

ABSTRACT

Title of Document: BIOCHEMICAL CHARACTERIZATION OF
THE MINICHROMOSOME MAINTENANCE
(MCM) HELICASE FROM
METHANOTHERMOBACTER
THERMAUTOTROPHICUS

Nozomi Sakakibara, Ph.D., 2009

Directed By: Professor Zvi Kelman,
Program in Molecular and Cell Biology

DNA replication requires coordination of numerous proteins to duplicate genetic information in a precise and timely manner. One of the key players in replication is the replicative helicase that unwinds the duplex DNA to provide the single-stranded template for the DNA polymerases. Minichromosome maintenance (MCM) protein is the replicative helicase in archaea. This dissertation focuses on the MCM helicase from the euryarchaeon *Methanothermobacter thermautotrophicus* (Mth).

Archaeal MCM proteins can be divided into two major parