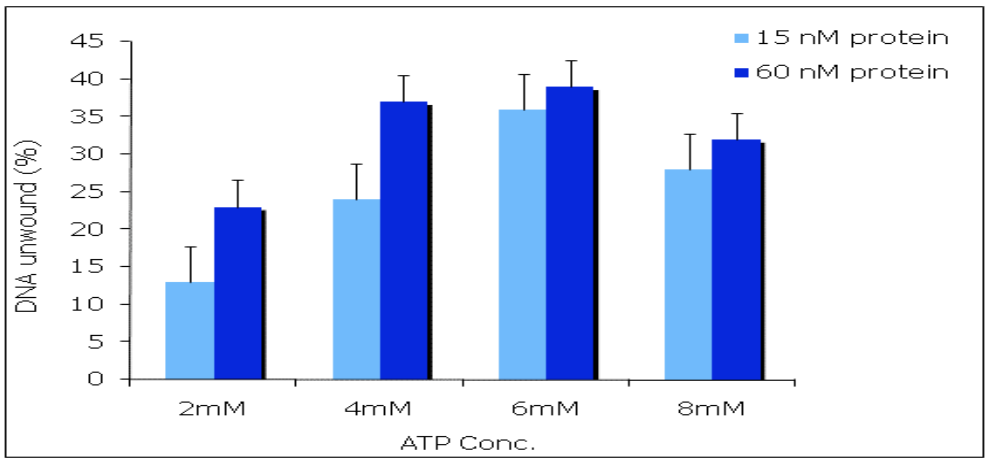


Figure 4.11 Effect of ATP concentrations on helicase activity of MMP McmA: Representative gel (A) under increasing ATP concentrations shown protein dependent helicase activity. (B) The helicase activity was quantified using densitometry. The x-axis shows an increasing concentration of ATP while the y-axis shows the percentage of DNA unwound. Maximum helicase activity was detected at 6 mM ATP concentration for both concentrations.

found to be between 2 fold to 3 fold.

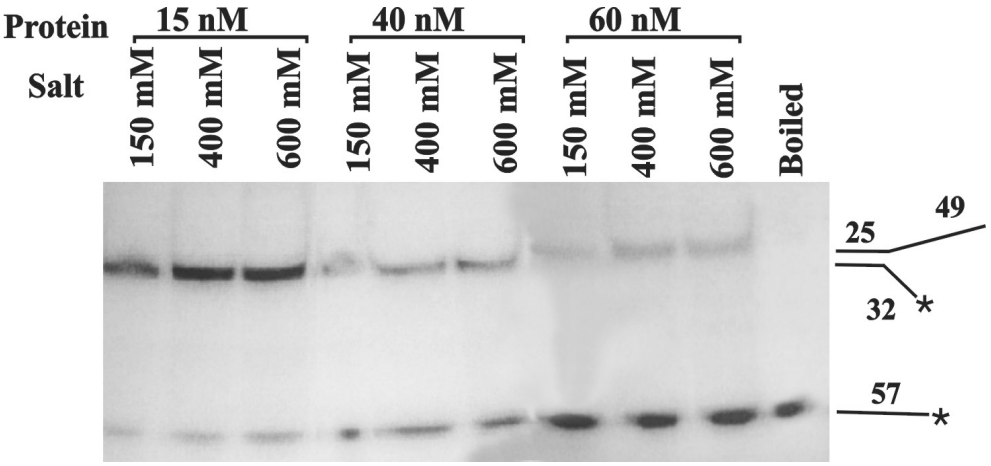
4.4.5 Effect of different salts and their concentration on helicase assays

Recent reports describe a significant effect of anions on Mcm helicase activity. More specifically this study suggested that chloride inhibits DNA unwinding by Mcms whereas glutamate supports Mcm helicase activity (Bochman and Schwacha, 2008). Therefore further investigation included the effect of other salts and their concentrations on unwinding ability of MMP McmA. The helicase activity of MMP McmA was detected in the presence of different salts and then quantified using densitometry. 15 nM and 60 protein concentrations were used to test the effect of different salts on unwinding. Sodium glutamate, potassium glutamate and sodium chloride were the salts to test on unwinding efficiency. Three different concentrations of each salt (150 mM, 400 mM, 600 mM) were tested. The rest of the assay was the same as described in Material and Method section 2.7.2. From the results, it was obvious that sodium chloride inhibited helicase activity and, sodium and potassium glutamate enhanced the activity (Figure 4.12) which further indicates that glutamate positively effects helicase activity while chloride inhibits helicase activity; this result is consistent with the recently published work discussed above.

4.5 Discussion

Although the purified MMP Mcm complex was available in the Chong lab, the concentration of complex was very low (0.258 mg/ml). Trials were conducted to purify co-expressed pairs of *M. maripaludis* Mcms protein. MMP McmA and MMP McmC were purified efficiently in a single step. MMP McmB and MMP McmD were purified in first step of GST affinity purification but the purified protein was not pure and additional lower molecular weight proteins were present. Also the concentration of MMP McmD was very low in affinity chromatography. To get rid of low molecular weight proteins, an additional step of anion exchange chromatography was carried out. In anion exchange chromatography, MMP

A



B

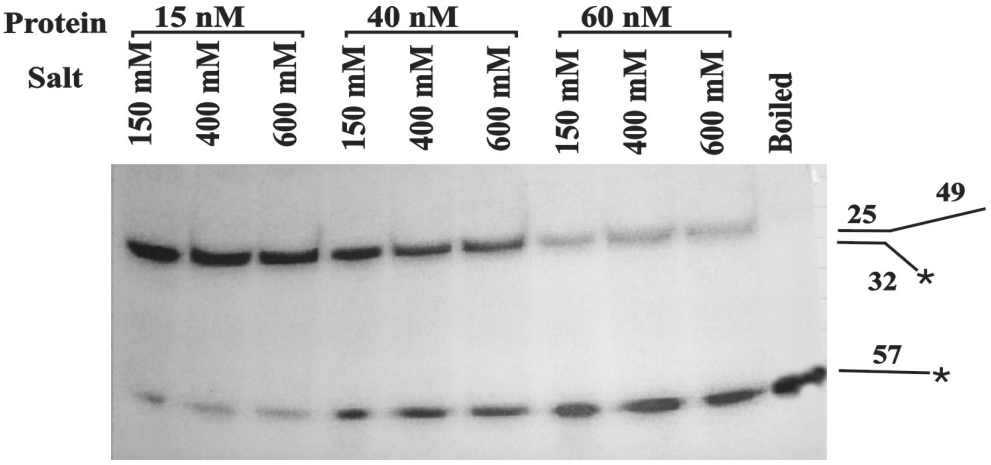
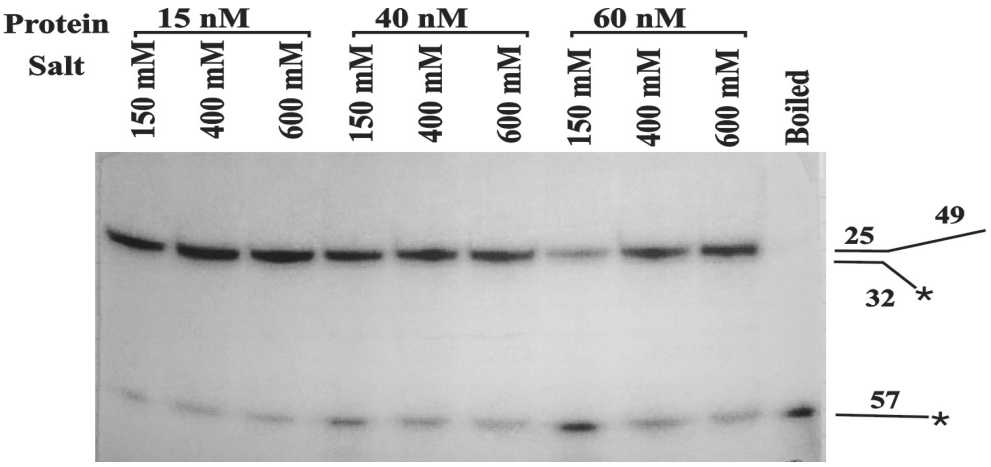
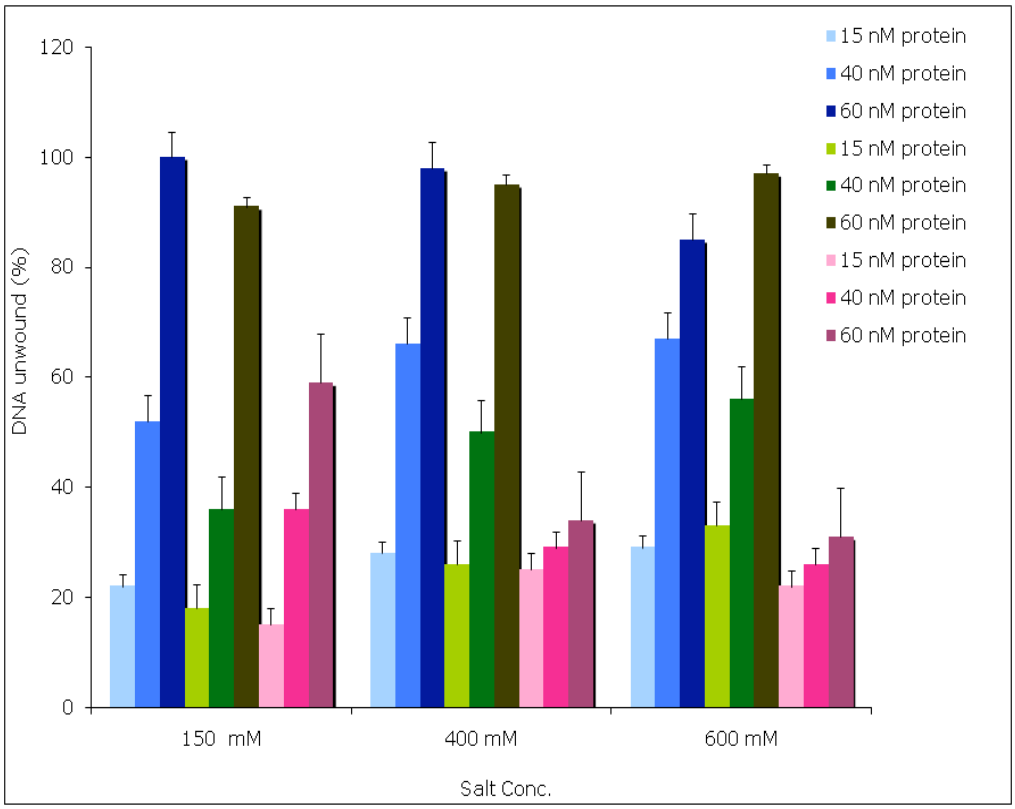


Figure 4.12 Effect of different salts and their concentrations on helicase activity of MMP McmA: The helicase activity of MMP McmA was measured in the presence of different salts and then quantified using densitometry. Three different salts (A) potassium glutamate, (B) sodium glutamate (C) sodium chloride were tested for their possible role in DNA unwinding. (D) The x-axis shows the increasing concentration of salts (150 mM, 400 mM, 600 mM) while the y-axis shows the percentage of DNA unwound (D). Shades of one colour represent the effect of one salt on different concentrations of protein. Consistent with published data sodium chloride inhibits helicase activity while sodium and potassium glutamate enhanced helicase activity.

C



D



Sodium glutamate (Blue)

Potassium glutamate (Green)

Sodium chloride (Pink)

Figure 4.12 Continued

McmB was eluted without any other contaminating proteins but MMP McmD was lost. But overall, concentrations of MMP Mcms (A, B and C) proteins were greatly enhanced by this method of purification. It was then decided to perform helicase assays on the available MMP Mcms.

Individual purified *M. maripaludis* Mcms proteins and complex were used to carry out helicase activity assays. Individually MMP McmA and MMP McmB showed robust protein dependent unwinding activity. MMP McmC did not display any helicase activity, while MMP Mcm complex (includes Mcms, A, B, C and D) also showed good helicase activity. MMP McmD was mainly contaminated and low in concentration, therefore helicase activity results from MMP McmD were largely non-informative. It is very interesting that although McmA is very similar to McmB and McmC in sequence but still showed more robust activity compared to McmB while McmC did not show any activity under same conditions. As MMP McmA efficiently unwound DNA substrate, therefore this protein was further used to find out optimum reaction conditions for unwinding. Magnesium chloride concentration is crucial for DNA unwinding, without magnesium chloride unwinding is not possible but at the same time high concentrations of chloride can inhibit unwinding. The optimum concentration found was 7 mM. Experiments to determine the best incubation period for unwinding suggested that 60 minutes of incubation gives best results. Furthermore the pH of helicase buffer used in unwinding also gives important insight. At lower pH, a low amount of DNA unwinding was measured. At pH 7.6 and pH 8.0, unwinding efficiency was very good. Helicase activity also depends on ATP in the reaction mixture. Without ATP, no helicase activity is possible. 6 mM ATP produced maximum unwinding and above this concentration, ATP inhibited helicase activity. The effect of different salts and their concentration on unwinding was also measured. Sodium chloride inhibited helicase activity and, sodium and potassium glutamate enhanced the activity which further indicates that the glutamate positively affects helicase activity while chloride inhibits helicase activity.

Chapter 5: The role of potential kinases in replication in archaea

5.1 Introduction

Although archaeal metabolism and operonic gene organization is certainly more similar to prokaryotic eubacteria, the archaeal factors for transcription, translation and DNA replication are more similar to those found in eukaryotes (Vas and Leatherwood, 2000). Most archaea replicate their circular genome from a single DNA replication origin as do bacteria, even though they may use eukaryotic-like proteins to do so (Myllykallio at al., 2000). For timely replication of big genomes in eukaryotes, it is important for replication to be started at multiple replication origins simultaneously. To coordinate replication initiation at various sites, there is a need for regulation that is in part carried out by cyclin-dependent kinases (Cdks) and cyclins (see reviews by Morgan, 1995; Roberts, 1999; Sherr and Roberts, 1999). Eukaryotic replication is also dependent on the replication specific kinase Cdc7-Dbf4 (Sclafani, 2000). Archaeal genomes are smaller and do not need as much coordination but archaea still need to couple replication with growth and division. Most archaea allocate less then 0.5% of their genome to signal transduction molecules (Galperin, 2005) compared to 4% in eukaryotes for kinases and phosphatases (Manning at al., 2002a and Manning at al., 2002b). Archaea lack recognizable homologues of cyclin dependent kinases, cyclins, or the kinase Cdc7.

Another reason to focus on kinases is that recently Dr Alison Walters in the Chong group co-precipitated MMP0004 together with MMP McmA (unpublished data, Chong lab). Bioinformatics studies give powerful information of particular proteins and their function. Based on sequence analysis studies, we tried to find the role of this hypothetical archaeal protein from *Methanococcus maripaludis* strain S2 (MMP0004). Another protein from *Methanocaldococcus jannaschii* strain DSM 2661 (Mj1073) that has 91% sequence homology with MMP0004 was also cloned and purified in order to examine the potential role of these proteins in replication of archaea.

5.2 Bioinformatics of *M. maripaludis* (MMP) 0004 and *M. jannaschii* (Mj) 1073

5.2.1 MMP0004

MMP0004 is a hypothetical protein that contains 298 amino acids and belongs to the RIO kinase family. RIO kinases are atypical protein serine kinases containing a kinase catalytic signature, but otherwise show very little sequence similarity to typical PKs (Angermayr et al., 2000). Serine kinases catalyze the transfer of the gamma-phosphoryl group from ATP to serine residues in protein substrates (Edelman et al., 1987). RIO2 is present in archaea and eukaryotes (Manning et al., 2002). The N-terminal domain found in RIO2 kinases is structurally homologous to a wHTH domain. It adopts a structure consisting of four alpha helices followed by two beta strands and a fifth alpha helix (LaRonde-LeBlanc and Wlodawer, 2004). The wHTH domain is primarily seen in DNA-binding proteins, although some wHTH domains may be involved in RNA recognition. RIO2 is essential for survival and is necessary for rRNA cleavage during 40S ribosomal subunit maturation (Vanrobays et al., 2001)**.** The biological substrates of RIO2 are still unknown (Campbell and Karbstein, 2010).

5.2.2 Mj1073

Mj1073 belongs to the RIO-type serine/threonine-protein kinase family. The BLAST searches showed that the similarity between MMP0004 and Mj1073 is 91%. From the protein database, it is shown to have metal binding, ATP binding and serine/threonine-protein kinase motifs. This protein belongs to a heterogenous group of serine/threonine protein kinases that are either non-specific or their specificity has not been determined to date but there is automated prediction of optimal substrate prediction available that use crystal structures, molecular modelling, and sequence analyses of kinases (Brinkworth et al., 2002). The functions associated with different regions of the protein and their corresponding amino acid positions according to bioinformatics studies are summarised in table 5.1 along with a cartoon representation (figure 5.1) of different features on sequence.

Feature Key	Position (s) Amino acids	Length Amino acids	Description
Molecular Processing			
Chain	1-270	270	RIO-type serine/threonine-protein kinase Rio2
Regions			
Domain	65-270	206	Protein kinase
Nucleotide binding	72-77	6	ATP
Nucleotide binding	158-164	7	ATP
Nucleotide binding	206-207	2	ATP
Sites			
Active site	202	1	Proton accepter
Metal binding	76	1	Magnesium 1
Metal binding	207	1	Magnesium 2
Metal binding	220	1	Magnesium 1
Metal binding	220	1	Magnesium 2
Binding site	93	1	ATP
Binding site	20	1	ATP

Table 5.1 Function of different regions of Mj1073 gene.

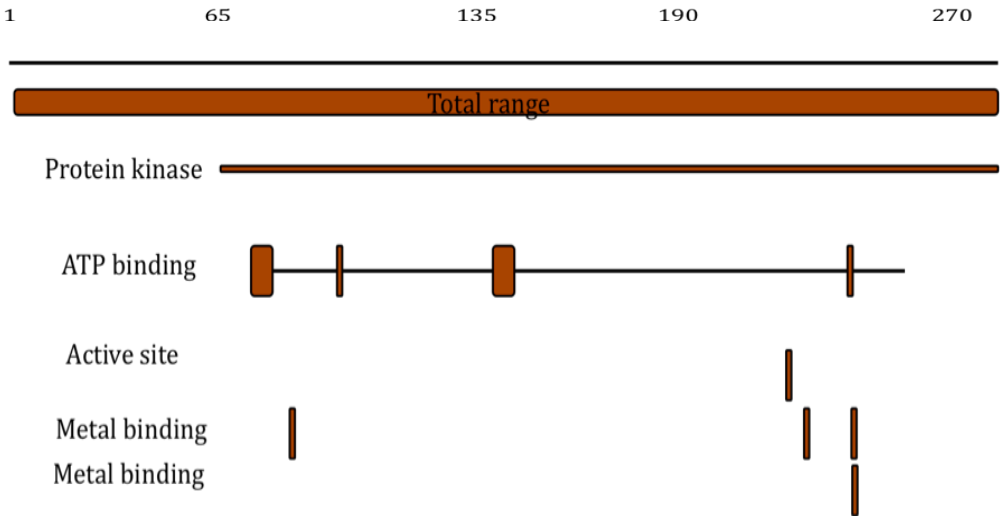


Figure 5.1 Cartoon representations of important regions and sites on Mj1073 gene sequence.

5.3 Phylogenetic analysis

Phylogenetic analysis was carried out on MMP0004 in order to find the closest homologues. The closest homologues of MMP0004 were found in *Methanocaldococcus jannaschii* and *Methanothermococcus okinawensis* (Figure 5.2). There are homologues of this protein present in lower eukaryotes and in yeast as well. Outside the *Methanococcales*, there is less homology. Although it is a hypothetical protein, the BLAST searches match this protein to other serine/threonine protein kinases that predict its role as a kinase.

5.4 Protein structure prediction

The prediction of the three-dimensional structure of a protein from its amino acid sequence was carried out using protein structure prediction software that include RaptorX (Peng and Xu, 2011), SWISS-MODEL (Guex and Peitsch, 1997; Schwede et al., 2003; Arnold et al., 2006) and M4T (Fernandez-Fuentes et al, 2007). These models are based on trying to fit the protein sequence to the most related structure available in the database. In MMP0004 and Mj 1073 case, this is the protein database (pdb) structure of 1zao, which is *A. fulgidus* RIO2 Kinase. It should be noted that this protein is quite distant compared to other available sequences as illustrated in figure 5.2. The generated model of MMMP0004 was evaluated by using QMEAN4 score in SwissModel Workspace. The global QMEAN4 scoring function (Benkert et al. 2008) is a linear combination of structural descriptors using statistical potentials. QMEAN4 is a reliability score for the whole model that can be used in order to compare and rank alternative models of the same target. This score is a composite score consisting of a linear combination of four statistical potential terms. The quality estimate ranges between 0 and 1 with higher values for better models. The QMEAN4 value for MMP0004 structure is 0.5 (Figure 5.3 A). Residue error plot showed two big error peaks (Figure 5.3 B) and rest of the plot seems fine and does have conserved residues (discussed later).

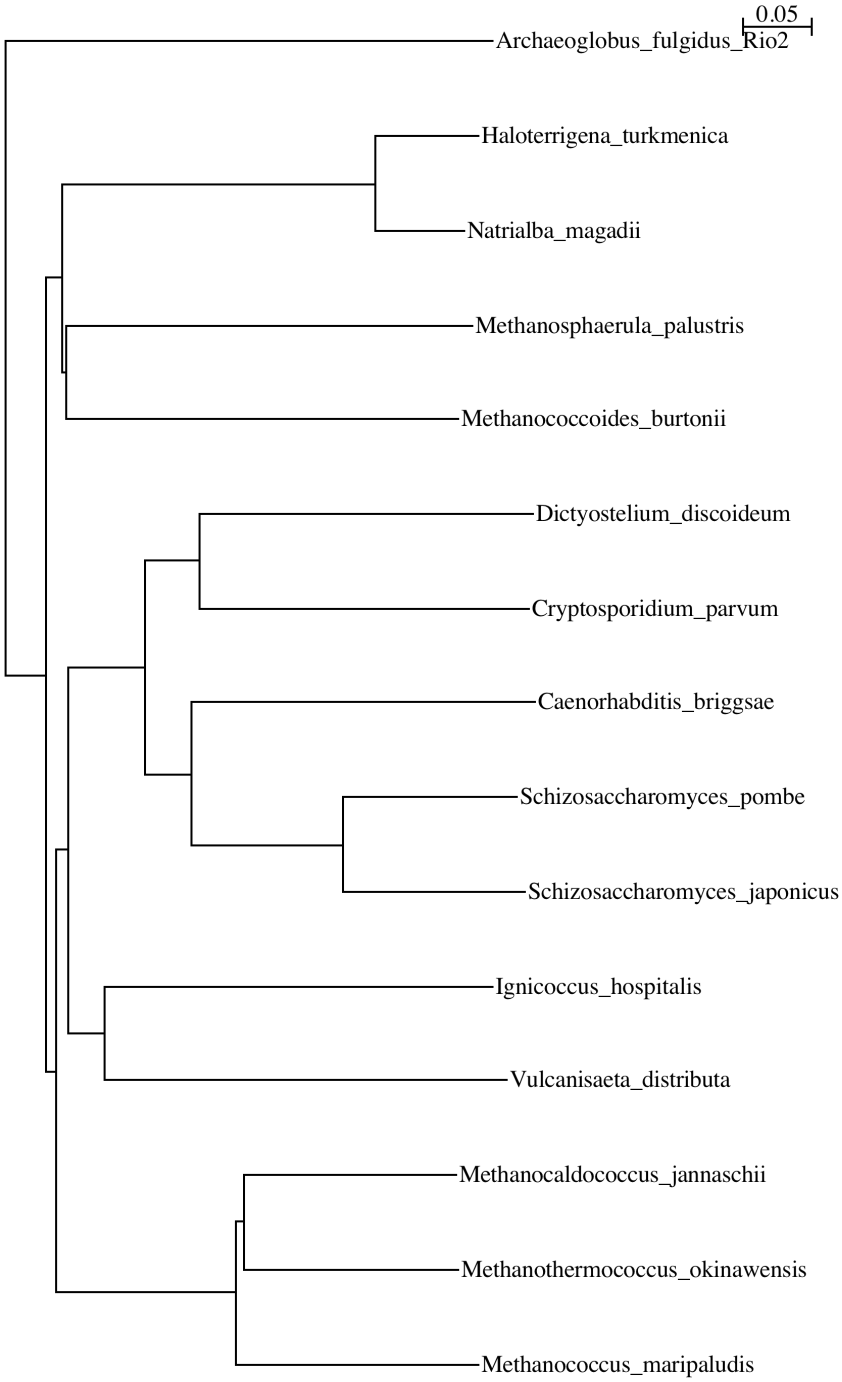
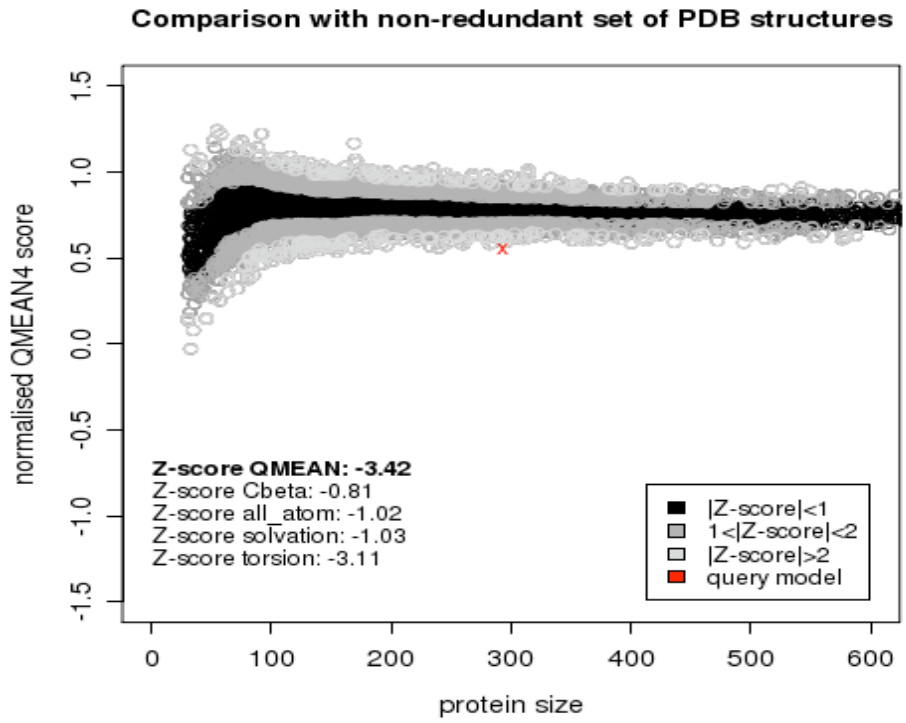


Figure 5.2 MMP0004 phylogenetic tree: A multiple sequence alignment shows that there is homology of MMP0004 with other proteins present in other archaea and lower eukaryotes. The closest homology is in *Methanocaldococcus jannashii* and *Methanothermococcus okinawensis.*

A



B

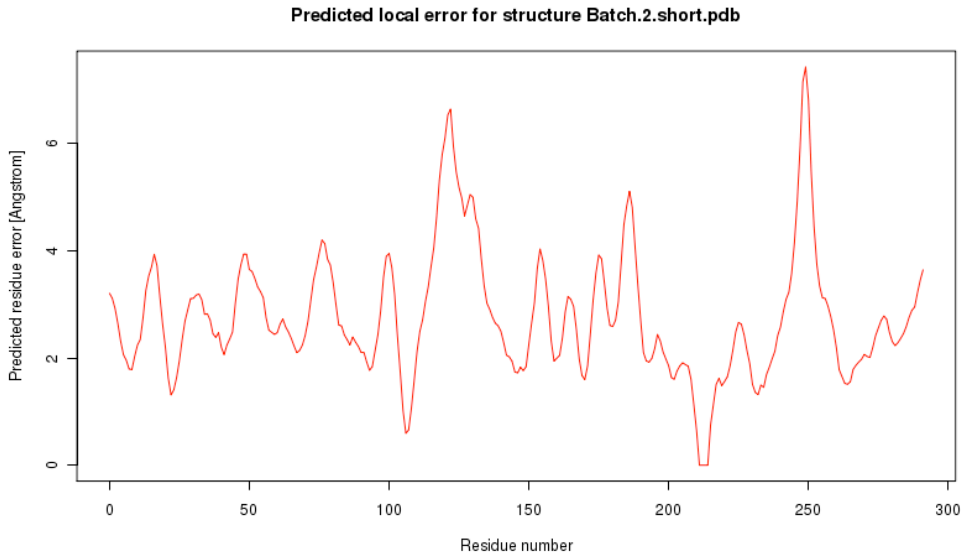


Figure 5.3 QMEAN4 score and residue error plot of MMP0004: QMEAN4 score (A) of MMMP0004 was evaluated by using SwissModel Workspace and 0.5 value of QMEAN score is obtained that lies exactly in the middle. Residue error plot (B) showed two big error peaks indicating regions of dissimilarity when compared to *A. fulgidus* RIO2 Kinase.

A number of structures of protein-serine/threonine and protein-tyrosine kinases have been determined, and the conserved subdomain residues have been shown to be involved in recognition and binding of ATP or substrate peptides, as well as in actual catalysis (Bossemeyer, 1995; Engh and Bossemeyer, 2002; Knighton et al., 1991a, 1991b). The predicted ribbon structures also have a specific ATP binding pocket that is conserved in all serine/threonine kinases. Predicted structures of both proteins (MMP0004 and Mj1073) are shown in figure 5.4 with the highlighted yellow residues on MMP0004 that may have role in DNA binding. Analysis of the structure of *A. fulgidus* (Af) RIO2 kinase revealed the presence of two domains, the N-terminal domain and the C-terminal domain (LaRonde-LeBlanc and Wlodawer, 2004). The two model structure of MMMP0004 (5.4 A) and Mj1073 (5.4 B) and the crystal structure *A. fulgidus* (Af) RIO2 Kinase (5.4 C) seems essentially the same. The N-terminal domain (wHTH domain) contains a combination of α helices and β strand in all three structures. Middle part of the ribbon structure is the same in all three proteins, that is five β strands with one long α helix. The C-terminal domain (RIO kinase domain) is the most different part in all three proteins as the number of β strands are different in all three proteins

5.4.1 Comparison of the wHTH domain with its structural equivalance

A multiple sequence alignment of the MMP0004 homologues from archaea showed that all these proteins possess the motifs believed to be important in kinase function (Figure 5.5 and 5.6).

The N-terminal domain is conserved in members of the Rio2 family. A strand that connects the second and third α helices combines with the other two β strands to form a β sheet, and the loop between the second and third β strands is called a “wing,” which gives the wHTH its name. The fifth α helix replaces the second wing seen in some proteins with the wHTH fold (LaRonde-LeBlanc and Wlodawer 2004). The most commonly reported function of such domains is DNA binding (Gajiwala et al., 2000; Gajiwala and Burley, 2000). From the predicted structural data (Figure 5.4) and

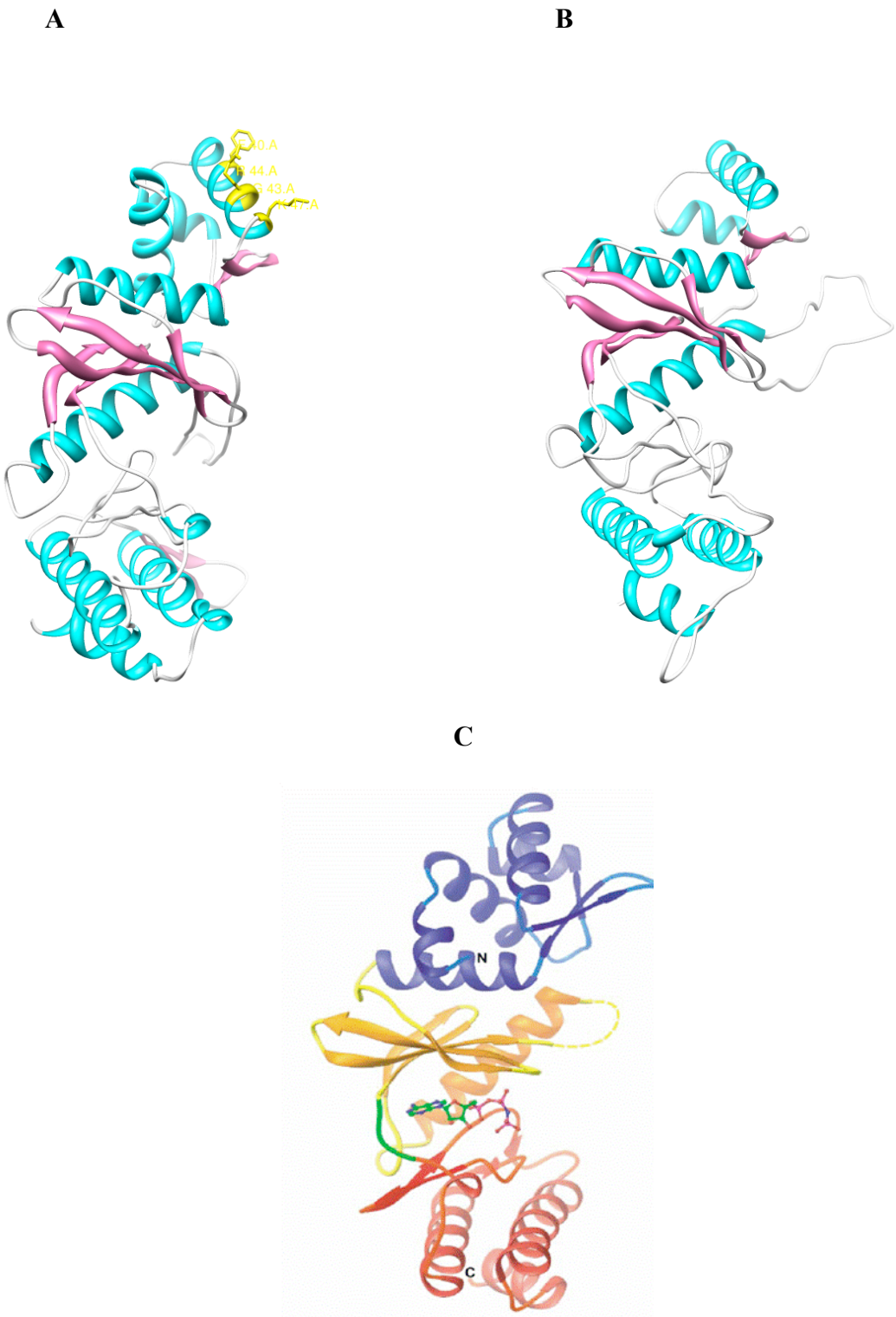


Figure 5.4 Overall structures: Predicted ribbon structures of MMP0004 (A) and Mj1073 (B) and crystal structure of *A. fulgidus* Rio2 (C). α Helices are shown in green colour while β Sheets are pink in colour in A and B. MMMP0004 model was generated in SwissModel while Mj1073 is regenerated from RaptorX while the crystal structure of *A. fulgidus* Rio2 was taken from LaRonde-LeBlanc and Wlodawer, 2004.

sequence alignment of the N- terminal domain (Figure 5.5), it is clear that the MMP0004 protein is missing one α helix named as α0 while Mj1073 is missing two α helices namely α0 and α1. Similarly there are three β strands present in Af. Rio2, this is not the case with MMP0004 and Mj1073, they both are missing one β strand and most probably that is one called β3 (Figure 5.5). One possible reason might be that a few residues are missing from this segment in the alignment and also this sequence does not contain any homology with any residue with *A. fulgidus* Rio2.

5.4.2 The RIO kinase domain

The C-terminal RIO2 domain, the sequence of which is conserved in both Rio1 and Rio2 proteins, is indeed structurally homologous to known protein kinase domains (LaRonde-LeBlanc and Wlodawer, 2005a, b). It is bilobal with a twisted five-stranded β sheet and a single long α helix (αC) in the N-terminal lobe whereas the C-terminal lobe consists of a combination of four β strands and three α helices (Figures 5.4 C and 5.6) in the case of *A. fulgidus* Rio2. The N-terminal lobe of RIO kinase domain in MMMP0004 and Mj1073 is exactly the same as Af. Rio2 in predicted structures but the C-terminal lobe consists of two β strands and four α helices incase of MMP0004 (Figures 5.4 A and 5.6). and three α helices but no β strands in Mj1073 predicted ribbon structures (Figures 5.4 B and 5.6).

A loop of amino acids between β strand 3 (β3) and α helix C (αC) contains a highly conserved threonine in all three of these proteins, possible targets for phosphorylation (Figure 5.6).

5.4.3 Conserved residues of the RIO proteins

The structure of Rio2 from *A.fulgidus* was analyzed to locate the residues that were conserved in the RIO protein family and when mapped on the surface, the conserved residues appeared to cluster around the active site on one face of the protein (LaRonde-LeBlanc and Wlodawer, 2004) and included the following residues. Tyr222 was conserved as either a Tyr or a

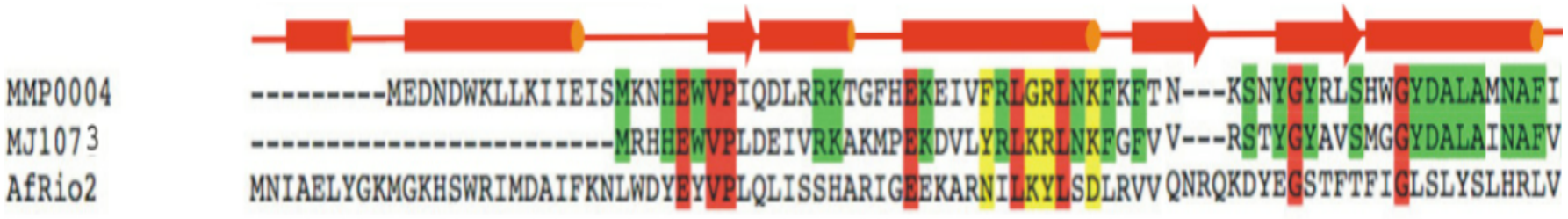


Figure 5.5 The Rio2 winged helix domain: Sequence alignment of the N-terminal domain of Rio2 from *A. fulgidus* (AfRio2), *M. maripaludis* (MMP0004) and *M. jannaschii* (Mj1073). Residues identical in all the sequences are highlighted in red, while green highlighted residues are identical in MMP0004 and Mj1073 proteins only. Yellow highlights indicate residues that may be involved in DNA binding in the Rio2 proteins. Phe and Tyr are colored as identical in all cases.

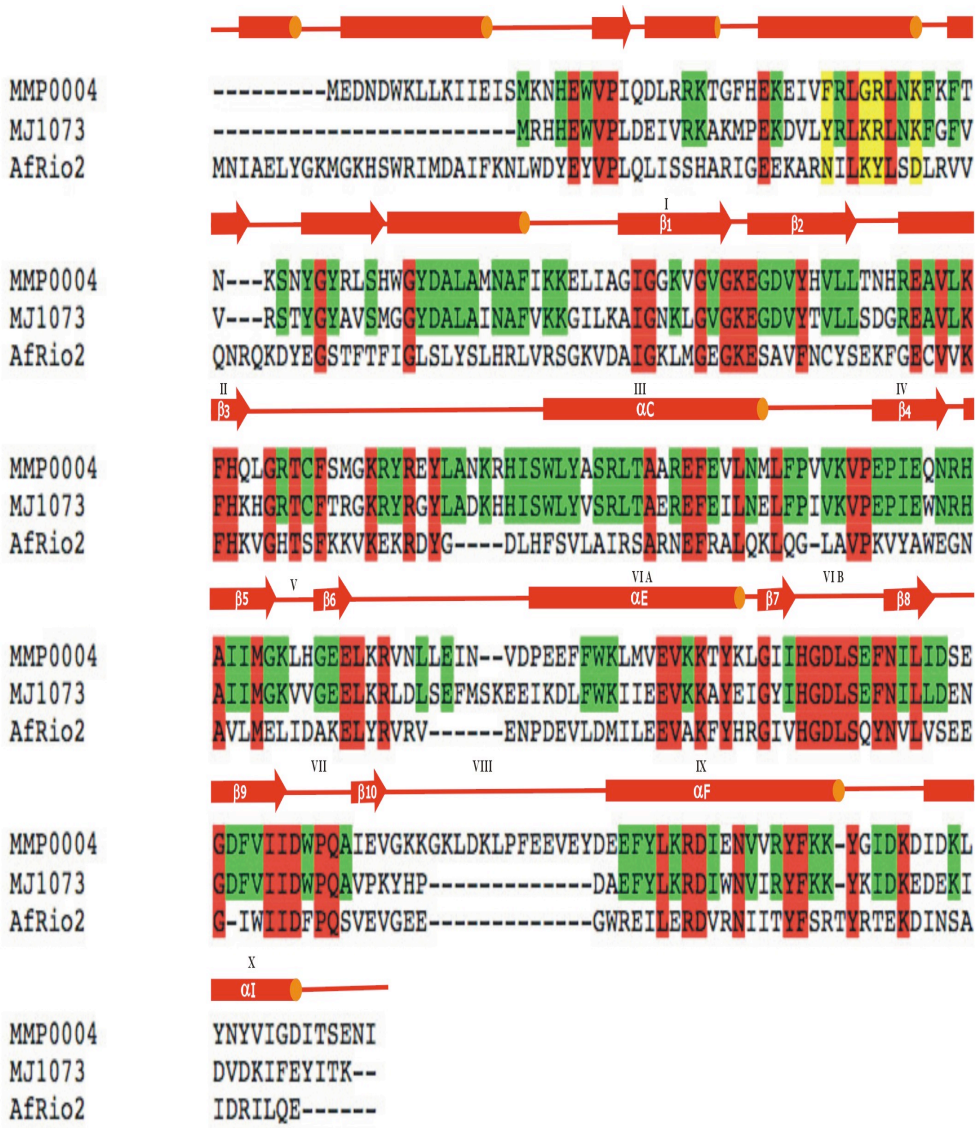


Figure 5.6 The RIO kinase domain: Sequence alignment of the kinase domains of Rio2 from *A. fulgidus* (AfRio2), *M. maripaludis* (MMMP0004) and *M. jannaschii* (Mj1073). Schematics of the secondary structure are shown in blue (Rio2). Residues highlighted in red are identical in AfRio2, MMP0004 and Mj1073. Residues highlighted in green are identical in MMP0004 and Mj1073 but not in AfRio2. Phe and Tyr are colored as identical in all cases. Kinase subdomains based on LaRonde-LeBlanc and Wlodawer, 2004 are indicated with black Roman numerals.

Phe (β3). Glu103 was located in the glycine-rich loop (β1) and is conserved in all RIO proteins. Lys102 (β1) was also conserved in all RIO domains but did not interact with the nucleotide and was disordered in the structure. The disordered loop (residues 125–141, between β3 and αC) was located near the conserved surface and contained Gly125, Thr127 and Tyr148, all of them are highly conserved among Rio2 family members including MMP0004 and Mj1073 (Figure 5.6).

5.5 Possible protein/protein interactions

Furthermore MMP0004 protein sequence was analysed in other bioinformatics studies for its possible interactions with other proteins within *M. maripaludis* S2 strain. For this purpose online the string database was used. STRING is a database of known and predicted protein interactions (www.string-db.org.). The interactions include direct (physical) and indirect (functional) associations. With a medium confidence level (score 0.400), MMP0004 shows predicted interactions with a number of different proteins that are either hypothetical or uncharacterised proteins. With a high level of confidence (score 0.700), MMP0004 interacts with 30s ribosomal proteins (Figure 5.7). The activity of Rio2 is necessary for rRNA cleavage in 40S ribosomal subunit maturation. The interaction of MMP0004 with ribosomal proteins, justifies its role as a serine/threonine kinase of the RIO2 family because the RIO2 family is necessary for rRNA cleavage. The network of interacting proteins generated is with medium confidence level (Figure 5.7). It is interesting that no interaction has been shown with any of the Mcms as Dr Alison Walters in the Chong group recently showed that the MMP McmA interacts with MMP0004 *in vivo* (unpublished data, Chong lab).

5.6 Cloning of MMP0004

The MMP0004 gene was cloned from genomic DNA of *M. maripaludis* S2 available in the Chong lab. The MMP0004 gene was amplified via touch master mixture instead of Taq polymerase. The PCR product was first cleaned using QIAquick PCR clean-up kit (QIAGEN) and blunt ligated into

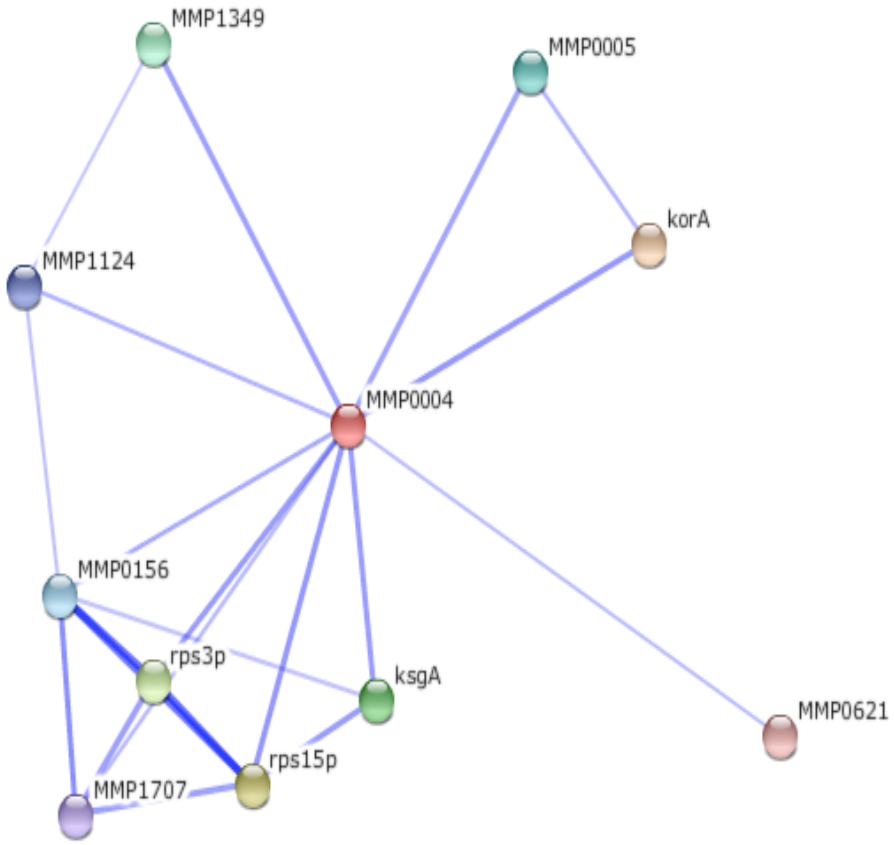


Figure 5.7 Possible interaction of MMP0004: *In vivo* interaction of MMP0004 with other proteins within *M. maripaludis* S2. Most of the interactions are with ribosomal proteins. No interaction has been shown with any of the Mcm. Thicker lines represent stronger associations.
http://string-db.org/

down PCR using primers containing appropriate restriction sites (see appendix 2 for primer details). FideliTaq DNA polymerase was used in cloning vector pJET 1.2 following manufacturer's protocol. After various failed trials of subcloning MMP0004 into an expression vector from pJET 1.2, MMP0004 was subcloned into another holding vector pGEM®-T using manufacturer's protocol. The resulting vector and pET 19b were digested with *BamH1* and *Xhol* and ligated overnight at 4 ^{o}C. To confirm the ligation had worked, pET 19b containing MMP0004 was digested (Figure 5.8). pET 19b construct containing MMP0004 was sent for sequencing to the sequencing laboratory of York University (Technology Facility). Transformation was carried out using *E.coli* BL21 (DE3) Rosetta competent cell. In transformation protocol LB was used instead of SOC media.

5.7 Expression of MMP0004

MMP0004 was expressed in 750 ml culture which yielded 4 gram of pellet. Cells, which were harvested by centrifugation, resuspended immediately in resuspension buffer (Section 2.5.1) before storing at -80 ^{o}C prior to protein purification.

5.8 Purification of MMP0004

Resuspended cells stored at -80 ^{o}C were thawed on ice and protein was purified in the same way as explained in section 2.5. Elution fraction samples were run on 15% SDS-PAGE as the size of this protein was small. Eluant fractions were pooled together and flash frozen in liquid nitrogen and stored at -80 ^{o}C. SDS-PAGE analysis showed clean and intense band at expected size (29.9 KDa) just above 25 KDa protein marker (Figure 5.9). Purified proteins were quantified using Bradford assays. The concentration was 3.3 mg/ml in a volume of 15 ml. The purified protein was concentrated as explained in section 2.9.7 in Materials and Methods.

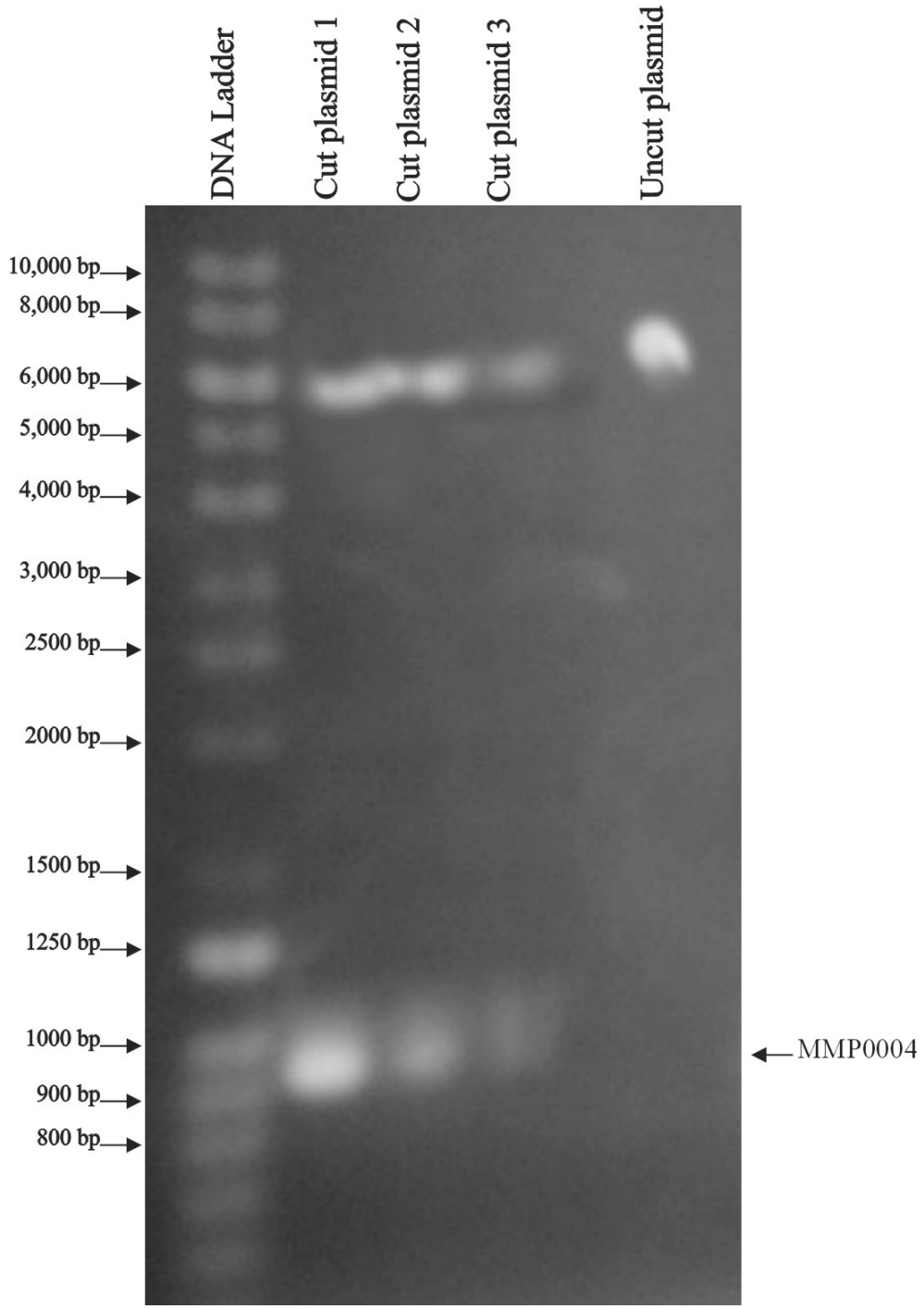


Figure 5.8 Restriction digestion of pET 19b containing MMP0004: 5 µg of DNA was digested over the period of 1 hour at 37 ^{o}C and loaded in three lanes of 2% agarose gel in TAE along with uncut vector of pET 19b. The agarose gel showed that MMP0004 (897 bp) was ligated into pET 19 b (5717 bp).

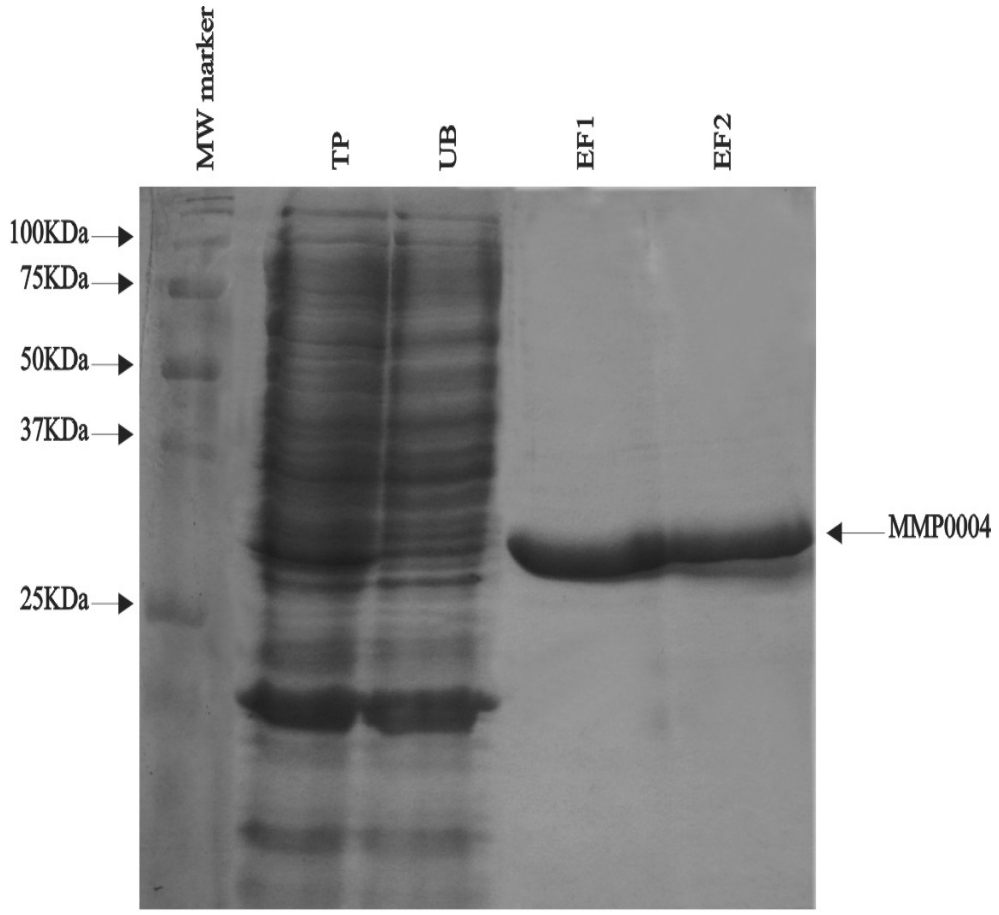


Figure 5.9 Purification of MMP0004: SDS-PAGE analyses of fractions from purification trials of MMP0004. Protein was expressed in *E.coli* BL21 (DE3) cells Rosetta at 37 ^{o}C overnight. Pellets from 750 ml of culture were resuspended, and the soluble protein fraction was bound, washed and eluted from nickel beads using 500 mM imidazole. A band of ~30 KDa was present in elution fraction visible after Coomassie blue staining. Protein ID was confirmed by mass spectrometry (MS).

TP, total protein; UB, unbound fraction; EF, elution fraction

5.9 Kinase assays for MMP0004

The protein database available on the NCBI website showed that MMP0004 is a putative protein serine/threonine kinase. Kinase assays were used to test whether MMP0004 acts as Mcm kinase. As the substrates for MMP0004 were unknown and MMP0004 has been pulled down together with MMP McmA from *M. maripaludis* extracts (unpublished data, Chong lab), it was hypothesized that the substrate for MMP0004 was Mcms. Kinase activity assays were carried out to test the phosphorylation on purified MMP McmA by MMP0004. MMP McmA was used instead of MMP Mcm complex as the purity and concentration of MMP McmA was much higher then MMP Mcm complex. Also MMP McmA showed robust helicase activity as discussed in previous chapter. The results from the kinase assays showed that there was not any phosphorylation by MMP0004 on MMP McmA. It may be due to substrate choice or requirement for other proteins or cofactors in these assays. The gel picture was not good quality (not shown), as the $\gamma[^{32}P]$ATP required 4-12 % gradient gel that was not available. Instead the samples were run on 12% TBE gel that made the whole gel picture black.

Further to confirm results of kinase activity of MMP0004 on McmA, helicase assays were carried out on kinase treated McmA. The helicase assay carried out on McmA is described in section 2.7.2. Interestingly McmA did not show any helicase activity after being included in kinase assay with MMP0004. The positive control was the boiled substrate while the negative control was substrate without any protein added (Figure 5.10). There may be the possibility that McmA proteolysis took place in kinase assays and then further in helicase assays, and that may be the reason it did not show any helicase activity afterwards.

5.10 Cloning of Mj1073

The Mj1073 gene from DSM 2661 strain was cloned from the whole genomic DNA of *Methanocaldococcus jannaschii* present in the Chong lab. The Mj1073 gene was amplified using touch down PCR with primers

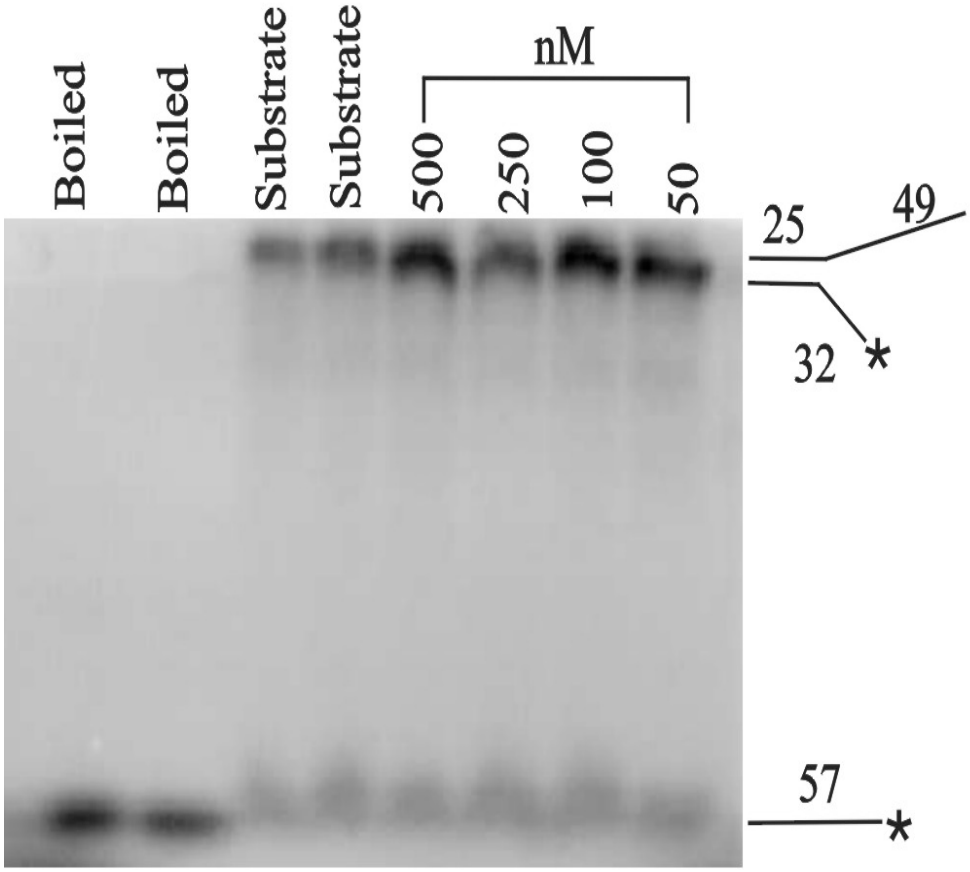


Figure 5.10 Helicase activity of kinase treated MMP McmA: PAGE analysis of strand displacement assays of kinase treated MMP McmA with various protein concentrations. Controls were generated by performing the helicase assays with intact substrate (negative control) and by boiling substrate (positive control). Kinase-treated MMP McmA did not show any helicase activity.

containing appropriate restriction sites (see appendix 2 for primer details). The method of cloning of Mj1073 was essentially the same as MMP0004 explained in section 5.2.2 except the initial annealing temperature of primers for Mj1073 was quite low (46 ^{o}C). PCR product was first cleaned using QIAquick PCR clean-up kit (QIAGEN) and blunt ligated into cloning vector pGEM®-T. The resulting vector and pET 19b were prepared in a similar manner with *BamH1* and *Xhol* restriction sites and ligated overnight at 4 ^{o}C. 2 µl of ligation mix was transformed into Rosetta DE3 *E.coli* competent cells. To confirm the ligation has worked, digested the pET 19b containing Mj1073 (Figure 5.11). Sequencing laboratory of York University (Technology Facility) confirmed the sequence of Mj1073 in pET 19b without any mutation. Transformation was carried out using *E.coli* BL21 (DE3) Rosetta competent cells.

Making glycerol stocks was not possible as storing these cells at -80 ^{o}C did not yielded much protein and resulted in the loss of plasmid, so direct purification step was taken for this protein from fresh transformation every time.

5.11 Expression of Mj1073

Mj1073 was expressed in 750 ml culture which yielded 3 grams of pellet. Cells, which were harvested by centrifugation, resuspended immediately in resuspension buffer (Section 2.5.1) before storing at -80 ^{o}C prior to protein purification.

5.12 Purification of Mj1073

The protocol for the purification of Mj1073 protocol was essentially the same as for MMP0004 (Section 5.8). Eluted fraction samples were run on 15% SDS-PAGE and protein was visualised using Coomassie blue staining. Elution fraction 1 and 2 displayed a pure band of Mj1073 at the expected size of 27 KDa (Figure 5.12). Purified proteins were quantified using Bradford assays. The concentration was 5.1 mg/ml in a volume of 12 ml.

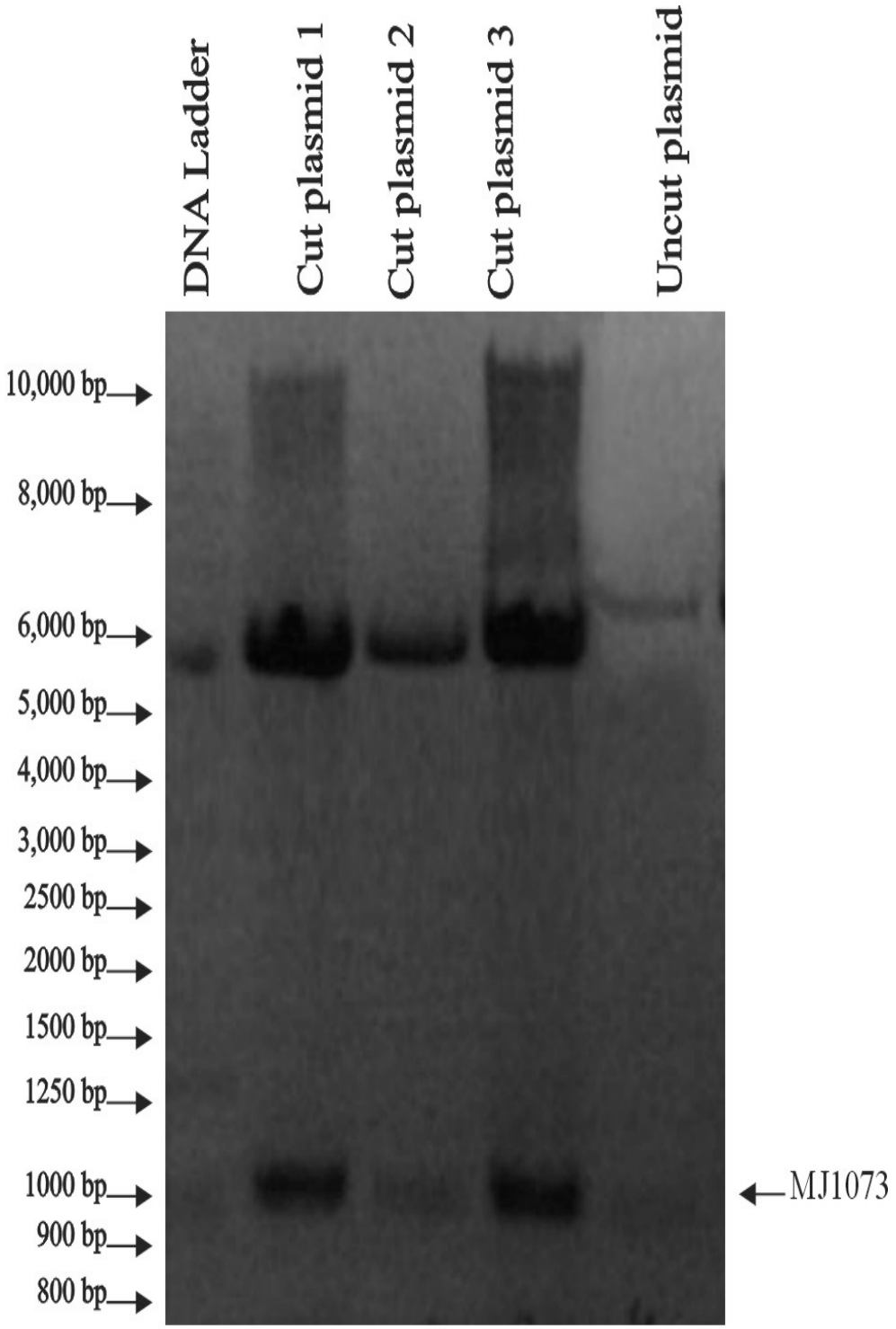


Figure 5.11 Restriction digestion of pET 19b containing Mj1073: 5 μg of DNA was digested over the period of 1 hour at 37 ^{o}C and loaded in three lanes of 2% agarose gel in TAE along with uncut vector of pET 19b. The agarose gel showed that Mj1073 (810 bp) was ligated into pET 19 b (5717 bp).

Eluant fractions were pooled together and flash frozen immediately in liquid nitrogen then stored at -80 ^{o}C.

5.13 Discussion

MMP0004 protein has been identified in *Methanococcus maripaludis* S2 as possibly interacting with Mcms. The role of MMP0004 has been suggested to be a kinase by bioinformatics studies based on its sequence homology with other serine/threonine kinases. The study of this protein in relation to Mcm leads to the hypothesis that this protein may act as a kinase using Mcm as a substrate. As this protein is hypothetical, and the only role that was assigned to it was as a putative kinase by BLAST searches, the substrates for this kinase are not known. The closest homology of this protein was present in *Methanocaldococcus jannaschii* (Mj1073). Predicted ribbon structures were obtained from different online data servers and all produced similar structures that were based on the most homologus known protein crystal structure of *A. fulgidus* RIO2 kinase. Both of these proteins are fairly small, comprising of 29 KDa and 27 KDa for MMP0004 and Mj1073 respectively. Mj1073 is likely to be more thermostable as *M. jannaschii* is an hyperthermophilic organism (Bult et. al., 1996). Both proteins were cloned and purified. Purified MMP0004 was further subjected to kinase assays but did not phosphorylate MMP McmA. Either MMP0004 does not use McmA as a substrate or it requires other proteins or cofactors to be present to produce kinase activity. It is also possibe that it does not have any kinase activity but the helicase assays of kinase-treated MMP McmA did not show any DNA unwinding activity. MMP McmA did show helicase activity under normal circumstances as described in the previous chapter. As the substrate for Mj1073 was unknown, so the kinase assays for Mj1073 were not performed. But both proteins were successfully purified and can be used in future for the identification of their role and further insight to the kinase activity related to these proteins.

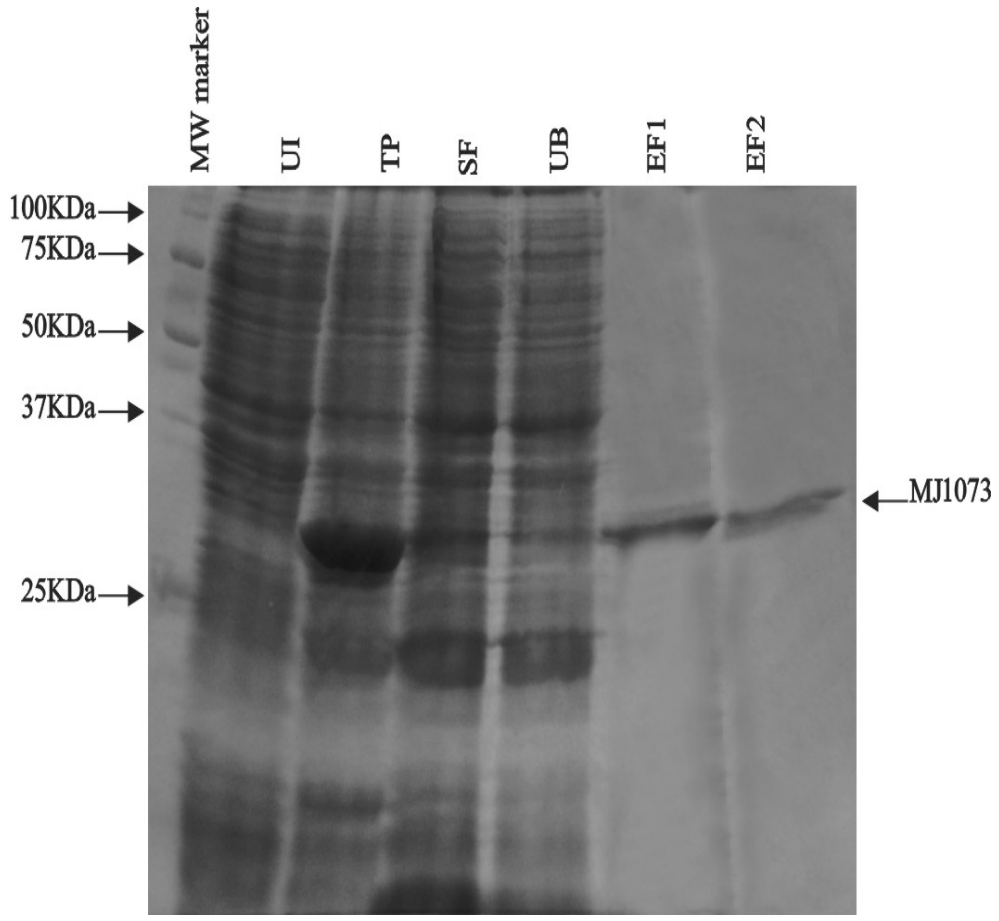


Figure 5.12 Purification of Mj1073: SDS-PAGE analyses of fraction from purification trials of Mj1073. Protein was expressed in *E.coli* BL21 (DE3) Rosetta cells at 37 ^{o}C overnight. 500 ml culture pellet was resuspended bound, washed and eluted from nickel bead using 500 mM imidazole. A band just above 25 KDa marker is visible at the expected size for Mj1073 (27KDa).

UI, uninduced; TP, total protein; SF, soluble fraction; UB, unbound fraction; EF, elution fraction

Chapter 6: Discussion

6.1 Reconstitution of hMcm2-7 complex

The minichromosome maintenance proteins are components of the prereplicative complex (preRC) that assembles on replication origins prior to S phase (Donovan et al., 1997; Tanaka et al., 1997). The six proteins Mcm2-7 form a heterohexameric ring that is the replicative helicase in eukaryotic cells (Schwacha and Bochman, 2008). Mcm proteins are required for replication initiation and elongation (Labib et al., 2000; Lei and Tye, 2001). In addition, they appear to travel with the replication fork *in vivo* (Aparicio et al., 1997). All six of the Mcm proteins are required for ongoing replication in *S. cerevisiae* (Labib et al., 2000). The purification of yeast Mcm2-7 as individual subunits by expressing individually in bacteria was reported (Davey et al., 2003). Recombinant individual human Mcm2-7 were expressed in *E.coli* in this study. It was not possible to obtained individual hMcm2-7 in purified form to conduct further assays. Only hMcm4, 6 and 7 were obtained in low concentrations after a number of different purification trials. Mcm4, 6 and 7 were used further to assemble a complex without success. hMcm2 was purified in very small amounts by expressing it at low temperature in Arctic Express (DE3) RIL cells. Despite various efforts to purify hMcm3 and 5, they remained largely insoluble.

At that point it was decided to try to co-express and co-purify all hMcms from *E.coli.* All six hMcms were co-purified through four purification steps successfully when co-expressed in *E.coli,* suggesting that they form a heterohexameric complex (Lee and Hurwitz, 2000; Davey et al., 2003). However the complex concentration was low and only sufficient to perform further activity assays but for stoichiometric analysis, the complex needed to be pure and more concentrated. Therefore stoichiometry of the heterohexameric complex was not ascertained. Quantification of the hMcm complex would be possible by using MALDI-MS fingerprinting if there would be equal abundance of all hMcm in final purification step with no degradation or contamination. For these reasons it was decided not to use

this complex for stoichiometric studies that remains an area of speculation as the stoichiometry of hMcm complex has not been measured accurately to date because it has been dependent on indirect method of antibody detection. All six proteins were present in the last chromatographic step as shown by western blot, but the amount of each protein was not equal as the intensity of different bands of each hMcm varied in SDS PAGE Although SDS PAGE and western blots indicated the presence of all six hMcms, after various co-purification steps, but there may be some protein degradation products visible in SDS PAGE. It is quite possible that the purified complex is a mixture of homohexamers and heterohexamers with various combinations as Mcms are isolated in various stable assemblies (Davey et al., 2003).

Another possibility is that they may not be forming heterohexamers and the complex is just a number of different homohexamers. As the protein complex is pull down with his-tag that was present on three of the proteins, that confirms Mcm2-7 protein complex do consists three dimers but does not confirm if the hMcm complex actually makes a heterohexamer containing all six of the hMcms. If the approach of co-purifying the hMcm with S-tag or GST-tag or Txr tag would have followed instead of his-tag, it would have confirmed whether they all make heterohexamers as these tags were present only on one of the hMcms.

In future this complex can be readily characterised further by electron microscopy to provide a three-dimensional reconstruction.

6.2 hMcm helicase

In eukaryotes, *in vivo* observations implicate the Mcm2-7 complex as the replicative helicase (reviewed in Bell and Dutta, 2002) but, yeast Mcm2-7 has been reported to lack *in vitro* helicase activity (Bochman and Schwacha, 2007). Interestingly, both an archaeal Mcm complex (reviewed in Kelman and White, 2005) and hexameric Mcm complex containing only three of the six eukaryotic Mcm subunits (Mcm4/6/7 complex) had DNA-unwinding activity (Ishimi, 1997; Kelman et al., 1999; Chong et al., 2000; Lee and

Hurwitz, 2001; Kaplan et al., 2003). The purified Cdc45/Mcm2-7/GINS complex was shown to associate with an active ATP-dependent DNA helicase function (Moyer et al., 2006). This discrepancy that Mcm2-7 lacks helicase activity was rationalized and Mcm2-7 complex from yeast was shown to have anion dependent helicase activity (Bochman and Schwacha, 2008). Very recently a single subunit of Mcm6 was isolated from pea and reported to form a homohexamer and function as a DNA helicase (Tran et al., 2010). To date no study of human Mcm2-7 complex has shown helicase activity. Purified hMcm2-7 complex was subjected to helicase activity assays in this study. Robust *in vitro* helicase activity of hMcm2-7 complex was confirmed by strand displacement assays. This result of hMcm2-7 complex having helicase activity is consistent with recently published papers (Bochman and Schwacha, 2008). More specifically Bochman and Schwacha demonstrated that chloride ions inhibit DNA unwinding by Mcms whereas glutamate ions support Mcm helicase activity (Bochman and Schwacha, 2008). Although in this study we have shown that the presence of chloride does decrease Mcm helicase activity but does not abolish it completely (See later). The anions present in previously published helicase assays would have blocked Mcm2-7-unwinding activity. This study further confirms that there is no need for any other proteins or cofactors for helicase activity of hMcm2-7 although it has been previously shown that purified Cdc45 and GINS are necessary for Mcm2-7 helicase activity in a complex purified from *Drosophila* (Moyer et al., 2006). Although this study does not preclude the possibility that their presence may increase the helicase activity or even processivity of helicase activity.

To further confirm that hMcm2-7 has *in vitro* helicase activity, it would have been clear if there were any doubt, if the elution fraction that does not contain any hMcm2-7 protein in it (after the hMcm2-7 comes off), would have been assayed for helicase activity in parallel with the elution fraction that does actually contain hMcm2-7. This would have confirmed that nothing in the reaction mixture or elution fraction is driving the helicase activity except hMcm2-7 complex.

Along with the hMcm2-7 complex, hMcm4, hMcm6, hMcm7 could have been individually tested for helicase assays as a single subunit of Mcm6 has been reported to form a homohexamer and function as a DNA helicase (Tran et al., 2010).

6.3 Phosphorylation of hMcm2-7

The hMcm complex was further tested as a target for kinase activity in the presence of CDK2/Cyclin A and CDK2/Cyclin E. Human Mcm2-7 complex showed phosphorylation with both complexes, CDK2/cyclin A and CDK2/cyclin E. CDK2/cyclin E showed more robust activity compared to CDK2/cyclin A. Various previous studies confirmed the phosphorylation of individual Mcms and Mcm2-7 complex by CDK (Ishimi and Komamura-Kohno, 2001; Wheeler et al., 2008). The site(s) of phosphorylation by CDK2/Cyclin A and CDK2/Cyclin E are not fully known. It has been demonstrated recently that Mcm3 is a substrate of cyclin E/Cdk2 and can be phosphorylated by cyclin E/Cdk2 at Thr722 (Li et al., 2011). Cdk2/cyclin A and Cdc2/cyclin E might be phosphorylating at the same or different sites, however it suggests that both Cdk2/cyclin A and Cdk2/cyclin E have similar specificity of substrate recognition, as both of them are phosphorylating the hMcm complex at various sites. Although it seems likely Cdk2/cyclin E has more sites for phosphorylation compared to Cdk2/cyclin A due to the signal strength observed in kinase assays.

Using mass spectrometry, phosphorylation sites of the human Mcm complex could have identified but due to PhD time frame as a limiting factor, samples were not sent for mass spectrometry. This could be done in future to identify known and novel phosphorylated sites on each of the hMcm.

6.4 Helicase activity of phosphorylated hMcm2-7

Phosphorylated hMcm2-7 via CDK2/cyclin E was subjected to helicase activity. The phosphorylated hMcm2-7 did not show any helicase activity, an observation supported by studies of Mcm4/6/7 complex (Ishimi et al.,

2000a; Ishimi et al., 2000b; kudoh et al., 2006). Recently it has been proved that Mcm2-7 helicase loading was inhibited by CDK (Chen and Bell, 2011).

These results suggest that one of the roles of phosphorylation by the cyclin-dependent kinase is to inactivate the DNA helicase activity of the hMcm2-7 complex that is probably involved in DNA replication *in vivo*. Based on these findings, it is considered that phosphorylated regions play a role in the interaction with DNA during DNA helicase action, and the phosphorylation of hMcm2-7 at specific sites by the cyclin-dependent kinase perturbs the interaction to inactivate the DNA helicase activity of the hMcm2-7 complex. This finding proves that Mcm proteins are substrate of the cyclin-dependent kinases.

A few positive controls were missing in the helicase assays of phosphorylated hMcm2-7. It would have been very informative to the results as a whole with more positive controls on it.

A number of sites of phosphorylation are already known on Mcm2, Mcm3 and Mcm4 subunits (Ishimi et al., 2000; Montagnoli et al., 2006; Li et al., 2011) that seem to regulate the function of Mcm complex. Further to this finding it is important to find out all the phosphorylation sites on hMcm2-7 complex and determine which sites are playing a role in the inhibition of hMcm2-7 complex helicase activity. It could further indicate whethet any sites are lethal for hMcm2-7 helicase activity.

Helicase related activity assays could have been performed with the hMcm2-7 complex including ATP hydrolysis, DNA binding assays and processivity assays. ATP hydrolysis is assessed in the presence and absence of closed circular DNA templates using $\alpha[^{32}P]$ATP and then separating the hydrolysis product by thin layer chromatography. DNA binding assays can be carried out using electrophoretic shift assay with an oligonuleotide substrate labelled with $\gamma[^{32}P]$ATP in the presence and absence of ATP. A processivity assay can be performed by using a substrate that covers a range of sizes of duplex DNA. All these above mentioned assays are quantified by

phosphorimaging (Jenkinson and Chong, 2006).

6.5 *M. maripaludis* S2 Mcms

M. maripaludis S2 is unique among other archaea as it contains four Mcm homologues. It has been suggested that multiple Mcms are functional in *M. maripaludis* (Walters and Chong 2010) compared to rest of the archaea where one Mcm is sufficient for replication (McGeoch and Bell, 2008; Bae et al., 2009). All four MMP Mcms were co-purified successfully. MMP McmA and MMP McmC were purified efficiently in a single affinity purification step with his-tag. MMP McmB and MMP McmD were purified by GST affinity purification where a number of lower molecular weight proteins were present. To get rid of low molecular weight proteins, an additional step of anion exchange chromatography was carried out. The western blot was carried out on the elution fractions of each of the Mcms that confirmed the presence of each of the Mcms although the western blots were performed without any controls, which would have more reliably confirmed the results. In this study further investigation of the role of each Mcm in helicase assays was carried out. Helicase assays performed on MMP McmA showed that this protein has robust *in vitro* helicase activity suggesting this Mcm is functional. Helicase assays performed on MMP McmB also demonstrated this Mcm is functional too as this Mcm has also showed protein dependent robust helicase activity *in vitro*. Helicase assays were carried out on MMP McmC and MMP McmD under the same conditions that were used to demonstrate the helicase activities of MMP McmA and MMP McmB, but neither of these Mcms showed any *in vitro* helicase activity. As the concentration of MMP McmD was very low and also it was not as pure as were the other Mcms, a firm conclusion cannot be drawn on that basis and MMP McmD purification needs to be optimized before performing these assays. Although it was very surprising that MMP McmC did not show any helicase activity as it contain all the necessary motif require for helicase activity and it is quite similar to MMP McmA and MMP McmB in sequence. One explanation might be a requirement for other proteins or cofactors for MMP McmC to show helicase activity as shown

previously that Cdc45 and GINS are necessary for *Drosophila* Mcm2-7 DNA helicase function (Moyer et al., 2006).

Furthermore MMP McmA was subjected to helicase activity assays under various conditions to optimize the measured activity. The most interesting finding identified was presence of glutamate increase the helicase activity while the chloride decreases the helicase activity. This result is somewhat consistent with the previous work (Bochman and Schwacha, 2008) where they have shown the presence of anions completely abolishes helicase activity. The data presented in this study clearly show that even in the presence of high salt hMcm2-7 complex still showed some helicase activity. In previously published results post translationally modified Mcm2-7 would have acted differently in the presence of anionic salts as the complex was expressed from yeast (Bochman and Schwacha, 2008) while hMcm2-7 complex used in this study was purified from *E.coli*.

6.6 Potential DNA kinases in archaea

Very little is known about the mechanisms that control DNA replication in archaea. Archaea lack recognizable homologues of cyclin dependent kinases, cyclins, or the kinase Cdc7. MMP0004 and Mj1073 are hypothetical proteins and bioinformatics has assigned them as RIO-type kinases. RIO-type kinases are family of conserved proteins present from archaea to eukarya and their biological substrate is still not known. Structural predictions of MMP0004 and Mj1073 were obtained from different online data servers and the model is based on trying to fit the protein sequence to the most related structure in the database. They all came up with similar structure that was based on the most homologous known protein of *A. fulgidus* RIO2 Kinase in protein database (LaRonde-LeBlanc and Wlodawer 2004). The crystal structure of *A. fulgidus* RIO2 Kinase has been solved and the predicted ribbon structure of MMP0004 and Mj1073 is very similar to *A. fulgidus* RIO2 Kinase structure. Clustal alignment has shown that there are a few conserved residues specific to RIO kinase along with a specific ATP binding pocket and kinase domains in predicted structures. MMP0004 and Mj1073 were

successfully cloned and purified from *E.coli*. The mass spectrometry confirmed the identity of the MMMP0004 protein. Purified MMP0004 was tested as possible kinase acting on MMP McmA as it was identified as a interacting partner of MMP McmA (Chong lab unpublished data). Results from kinase assays showed that there was not any phosphorylation by MMP0004 on MMP McmA. The possible explanation for MMP0004 not acting as a kinase may be due to requirement of other proteins or cofactors in assay. MMP McmA might not be a substrate for MMP0004 or MMP0004 might not be an active kinase though its sequence contains RIO type kinase motifs. As the purification of MMP0004 and Mj1073 has been successful and the purification concentrations of both proteins are very good, both of these proteins can be use in future in studies to ascertain their roles.

Appendix 1

Abbreviations:

A_{260}, A_{280}	Absorbance measured a 260 nm or 280 nm
AAA+	ATPases Associated with various cellular Activities
Amp	Ampicillin
APS	Ammonium persulfate
ARS	Autonomously replicating sequences
ATM	Ataxia Telangiectasia Mutated
ATP	Adenosine triphosphate
ATPase	ATP hydrolysis (activity or domain)
ATR	ATM and Rad3 related
AX	Anion exchange
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
CDC	Cell Division Cycle
Cdt	Cyclin dependent transcript
CDKs	Cyclin Dependent Kinases
CV	Column Volumn
DDKs	Dbf4 Dependent kinases
DIG	Digoxigenin
dNTPs	Deoxynucleoside triphosphates
DTT	Dithiothreitol
DUE	DNA Unwinding Element
e	Expect score
ECL	Enhance Chemiluminescence
EDTA	Ethylenediaminetetraacetic Acid
EF	Elution Fraction
EGTA	Ethylene Glycol Tetraacetic Acid
EM	Electron Microscopy
FPLC	Fast Protein Liquid Chromatography
FRET	Fluorescene resonance energy transfer
GF	Gel Filteration
GINS	Go, Ichi, Nii, San
GST	Glutathione S-Transferase
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HPLC	High Performance Liquid Chromatography
H2i	Helix-2 insert
his-tag	Histidine tag
Hsp	Heat shock protein
HTH	Helix Turn Helix
IAA	Isoamyl Alcohol
IBs	Inclusion Bodies
IPTG	Isopropyl β-D-1-thiogalactopyranoside
iTRAQ	Isobaric Tag for Relative and Absolute Quantitation
K_d	Dissociation constant
KDa	Killo Dalton

LB	Luria-Bertani medium
LC-MS	Liquid Chromatography Mass Spectrometry
LIC	Ligation Independent Cloning
MALDI	Matrix-Assisted Laser Desorption/Ionization
MASCOT	Multiple Alignment System developed by iCOT
MCM	Minichromosme maintenance
MCM3AP	Minichromocome Maintenance Acetylating Protein
MES	2-(*N*-morpholino)ethanesulfonic acid
Mj	*Methanocaldococcus jannaschii*
MMP	*Methanococcus maripaludi* proteins
M-OBP	*Methanococcales* Origin Binding Protein
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
*Mth*Mcm	*Methanothermobacter thermautotrophicus* Mcm protein
NCBI	North Carolina Banking Institute
Ni-NTA	Nickel-Nitrilotriacetic Acid
OBP	Origin Binding Protein
OD_{600}	Optical Density measured at 600 nm
ODNs	Oligodeoxynucleotides
ORB	Origin Recognition Box
ORC	Origin Recognition Complex
ORF	Open Reading Frame
Ori	Origin of DNA replication
OriC	Bacterial origin of replication
PDB	Protein database/Data Bank
PFD	Process Flow Diagram
PEG	Polyethylene Glycol
PCR	Polymerase Chain Reaction
PMSF	Phenylmethylsulphonyl Fluoride
PNK	Polynucleotide Kinase
PS1BH	Pre-sensor 1 beta-hairpin
PVDF	Polyvinylidene Fluoride
R-finger	Arginine finger
RIO	Right Open Reading frame (yeast kinase gene)
RPC	Replisome progression complex
RNAi	RNA interference
RT-PCR	Real-Time Polymerase Chain Reaction
SDM	Site-Directed Mutagenesis
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
S-layer	Surface layer
SOC	Super Optimal Broth “catabolite repression”
Spec	Spectinomycin
SSB	Single-strand binding protein
SSC	Saline Sodium Citrate buffer
Sso Mcm	*Sulfolobus solfataricus* Mcm
SV40	Simian Virus 40

TAE	Tris-Acetate EDTA
Tag	SV40 large tumour antigen
TBE	Tris Boric acid EDTA
TE	Tris EDTA
TEAB	Triethylammonium Bicarbonate
TEMED	Tetramethylethylenediamine
TOF	Time Of Flight
TP	Total Protein
Trx	Thioredoxin
UB	Unbound
UV	Ultra Violet
wHTH	winged Helix Turn Helix
WT	Wild Type
w/v	weight/volume

Appendix 2

List of chapter 3 primers

Table A1: Primer sequences for SDM of hMcm 2-7 genes

Primer Name	Direction of Primer	Sequence 5' →3' of Primer
SDM hMcm2	Forward	GATTCCGCC**ATG**GTATCCTGGTACCCGGG
SDM hMcm2	Reverse	CCCGGGTACCAGGATACCATGGCGGAATC
SDM hMcm3	Forward	CGGTACCCGCC**ATG**AATTTCGGATCCTGGTACC
SDM hMcm3	Reverse	GGTACCAGGATCCGAAATTCATGGCGGGTACCG
SDM hMcm4	Forward	GCCGGGGACGAC**ATG**AAATTCGGATCCTGG
SDM hMcm4	Reverse	CCAGGATCCGAATTTCATGTCGTCCCCGGC
SDM hMcm5	Forward	CGAATCCCGAC**ATG**AATTTCGGATCCTGGTACC
SDM hMcm5	Reverse	GGTACCAGGATCCGAAATTCATGTCGGGATTCG
SDM hMcm6	Forward	CCGCGAGGTCC**ATG**AATTTCGGATCCTGGTACC
SDM hMcm6	Reverse	GGTACCAGGATCCGAAATTCATGGACCTCGCGG
SDM hMcm7	Forward	CCTTCAGTGCC**ATG**AATTTCGGATCCTGGTACC
SDM hMcm7	Reverse	GGTACCAGGATCCGAAATTCATGGCAC**TGA**AGG

Bold letters indicate start and stop codons.

Table A2: Primer sequences for LIC of hMcm 2-7 genes

Primer Name	Direction of Primer	Sequence 5' →3' of Primer
LIC hMcm2	Forward	GCGGGCCCGGCCTCC**ATG**GCGGAATCAT CGGAATCCTTCACC
LIC hMcm2	Reverse	GAGGAGAAGCCCGGTCAGAACTGCTGCA GGATCATTTTCC
LIC hMcm3	Forward	GCGGGCCCGGCCTTC**ATG**GCGGGTACCG TGGTGCTGGAC
LIC hMcm3	Reverse	GAGGAGAAGCCCGGTCAGATGAGGAAGA **TGA**TGCCCTCAG
LIC hMcm4	Forward	GCGGGCCCGGCCTTC**ATG**TCGTCCCCGGC GTCGACCCC
LIC hMcm4	Reverse	GAGGAGAAGCCCGGTCAGAGCAAGCGCA CGGTCTTCCC
LIC hMcm5	Forward	GACGACGACAAG**ATG**TCGGGATTCGACG ATCCTGGC
LIC hMcm5	Reverse	CGCGGGCGGCCGTCACTTGAGGCGG**TAG** AGAACCTTGC
LIC hMcm6	Forward	GACGACGACAAG**ATG**GACCTCGCGGCGG CAGCGG
LIC hMcm6	Reverse	CGCGGGCGGCCGTCAATCTTCGAGCAAG **TAG**TTAGGG
LIC hMcm7	Forward	GACGACGACAAG**ATG**GCACTGAAGGACT ACGCGCTAG
LIC hMcm7	Reverse	CGCGGGCGGCCGTCAGACAAAAG**TGA**TC CGTGTCCGGG

Bold letters indicate start and stop codons.

Table A3: Primer sequence of helicase substrate

Primer Name	Sequence 5' →3' of Primer
HS2	$(TTTG)_8$CCGACGTGCCAGGCCGACGCGTCCC
HS1	GGGACGCGTCGGCCTGGCACGTCGGCCGCTGCGGCCA GGCACCCGATGGC$(GTTT)_6$

List of chapter 5 Primers

Table A4: Primer sequences for cloning of MMP0004 and Mj1073

Primer Name	Direction of Primer	Sequence 5' →3' of Primer
MMP0004	Forward	TTCAATGGATCCAA**ATG**GAAGATAACG ACTGGAAATTG
MMP0004	Reverse	AAACTCGAGTTAAATTTCGC**TAG**TTA TATCTCCAATAAC
Mj1073	Forward	AGAGGATCCAA**ATG**AGACATCATGAG TG
Mj1073	Reverse	CCACTCGAGTTATT**TAG**TTATATACTC AAAGATTTT

Bold letters indicate start and stop codons, while underline letters indicate restrictions sites.

References:

Alberts, B.; A. Johnson, et al. (2002). "Molecular Biology of the Cell". Chapter 5; DNA Replication Mechanisms Garland Science. ISBN 0-8153-3218-1

Adachi, Y., J. Usukura, et al. (1997). "A globular complex formation by Nda1 and the other five members of the MCM protein family in fission yeast." Genes Cells **2**(7): 467-479.

Alison, M. R., T. Hunt, et al. (2002). "Minichromosome maintenance (MCM) proteins may be pre-cancer markers." Gut **50**(3): 290-291.

Allers, T. and M. Mevarech (2005). "Archaeal genetics - the third way." Nat Rev Genet **6**(1): 58-73.

Angermayr, M., G. Strobel, et al. (2000). "Stable plasma membrane expression of the soluble domain of the human insulin receptor in yeast." FEBS Lett **481**(1): 8-12.

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): 9130-9135.

Aparicio, O. M., D. M. Weinstein, et al. (1997). "Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase." Cell **91**(1): 59-69.

Aravind, L. and E. V. Koonin (1999). "DNA-binding proteins and evolution of transcription regulation in the archaea." Nucleic Acids Res **27**(23): 4658-4670.

Ariizumi, K., M. R. Ghosh, et al. (1993). "Elements in the immunoglobulin heavy-chain enhancer directly regulate simian virus 40 ori-dependent DNA replication." Mol Cell Biol **13**(9): 5629-5636.

Ariizumi, K., Z. Wang, et al. (1993). "Immunoglobulin heavy chain enhancer is located near or in an initiation zone of chromosomal DNA replication." Proc Natl Acad Sci U S A **90**(8): 3695-3699.

Arnold, K., L. Bordoli, et al. (2006). "The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling." Bioinformatics **22**(2): 195-201.

Atanassova, N. and I. Grainge (2008). "Biochemical characterization of the minichromosome maintenance (MCM) protein of the crenarchaeote Aeropyrum pernix and its interactions with the origin recognition complex (ORC) proteins." Biochemistry **47**(50): 13362-13370.

Bae, B., Y. H. Chen, et al. (2009). "Insights into the architecture of the replicative helicase from the structure of an archaeal MCM homolog." Structure **17**(2): 211-222.

Barry, E. R. and S. D. Bell (2006). "DNA replication in the archaea." Microbiol Mol Biol Rev **70**(4): 876-887.

Barry, E. R., A. T. McGeoch, et al. (2007). "Archaeal MCM has separable processivity, substrate choice and helicase domains." Nucleic Acids Res **35**(3): 988-998.

Bell, S. P. and A. Dutta (2002). "DNA replication in eukaryotic cells." Annu Rev Biochem **71**: 333-374.

Bell, S. P., R. Kobayashi, et al. (1993). "Yeast origin recognition complex functions in transcription silencing and DNA replication." Science **262**(5141): 1844-1849.

Bell, S. P., Y. Marahrens, et al. (1993). "The replicon model and eukaryotic chromosomes." Cold Spring Harb Symp Quant Biol **58**: 435-442.

Bell, S. P., J. Mitchell, et al. (1995). "The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing." Cell **83**(4): 563-568.

Bell, S. P. and B. Stillman (1992). "ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex." Nature **357**(6374): 128-134.

Berg, J. M.; J. L. Tymoczko, et al. (2002). Biochemistry. Chapter 27, Section 4: DNA Replication of Both Strands Proceeds Rapidly from Specific Start Sites W.H. Freeman and Company ISBN 0-7167-3051-0

Blow, J. J. and R. A. Laskey (1988). "A role for the nuclear envelope in controlling DNA replication within the cell cycle." Nature **332**(6164): 546-548.

Bochman, M. L., S. P. Bell, et al. (2008). "Subunit organization of Mcm2-7 and the unequal role of active sites in ATP hydrolysis and viability." Mol Cell Biol **28**(19): 5865-5873.

Bochman, M. L. and A. Schwacha (2007). "Differences in the single-stranded DNA binding activities of MCM2-7 and MCM467: MCM2 and MCM5 define a slow ATP-dependent step." J Biol Chem **282**(46): 33795-33804.

Bochman, M. L. and A. Schwacha (2008). "The Mcm2-7 complex has in vitro helicase activity." Mol Cell **31**(2): 287-293.

Bochman, M. L. and A. Schwacha (2010). "The Saccharomyces cerevisiae Mcm6/2 and Mcm5/3 ATPase active sites contribute to the function of the putative Mcm2-7 'gate'." Nucleic Acids Res **38**(18): 6078-6088.

Bossemeyer, D. (1995). "Protein kinases--structure and function." FEBS Lett **369**(1): 57-61.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem **72**: 248-254.

Brewer, B. J. and W. L. Fangman (1991). "Mapping replication origins in yeast chromosomes." Bioessays **13**(7): 317-322.

Brinkmann, U., R. E. Mattes, et al. (1989). "High-level expression of recombinant genes in Escherichia coli is dependent on the availability of the dnaY gene product." Gene **85**(1): 109-114.

Bult, C. J., O. White, et al. (1996). "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii." Science **273**(5278): 1058-1073.

Burkhart, R., D. Schulte, et al. (1995). "Interactions of human nuclear proteins P1Mcm3 and P1Cdc46." Eur J Biochem **228**(2): 431-438.

Campbell, M. G. and K. Karbstein (2011). "Protein-protein interactions within late pre-40S ribosomes." PLoS One **6**(1): e16194.

Carpentieri, F., M. De Felice, et al. (2002). "Physical and functional interaction between the mini-chromosome maintenance-like DNA helicase and the single-stranded DNA binding protein from the crenarchaeon Sulfolobus solfataricus." J Biol Chem **277**(14): 12118-12127.

Carr, K. M. and J. M. Kaguni (2001). "Stoichiometry of DnaA and DnaB protein in initiation at the Escherichia coli chromosomal origin." J Biol Chem **276**(48): 44919-44925.

Chen, S. and S. P. Bell (2011). "CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6." Genes Dev **25**(4): 363-372.

Chen, Y., K. M. Hennessy, et al. (1992). "CDC46/MCM5, a yeast protein whose subcellular localization is cell cycle-regulated, is involved in DNA replication at autonomously replicating sequences." Proc Natl Acad Sci U S A **89**(21): 10459-10463.

Chen, Y., A. Sittler, et al. (1998). "[Screening of proteins interact with FMR1 by yeast two-hybrid system]." Zhongguo Yi Xue Ke Xue Yuan Xue Bao **20**(3): 173-178.

Chen, Y. J., X. Yu, et al. (2005). "Structural polymorphism of Methanothermobacter thermautotrophicus MCM." J Mol Biol **346**(2): 389-394.

Chick, H. and C. J. Martin (1913). "The Precipitation of Egg-Albumin by Ammonium Sulphate. A Contribution to the Theory of the "Salting-out" of Proteins." Biochem J **7**(4): 380-398.

Chong, J. P. (2005). "Learning to unwind." Nat Struct Mol Biol **12**(9): 734-736.

Chong, J. P., M. K. Hayashi, et al. (2000). "A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase." Proc Natl Acad Sci U S A **97**(4): 1530-1535.

Chong, J. P., H. M. Mahbubani, et al. (1995). "Purification of an MCM-containing complex as a component of the DNA replication licensing system." Nature **375**(6530): 418-421.

Chong, J. P., P. Thommes, et al. (1996). "The role of MCM/P1 proteins in the licensing of DNA replication." Trends Biochem Sci **21**(3): 102-106.

Coleman, L. W., I. B. Bronstein, et al. (2001). "Immunohistochemical staining for DNA topoisomerase I, DNA topoisomerase II-alpha and p53 in gastric carcinomas." Anticancer Res **21**(2A): 1167-1172.

Cortez, D., G. Glick, et al. (2004). "Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases." Proc Natl Acad Sci U S A **101**(27): 10078-10083.

Costa, A., T. Pape, et al. (2006). "Structural basis of the Methanothermobacter thermautotrophicus MCM helicase activity." Nucleic Acids Res **34**(20): 5829-5838.

Costa, A., T. Pape, et al. (2006). "Structural studies of the archaeal MCM complex in different functional states." J Struct Biol **156**(1): 210-219.

Costa, M., A. Ochem, et al. (1999). "Human DNA helicase VIII: a DNA and RNA helicase corresponding to the G3BP protein, an element of the ras transduction pathway." Nucleic Acids Res **27**(3): 817-821.

Coue, M., S. E. Kearsey, et al. (1996). "Chromotin binding, nuclear localization and phosphorylation of Xenopus cdc21 are cell-cycle dependent and associated with the control of initiation of DNA replication." EMBO J **15**(5): 1085-1097.

Coxon, A., K. Maundrell, et al. (1992). "Fission yeast cdc21+ belongs to a family of proteins involved in an early step of chromosome replication." Nucleic Acids Res **20**(21): 5571-5577.

Crevel, G., A. Ivetic, et al. (2001). "Nearest neighbour analysis of MCM protein complexes in Drosophila melanogaster." Nucleic Acids Res **29**(23): 4834-4842.

Dahmann, C., J. F. Diffley, et al. (1995). "S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state." Curr Biol **5**(11): 1257-1269.

Damodaran, S. (1986). "Kinetics of formation of hydrophobic regions during refolding of bovine serum albumin." Int J Pept Protein Res **27**(6): 589-596.

Davey, M. J., C. Indiani, et al. (2003). "Reconstitution of the Mcm2-7p heterohexamer, subunit arrangement, and ATP site architecture." J Biol Chem **278**(7): 4491-4499.

Davey, M. J., D. Jeruzalmi, et al. (2002). "Motors and switches: AAA+ machines within the replisome." Nat Rev Mol Cell Biol **3**(11): 826-835.

De Felice, M., L. Esposito, et al. (2003). "Biochemical characterization of a CDC6-like protein from the crenarchaeon Sulfolobus solfataricus." J Biol Chem **278**(47): 46424-46431.

de Fougerolles, A. R. (2008). "Delivery vehicles for small interfering RNA in vivo." Hum Gene Ther **19**(2): 125-132.

Diffley, J. F., (2004). " Regulation of early events in chromosome replication." Curr Biol **14**(18):778-786.

Diffley, J. F. and J. H. Cocker (1992). "Protein-DNA interactions at a yeast replication origin." Nature **357**(6374): 169-172.

Diffley, J. F., J. H. Cocker, et al. (1994). "Two steps in the assembly of complexes at yeast replication origins in vivo." Cell **78**(2): 303-316.

Donovan, S., J. Harwood, et al. (1997). "Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast." Proc Natl Acad Sci U S A **94**(11): 5611-5616.

Dubendorff, J. W. and F. W. Studier (1991). "Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor." J Mol Biol **219**(1): 45-59.

Dubendorff, J. W. and F. W. Studier (1991). "Creation of a T7 autogene. Cloning and expression of the gene for bacteriophage T7 RNA polymerase under control of its cognate promoter." J Mol Biol **219**(1): 61-68.

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

Echols, H. (1990). "Nucleoprotein structures initiating DNA replication, transcription, and site-specific recombination." J Biol Chem **265**(25): 14697-14700.

Edelman, A. M., D. K. Blumenthal, et al. (1987). "Protein serine/threonine kinases." Annu Rev Biochem **56**: 567-613.

Edgell, D. R. and W. F. Doolittle (1997). "Archaea and the origin(s) of DNA replication proteins." Cell **89**(7): 995-998.

Endl, E., I. Kausch, et al. (2001). "The expression of Ki-67, MCM3, and p27 defines distinct subsets of proliferating, resting, and differentiated cells." J Pathol **195**(4): 457-462.

Enemark, E. J. and L. Joshua-Tor (2008). "On helicases and other motor proteins." Curr Opin Struct Biol **18**(2): 243-257.

Engh, R. A. and D. Bossemeyer (2002). "Structural aspects of protein kinase control-role of conformational flexibility." Pharmacol Ther **93**(2-3): 99-111.

Fang, L., M. J. Davey, et al. (1999). "Replisome assembly at oriC, the replication origin of E. coli, reveals an explanation for initiation sites outside an origin." Mol Cell **4**(4): 541-553.

Feger, G., H. Vaessin, et al. (1995). "dpa, a member of the MCM family, is required for mitotic DNA replication but not endoreplication in Drosophila." EMBO J **14**(21): 5387-5398.

Fernandez-Fuentes, N., C. J. Madrid-Aliste, et al. (2007). "M4T: a comparative protein structure modeling server." Nucleic Acids Res **35**(Web Server issue): W363-368.

Fletcher, R. J., B. E. Bishop, et al. (2003). "The structure and function of MCM from archaeal M. Thermoautotrophicum." Nat Struct Biol **10**(3): 160-167.

Forsburg, S. L. (2004). "Eukaryotic MCM proteins: beyond replication initiation." Microbiol Mol Biol Rev **68**(1): 109-131.

Forsburg, S. L., D. A. Sherman, et al. (1997). "Mutational analysis of Cdc19p, a Schizosaccharomyces pombe MCM protein." Genetics **147**(3): 1025-1041.

Foss, M., F. J. McNally, et al. (1993). "Origin recognition complex (ORC) in transcriptional silencing and DNA replication in S. cerevisiae." Science **262**(5141): 1838-1844.

Freeman, A., L. S. Morris, et al. (1999). "Minichromosome maintenance proteins as biological markers of dysplasia and malignancy." Clin Cancer Res **5**(8): 2121-2132.

Frick, D. N. and C. C. Richardson (2001). "DNA primases." Annu Rev Biochem **70**: 39-80.

Fujita, M., T. Kiyono, et al. (1997). "In vivo interaction of human MCM heterohexameric complexes with chromatin. Possible involvement of ATP." J Biol Chem **272**(16): 10928-10935.

Gajiwala, K. S. and S. K. Burley (2000). "Winged helix proteins." Curr Opin Struct Biol **10**(1): 110-116.

Gajiwala, K. S., H. Chen, et al. (2000). "Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding." Nature **403**(6772): 916-921.

Galperin, M. Y. (2005). "The Molecular Biology Database Collection: 2005 update." Nucleic Acids Res **33**(Database issue): D5-24.

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): 358-366.

Gibson, S. I., R. T. Surosky, et al. (1990). "The phenotype of the minichromosome maintenance mutant mcm3 is characteristic of mutants defective in DNA replication." Mol Cell Biol **10**(11): 5707-5720.

Going, J. J., W. N. Keith, et al. (2002). "Aberrant expression of minichromosome maintenance proteins 2 and 5, and Ki-67 in dysplastic squamous oesophageal epithelium and Barrett's mucosa." Gut **50**(3): 373-377.

Gomez-Llorente, Y., R. J. Fletcher, et al. (2005). "Polymorphism and double hexamer structure in the archaeal minichromosome maintenance (MCM) helicase from Methanobacterium thermoautotrophicum." J Biol Chem **280**(49): 40909-40915.

Gorbalenya, A. E., E. V. Koonin, et al. (1988). "A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are

probably involved in duplex unwinding in DNA and RNA replication and recombination." FEBS Lett **235**(1-2): 16-24.

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): 570-579.

Grabowski, B. and Z. Kelman (2001). "Autophosphorylation of archaeal Cdc6 homologues is regulated by DNA." J Bacteriol **183**(18): 5459-5464.

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

Grainge, I., S. Scaife, et al. (2003). "Biochemical analysis of components of the pre-replication complex of Archaeoglobus fulgidus." Nucleic Acids Res **31**(16): 4888-4898.

Grallert, B. and P. Nurse (1996). "The ORC1 homolog orp1 in fission yeast plays a key role in regulating onset of S phase." Genes Dev **10**(20): 2644-2654.

Guex, N. and M. C. Peitsch (1997). "SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling." Electrophoresis **18**(15): 2714-2723.

Hall, M. C. and S. W. Matson (1999). "Helicase motifs: the engine that powers DNA unwinding." Mol Microbiol **34**(5): 867-877.

Hamlin, J. L., L. D. Mesner, et al. (2008). "A revisionist replicon model for higher eukaryotic genomes." J Cell Biochem **105**(2): 321-329.

Hardy, C. F. (1997). "Identification of Cdc45p, an essential factor required for DNA replication." Gene **187**(2): 239-246.

Haugland, G. T., M. Innselset, et al. (2008). "Characterization of the Cdc6 Homologues from the Euryarchaeon Thermoplasma acidophilum." Open Biochem J **2**: 129-134.

Haugland, G. T., C. R. Rollor, et al. (2009). "Biochemical characterization of the minichromosome maintenance protein from the archaeon Thermoplasma acidophilum." Extremophiles **13**(1): 81-88.

Haugland, G. T., J. H. Shin, et al. (2006). "Stimulation of MCM helicase activity by a Cdc6 protein in the archaeon Thermoplasma acidophilum." Nucleic Acids Res **34**(21): 6337-6344.

Hendrickson, E. L., R. Kaul, et al. (2004). "Complete genome sequence of the genetically tractable hydrogenotrophic methanogen Methanococcus maripaludis." J Bacteriol **186**(20): 6956-6969.

Hendrickson, M., M. Madine, et al. (1996). "Phosphorylation of MCM4 by cdc2 protein kinase inhibits the activity of the minichromosome maintenance complex." Proc Natl Acad Sci U S A **93**(22): 12223-12228.

Hennessy, K. M., A. Lee, et al. (1991). "A group of interacting yeast DNA replication genes." Genes Dev **5**(6): 958-969.

Hiraiwa, A., M. Fujita, et al. (1998). "Specific distribution patterns of hCDC47 expression in cutaneous diseases." J Cutan Pathol **25**(6): 285-290.

Hiraiwa, A., M. Fujita, et al. (1997). "Immunolocalization of hCDC47 protein in normal and neoplastic human tissues and its relation to growth." Int J Cancer **74**(2): 180-184.

Holthoff, H. P., M. Baack, et al. (1998). "Human protein MCM6 on HeLa cell chromatin." J Biol Chem **273**(13): 7320-7325.

Hopwood, B. and S. Dalton (1996). "Cdc45p assembles into a complex with Cdc46p/Mcm5p, is required for minichromosome maintenance, and is essential for chromosomal DNA replication." Proc Natl Acad Sci U S A **93**(22): 12309-12314.

Hu, Y., J. Kazenwadel, et al. (1993). "Isolation and characterization of the murine homeobox gene Cdx-1. Regulation of expression in intestinal epithelial cells." J Biol Chem **268**(36): 27214-27225.

Hua, X. H., H. Yan, et al. (1997). "A role for Cdk2 kinase in negatively regulating DNA replication during S phase of the cell cycle." J Cell Biol **137**(1): 183-192.

Huberman, J. A., L. D. Spotila, et al. (1987). "The in vivo replication origin of the yeast 2 microns plasmid." Cell **51**(3): 473-481.

Hunt, D. P., A. Freeman, et al. (2002). "Early recurrence of benign meningioma correlates with expression of mini-chromosome maintenance-2 protein." Br J Neurosurg **16**(1): 10-15.

Ibarra, A., E. Schwob, et al. (2008). "Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication." Proc Natl Acad Sci U S A **105**(26): 8956-8961.

Ishimi, Y. (1997). "A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex." J Biol Chem **272**(39): 24508-24513.

Ishimi, Y., S. Ichinose, et al. (1996). "Binding of human minichromosome maintenance proteins with histone H3." J Biol Chem **271**(39): 24115-24122.

Ishimi, Y. and Y. Komamura-Kohno (2001). "Phosphorylation of Mcm4 at specific sites by cyclin-dependent kinase leads to loss of Mcm4,6,7 helicase activity." J Biol Chem **276**(37): 34428-34433.

Ishimi, Y., Y. Komamura-Kohno, et al. (2000). "Inhibition of Mcm4,6,7 helicase activity by phosphorylation with cyclin A/Cdk2." J Biol Chem **275**(21): 16235-16241.

Ishimi, Y., I. Okayasu, et al. (2003). "Enhanced expression of Mcm proteins in cancer cells derived from uterine cervix." Eur J Biochem **270**(6): 1089-1101.

Iyer, L. M., D. D. Leipe, et al. (2004). "Evolutionary history and higher order classification of AAA+ ATPases." J Struct Biol **146**(1-2): 11-31.

Jacob, F. and S. Brenner (1963). "[On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon]." C R Hebd Seances Acad Sci **256**: 298-300.

Jenkinson, E. R. and J. P. Chong (2006). "Minichromosome maintenance helicase activity is controlled by N- and C-terminal motifs and requires the ATPase domain helix-2 insert." Proc Natl Acad Sci U S A **103**(20): 7613-7618.

Jenkinson, E. R., A. Costa, et al. (2009). "Mutations in subdomain B of the minichromosome maintenance (MCM) helicase affect DNA binding and modulate conformational transitions." J Biol Chem **284**(9): 5654-5661.

Kane, J. F. (1995). "Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli." Curr Opin Biotechnol **6**(5): 494-500.

Kanemaki, M. and K. Labib (2006). "Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks." EMBO J **25**(8): 1753-1763.

Kanemaki, M., A. Sanchez-Diaz, et al. (2003). "Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo." Nature **423**(6941): 720-724.

Kanter, D. M., I. Bruck, et al. (2008). "Mcm subunits can assemble into two different active unwinding complexes." J Biol Chem **283**(45): 31172-31182.

Kaplan, D. L., M. J. Davey, et al. (2003). "Mcm4,6,7 uses a "pump in ring" mechanism to unwind DNA by steric exclusion and actively translocate along a duplex." J Biol Chem **278**(49): 49171-49182.

Kasiviswanathan, R., J. H. Shin, et al. (2005). "Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties." Nucleic Acids Res **33**(15): 4940-4950.

Kasiviswanathan, R., J. H. Shin, et al. (2004). "Biochemical characterization of the Methanothermobacter thermautotrophicus minichromosome maintenance (MCM) helicase N-terminal domains." J Biol Chem **279**(27): 28358-28366.

Kearsey, S. E. and K. Labib (1998). "MCM proteins: evolution, properties, and role in DNA replication." Biochim Biophys Acta **1398**(2): 113-136.

Kearsey, S. E., D. Maiorano, et al. (1996). "The role of MCM proteins in the cell cycle control of genome duplication." Bioessays **18**(3): 183-190.

Kelman, L. M. and Z. Kelman (2003). "Archaea: an archetype for replication initiation studies?" Mol Microbiol **48**(3): 605-615.

Kelman, Z., J. K. Lee, et al. (1999). "The single minichromosome maintenance protein of Methanobacterium thermoautotrophicum DeltaH contains DNA helicase activity." Proc Natl Acad Sci U S A **96**(26): 14783-14788.

Kelman, Z., S. Pietrokovski, et al. (1999). "Isolation and characterization of a split B-type DNA polymerase from the archaeon Methanobacterium thermoautotrophicum deltaH." J Biol Chem **274**(40): 28751-28761.

Kelman, Z. and M. F. White (2005). "Archaeal DNA replication and repair." Curr Opin Microbiol **8**(6): 669-676.

Kimura, H., N. Nozaki, et al. (1994). "DNA polymerase alpha associated protein P1, a murine homolog of yeast MCM3, changes its intranuclear distribution during the DNA synthetic period." EMBO J **13**(18): 4311-4320.

Kimura, H., T. Ohtomo, et al. (1996). "Mouse MCM proteins: complex formation and transportation to the nucleus." Genes Cells **1**(11): 977-993.

Knighton, D. R., J. H. Zheng, et al. (1991). "Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase." Science **253**(5018): 407-414.

Knighton, D. R., J. H. Zheng, et al. (1991). "Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase." Science **253**(5018): 414-420.

Koonin, E. V. (1993). "A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication." Nucleic Acids Res **21**(11): 2541-2547.

Krause, M. and W. Messer (1999). "DnaA proteins of Escherichia coli and Bacillus subtilis: coordinate actions with single-stranded DNA-binding protein and interspecies inhibition during open complex formation at the replication origins." Gene **228**(1-2): 123-132.

Kubota, Y., S. Mimura, et al. (1995). "Identification of the yeast MCM3-related protein as a component of Xenopus DNA replication licensing factor." Cell **81**(4): 601-609.

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): 1141-1152.

Kudoh, A., T. Daikoku, et al. (2006). "Phosphorylation of MCM4 at sites inactivating DNA helicase activity of the MCM4-MCM6-MCM7 complex during Epstein-Barr virus productive replication." J Virol **80**(20): 10064-10072.

Kurland, C. and J. Gallant (1996). "Errors of heterologous protein expression." Curr Opin Biotechnol **7**(5): 489-493.

Kurochkina, L. P. and V. V. Mesyanzhinov (1999). "Co-expression of gene 31 and 23 products of bacteriophage T4." Biochemistry (Mosc) **64**(4): 379-383.

Labib, K. and A. Gambus (2007). "A key role for the GINS complex at DNA replication forks." Trends Cell Biol **17**(6): 271-278.

Labib, K., S. E. Kearsey, et al. (2001). "MCM2-7 proteins are essential components of prereplicative complexes that accumulate cooperatively in the nucleus during G1-phase and are required to establish, but not maintain, the S-phase checkpoint." Mol Biol Cell **12**(11): 3658-3667.

Labib, K., J. A. Tercero, et al. (2000). "Uninterrupted MCM2-7 function required for DNA replication fork progression." Science **288**(5471): 1643-1647.

LaRonde-LeBlanc, N. and A. Wlodawer (2004). "Crystal structure of A. fulgidus Rio2 defines a new family of serine protein kinases." Structure **12**(9): 1585-1594.

LaRonde-LeBlanc, N. and A. Wlodawer (2005). "A family portrait of the RIO kinases." J Biol Chem **280**(45): 37297-37300.

LaRonde-LeBlanc, N. and A. Wlodawer (2005). "The RIO kinases: an atypical protein kinase family required for ribosome biogenesis and cell cycle progression." Biochim Biophys Acta **1754**(1-2): 14-24.

Leatherwood, J. (1998). "Emerging mechanisms of eukaryotic DNA replication initiation." Curr Opin Cell Biol **10**(6): 742-748.

Lee, J. K. and J. Hurwitz (2000). "Isolation and characterization of various complexes of the minichromosome maintenance proteins of Schizosaccharomyces pombe." J Biol Chem **275**(25): 18871-18878.

Lei, M. (2005). "The MCM complex: its role in DNA replication and implications for cancer therapy." Curr Cancer Drug Targets **5**(5): 365-380.

Lei, M., Y. Kawasaki, et al. (1996). "Physical interactions among Mcm proteins and effects of Mcm dosage on DNA replication in Saccharomyces cerevisiae." Mol Cell Biol **16**(9): 5081-5090.

Lei, M. and B. K. Tye (2001). "Initiating DNA synthesis: from recruiting to activating the MCM complex." J Cell Sci **114**(Pt 8): 1447-1454.

Li, J., M. Deng, et al. (2011). "Phosphorylation of MCM3 by cycline/CDK2 regulates its function in cell cycle." J Biol Chem.

Li, J. J. and I. Herskowitz (1993). "Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system." Science **262**(5141): 1870-1874.

Liu, J., C. L. Smith, et al. (2000). "Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control." Mol Cell **6**(3): 637-648.

Liu, W., B. Pucci, et al. (2008). "Structural analysis of the Sulfolobus solfataricus MCM protein N-terminal domain." Nucleic Acids Res **36**(10): 3235-3243.

Loo, S., C. A. Fox, et al. (1995). "The origin recognition complex in silencing, cell cycle progression, and DNA replication." Mol Biol Cell **6**(6): 741-756.

Lund, P. (1994). "The chaperonin cycle and protein folding." Bioessays **16**(4): 229-231.

Lundgren, M., A. Andersson, et al. (2004). "Three replication origins in Sulfolobus species: synchronous initiation of chromosome replication and asynchronous termination." Proc Natl Acad Sci U S A **101**(18): 7046-7051.

Luo, J. H. (2011). "Oncogenic activity of MCM7 transforming cluster." World J Clin Oncol **2**(2): 120-124.

Lutzmann, M., D. Maiorano, et al. (2005). "Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2-8 protein family." Gene **362**: 51-56.

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

MacNeill, S. A. (2010). "Structure and function of the GINS complex, a key component of the eukaryotic replisome." Biochem J **425**(3): 489-500.

Madine, M. A., M. Swietlik, et al. (2000). "The roles of the MCM, ORC, and Cdc6 proteins in determining the replication competence of chromatin in quiescent cells." J Struct Biol **129**(2-3): 198-210.

Mahbubani, H. M., J. P. Chong, et al. (1997). "Cell cycle regulation of the replication licensing system: involvement of a Cdk-dependent inhibitor." J Cell Biol **136**(1): 125-135.

Maine, G. T., P. Sinha, et al. (1984). "Mutants of S. cerevisiae defective in the maintenance of minichromosomes." Genetics **106**(3): 365-385.

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): 315-328.

Maiorano, D., M. Lutzmann, et al. (2006). "MCM proteins and DNA replication." Curr Opin Cell Biol **18**(2): 130-136.

Maiorano, D. and M. Mechali (2002). "Many roads lead to the origin." Nat Cell Biol **4**(3): E58-59.

Maiti, A. K. and P. Sinha (1992). "The mcm2 mutation of yeast affects replication, rather than segregation or amplification of the two micron plasmid." J Mol Biol **224**(3): 545-558.

Majernik, A. I., E. R. Jenkinson, et al. (2004). "DNA replication in thermophiles." Biochem Soc Trans **32**(Pt 2): 236-239.

Makarova, K. S., E. V. Koonin, et al. (2012). "The CMG (CDC45/RecJ, MCM, GINS) complex is a conserved component of the DNA replication system in all archaea and eukaryotes." Biol Direct **7**: 7.

Manning, G., G. D. Plowman, et al. (2002). "Evolution of protein kinase signaling from yeast to man." Trends Biochem Sci **27**(10): 514-520.

Manning, G., D. B. Whyte, et al. (2002). "The protein kinase complement of the human genome." Science **298**(5600): 1912-1934.

Marinsek, N., E. R. Barry, et al. (2006). "GINS, a central nexus in the archaeal DNA replication fork." EMBO Rep **7**(5): 539-545.

Masai, H., E. Matsui, et al. (2000). "Human Cdc7-related kinase complex. In vitro phosphorylation of MCM by concerted actions of Cdks and Cdc7 and that of a criticial threonine residue of Cdc7 bY Cdks." J Biol Chem **275**(37): 29042-29052.

McGeoch, A. T. and S. D. Bell (2008). "Extra-chromosomal elements and the evolution of cellular DNA replication machineries." Nat Rev Mol Cell Biol **9**(7): 569-574.

McGeoch, A. T., M. A. Trakselis, et al. (2005). "Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism." Nat Struct Mol Biol **12**(9): 756-762.

Meng, M. V., G. D. Grossfeld, et al. (2001). "Minichromosome maintenance protein 2 expression in prostate: characterization and association with outcome after therapy for cancer." Clin Cancer Res **7**(9): 2712-2718.

Micklem, G., A. Rowley, et al. (1993). "Yeast origin recognition complex is involved in DNA replication and transcriptional silencing." Nature **366**(6450): 87-89.

Mincheva, A., I. Karag'ozov, et al. (1992). "[The demonstration of HPV DNA in cervical intraepithelial neoplasms and in invasive carcinomas of the cervix uteri by DNA-DNA in situ hybridization]." Akush Ginekol (Sofiia) **31**(1): 20-22.

Moffatt, B. A. and F. W. Studier (1987). "T7 lysozyme inhibits transcription by T7 RNA polymerase." Cell **49**(2): 221-227.

Moir, D., S. E. Stewart, et al. (1982). "Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies." Genetics **100**(4): 547-563.

Montagnoli, A., B. Valsasina, et al. (2006). "Identification of Mcm2 phosphorylation sites by S-phase-regulating kinases." J Biol Chem **281**(15): 10281-10290.

Moreau, M. J., A. T. McGeoch, et al. (2007). "ATPase site architecture and helicase mechanism of an archaeal MCM." Mol Cell **28**(2): 304-314.

Morgan, D. O. (1995). "Principles of CDK regulation." Nature **374**(6518): 131-134.

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): 10236-10241.

Murzin, A. G. (1993). "OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences." EMBO J **12**(3): 861-867.

Musahl, C., D. Schulte, et al. (1995). "A human homologue of the yeast replication protein Cdc21. Interactions with other Mcm proteins." Eur J Biochem **230**(3): 1096-1101.

Myllykallio, H. and P. Forterre (2000). "Mapping of a chromosome replication origin in an archaeon: response." Trends Microbiol **8**(12): 537-539.

Myllykallio, H., P. Lopez, et al. (2000). "Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon." Science **288**(5474): 2212-2215.

Nigg, E. A. (1995). "Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle." Bioessays **17**(6): 471-480.

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): 581-587.

Pape, T., H. Meka, et al. (2003). "Hexameric ring structure of the full-length archaeal MCM protein complex." EMBO Rep **4**(11): 1079-1083.

Peng, J. and J. Xu (2011). "RaptorX: exploiting structure information for protein alignment by statistical inference." Proteins **79 Suppl 10**: 161-171.

Piatti, S. (1997). "Cell cycle regulation of S phase entry in Saccharomyces cerevisiae." Prog Cell Cycle Res **3**: 143-156.

Poplawski, A., B. Grabowski, et al. (2001). "The zinc finger domain of the archaeal minichromosome maintenance protein is required for helicase activity." J Biol Chem **276**(52): 49371-49377.

Prokhorova, T. A. and J. J. Blow (2000). "Sequential MCM/P1 subcomplex assembly is required to form a heterohexamer with replication licensing activity." J Biol Chem **275**(4): 2491-2498.

Ramnath, N., F. J. Hernandez, et al. (2001). "MCM2 is an independent predictor of survival in patients with non-small-cell lung cancer." J Clin Oncol **19**(22): 4259-4266.

Richter, A. and R. Knippers (1997). "High-molecular-mass complexes of human minichromosome-maintenance proteins in mitotic cells." Eur J Biochem **247**(1): 136-141.

Roberts, J. M. (1999). "Evolving ideas about cyclins." Cell **98**(2): 129-132.

Robinson, N. P., I. Dionne, et al. (2004). "Identification of two origins of replication in the single chromosome of the archaeon Sulfolobus solfataricus." Cell **116**(1): 25-38.

Rodins, K., M. Cheale, et al. (2002). "Minichromosome maintenance protein 2 expression in normal kidney and renal cell carcinomas: relationship to tumor dormancy and potential clinical utility." Clin Cancer Res **8**(4): 1075-1081.

Rykunov, D., E. Steinberger, et al. (2009). "Improved scoring function for comparative modeling using the M4T method." J Struct Funct Genomics **10**(1): 95-99.

Sabelli, P. A., S. R. Burgess, et al. (1996). "cDNA cloning and characterisation of a maize homologue of the MCM proteins required for the initiation of DNA replication." Mol Gen Genet **252**(1-2): 125-136.

Sakakibara, N., R. Kasiviswanathan, et al. (2008). "Coupling of DNA binding and helicase activity is mediated by a conserved loop in the MCM protein." Nucleic Acids Res **36**(4): 1309-1320.

Sato, M., T. Gotow, et al. (2000). "Electron microscopic observation and single-stranded DNA binding activity of the Mcm4,6,7 complex." J Mol Biol **300**(3): 421-431.

Schaper, S. and W. Messer (1995). "Interaction of the initiator protein DnaA of Escherichia coli with its DNA target." J Biol Chem **270**(29): 17622-17626.

Schwacha, A. and S. P. Bell (2001). "Interactions between two catalytically distinct MCM subgroups are essential for coordinated ATP hydrolysis and DNA replication." Mol Cell **8**(5): 1093-1104.

Schwede, T., J. Kopp, et al. (2003). "SWISS-MODEL: An automated protein homology-modeling server." Nucleic Acids Res **31**(13): 3381-3385.

Sclafani, R. A. (2000). "Cdc7p-Dbf4p becomes famous in the cell cycle." J Cell Sci **113 (Pt 12)**: 2111-2117.

Seidel, H. M., D. L. Pompliano, et al. (1992). "Phosphonate biosynthesis: molecular cloning of the gene for phosphoenolpyruvate mutase from

Tetrahymena pyriformis and overexpression of the gene product in Escherichia coli." Biochemistry **31**(9): 2598-2608.

Sekimizu, K., D. Bramhill, et al. (1987). "ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the E. coli chromosome." Cell **50**(2): 259-265.

Shechter, D. F., C. Y. Ying, et al. (2000). "The intrinsic DNA helicase activity of Methanobacterium thermoautotrophicum delta H minichromosome maintenance protein." J Biol Chem **275**(20): 15049-15059.

Sherman, D. A. and S. L. Forsburg (1998). "Schizosaccharomyces pombe Mcm3p, an essential nuclear protein, associates tightly with Nda4p (Mcm5p)." Nucleic Acids Res **26**(17): 3955-3960.

Sherr, C. J. and J. M. Roberts (1999). "CDK inhibitors: positive and negative regulators of G1-phase progression." Genes Dev **13**(12): 1501-1512.

Sheu, Y. J. and B. Stillman (2010). "The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4." Nature **463**(7277): 113-117.

Shi, Y., G. E. Dodson, et al. (2007). "Identification of carboxyl-terminal MCM3 phosphorylation sites using polyreactive phosphospecific antibodies." J Biol Chem **282**(12): 9236-9243.

Shin, J. H., B. Grabowski, et al. (2003). "Regulation of minichromosome maintenance helicase activity by Cdc6." J Biol Chem **278**(39): 38059-38067.

Shin, J. H., G. Y. Heo, et al. (2008). "The Methanothermobacter thermautotrophicus Cdc6-2 protein, the putative helicase loader, dissociates the minichromosome maintenance helicase." J Bacteriol **190**(11): 4091-4094.

Singleton, M. R., R. Morales, et al. (2004). "Conformational changes induced by nucleotide binding in Cdc6/ORC from Aeropyrum pernix." J Mol Biol **343**(3): 547-557.

Speck, C., Z. Chen, et al. (2005). "ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA." Nat Struct Mol Biol **12**(11): 965-971.

Speck, C. and W. Messer (2001). "Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA." EMBO J **20**(6): 1469-1476.

Speck, C., C. Weigel, et al. (1999). "ATP- and ADP-dnaA protein, a molecular switch in gene regulation." EMBO J **18**(21): 6169-6176.

Springer, P. S., W. R. McCombie, et al. (1995). "Gene trap tagging of PROLIFERA, an essential MCM2-3-5-like gene in Arabidopsis." Science **268**(5212): 877-880.

Starborg, M., E. Brundell, et al. (1995). "A murine replication protein accumulates temporarily in the heterochromatic regions of nuclei prior to initiation of DNA replication." J Cell Sci **108 (Pt 3)**: 927-934.

Stillman, B. (1994). "Initiation of chromosomal DNA replication in eukaryotes. Lessons from lambda." J Biol Chem **269**(10): 7047-7050.

Stillman, B. (1996). "Cell cycle control of DNA replication." Science **274**(5293): 1659-1664.

Stoeber, K., I. Halsall, et al. (1999). "Immunoassay for urothelial cancers that detects DNA replication protein Mcm5 in urine." Lancet **354**(9189): 1524-1525.

Stoeber, K., T. D. Tlsty, et al. (2001). "DNA replication licensing and human cell proliferation." J Cell Sci **114**(Pt 11): 2027-2041.

Studier, F. W. (1991). "Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system." J Mol Biol **219**(1): 37-44.

Studier, F. W., A. H. Rosenberg, et al. (1990). "Use of T7 RNA polymerase to direct expression of cloned genes." Methods Enzymol **185**: 60-89.

Su, T. T. and P. H. O'Farrell (1997). "Chromosome association of minichromosome maintenance proteins in Drosophila mitotic cycles." J Cell Biol **139**(1): 13-21.

Sykes, D. E. and M. M. Weiser (1995). "Rat intestinal crypt-cell replication factor with homology to early S-phase proteins required for cell division." Gene **163**(2): 243-247.

Taira, T., S. M. Iguchi-Ariga, et al. (1994). "A novel DNA replication origin identified in the human heat shock protein 70 gene promoter." Mol Cell Biol **14**(9): 6386-6397.

Takayama, Y., Y. Kamimura, et al. (2003). "GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast." Genes Dev **17**(9): 1153-1165.

Takei, Y., M. Assenberg, et al. (2002). "The MCM3 acetylase MCM3AP inhibits initiation, but not elongation, of DNA replication via interaction with MCM3." J Biol Chem **277**(45): 43121-43125.

Tanaka, S., T. Umemori, et al. (2007). "CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast." Nature **445**(7125): 328-332.

Tanaka, T., D. Knapp, et al. (1997). "Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs." Cell **90**(4): 649-660.

Thauer, R. K. (1998). "Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture." Microbiology **144 (Pt 9)**: 2377-2406.

Thommes, P., R. Fett, et al. (1992). "Properties of the nuclear P1 protein, a mammalian homologue of the yeast Mcm3 replication protein." Nucleic Acids Res **20**(5): 1069-1074.

Thommes, P., Y. Kubota, et al. (1997). "The RLF-M component of the replication licensing system forms complexes containing all six MCM/P1 polypeptides." EMBO J **16**(11): 3312-3319.

Thommes, P., T. A. Osman, et al. (1993). "TGMV replication protein AL1 preferentially binds to single-stranded DNA from the common region." FEBS Lett **319**(1-2): 95-99.

Thompson, J. D., T. J. Gibson, et al. (1997). "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." Nucleic Acids Res **25**(24): 4876-4882.

Todorov, I. T., A. Attaran, et al. (1995). "BM28, a human member of the MCM2-3-5 family, is displaced from chromatin during DNA replication." J Cell Biol **129**(6): 1433-1445.

Todorov, I. T., R. Pepperkok, et al. (1994). "A human nuclear protein with sequence homology to a family of early S phase proteins is required for entry into S phase and for cell division." J Cell Sci **107 (Pt 1)**: 253-265.

Todorov, I. T., B. A. Werness, et al. (1998). "HsMCM2/BM28: a novel proliferation marker for human tumors and normal tissues." Lab Invest **78**(1): 73-78.

Tran, N. Q., H. Q. Dang, et al. (2010). "A single subunit MCM6 from pea forms homohexamer and functions as DNA helicase." Plant Mol Biol **74**(4-5): 327-336.

Treisman, J. E., P. J. Follette, et al. (1995). "Cell proliferation and DNA replication defects in a Drosophila MCM2 mutant." Genes Dev **9**(14): 1709-1715.

Tsuruga, H., N. Yabuta, et al. (1997). "Expression, nuclear localization and interactions of human MCM/P1 proteins." Biochem Biophys Res Commun **236**(1): 118-125.

Tumbula, D. L. and W. B. Whitman (1999). "Genetics of Methanococcus: possibilities for functional genomics in Archaea." Mol Microbiol **33**(1): 1-7.

Tye, B. K. (1999). "MCM proteins in DNA replication." Annu Rev Biochem **68**: 649-686.

Tye, B. K. (2000). "Insights into DNA replication from the third domain of life." Proc Natl Acad Sci U S A **97**(6): 2399-2401.

Vanrobays, E., P. E. Gleizes, et al. (2001). "Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein." EMBO J **20**(15): 4204-4213.

Vas, A. and J. Leatherwood (2000). "Where does DNA replication start in archaea?" Genome Biol **1**(3): REVIEWS1020.

Wahle, E., R. S. Lasken, et al. (1989). "The dnaB-dnaC replication protein complex of Escherichia coli. I. Formation and properties." J Biol Chem **264**(5): 2463-2468.

Wahle, E., R. S. Lasken, et al. (1989). "The dnaB-dnaC replication protein complex of Escherichia coli. II. Role of the complex in mobilizing dnaB functions." J Biol Chem **264**(5): 2469-2475.

Walters, A. D. and J. P. Chong (2009). "Methanococcus maripaludis: an archaeon with multiple functional MCM proteins?" Biochem Soc Trans **37**(Pt 1): 1-6.

Walters, A. D. and J. P. Chong (2010). "An archaeal order with multiple minichromosome maintenance genes." Microbiology **156**(Pt 5): 1405-1414.

Weigel, C., A. Schmidt, et al. (1997). "DnaA protein binding to individual DnaA boxes in the Escherichia coli replication origin, oriC." EMBO J **16**(21): 6574-6583.

Wharton, S. B., K. K. Chan, et al. (2001). "Replicative Mcm2 protein as a novel proliferation marker in oligodendrogliomas and its relationship to Ki67 labelling index, histological grade and prognosis." Neuropathol Appl Neurobiol **27**(4): 305-313.

Wheeler, L. W., N. H. Lents, et al. (2008). "Cyclin A-CDK activity during G1 phase impairs MCM chromatin loading and inhibits DNA synthesis in mammalian cells." Cell Cycle **7**(14): 2179-2188.

Williams, K., C. Chubb, et al. (1998). "Analysis of differential protein expression in normal and neoplastic human breast epithelial cell lines." Electrophoresis **19**(2): 333-343.

Woese, C. R. and G. E. Fox (1977). "The concept of cellular evolution." J Mol Evol **10**(1): 1-6.

Woese, C. R. and G. E. Fox (1977). "Phylogenetic structure of the prokaryotic domain: the primary kingdoms." Proc Natl Acad Sci U S A **74**(11): 5088-5090.

Won, K. A. and S. I. Reed (1996). "Activation of cyclin E/CDK2 is coupled to site-specific autophosphorylation and ubiquitin-dependent degradation of cyclin E." EMBO J **15**(16): 4182-4193.

Xia, Q., E. L. Hendrickson, et al. (2006). "Quantitative proteomics of the archaeon Methanococcus maripaludis validated by microarray analysis and real time PCR." Mol Cell Proteomics **5**(5): 868-881.

Yoshimochi, T., R. Fujikane, et al. (2008). "The GINS complex from Pyrococcus furiosus stimulates the MCM helicase activity." J Biol Chem **283**(3): 1601-1609.

You, Z., Y. Komamura, et al. (1999). "Biochemical analysis of the intrinsic Mcm4-Mcm6-mcm7 DNA helicase activity." Mol Cell Biol **19**(12): 8003-8015.

Young, M. R. and B. K. Tye (1997). "Mcm2 and Mcm3 are constitutive nuclear proteins that exhibit distinct isoforms and bind chromatin during specific cell cycle stages of Saccharomyces cerevisiae." Mol Biol Cell **8**(8): 1587-1601.

Yu, X., M. S. VanLoock, et al. (2002). "The Methanobacterium thermoautotrophicum MCM protein can form heptameric rings." EMBO Rep **3**(8): 792-797.

Yu, Z., D. Feng, et al. (2004). "Pairwise interactions of the six human MCM protein subunits." J Mol Biol **340**(5): 1197-1206.

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): 281-285.

Zegerman, P. and J. F. Diffley (2010). "Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation." Nature **467**(7314): 474-478.

Zou, L., J. Mitchell, et al. (1997). "CDC45, a novel yeast gene that functions with the origin recognition complex and Mcm proteins in initiation of DNA replication." Mol Cell Biol **17**(2): 553-563.

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): 3086-3096.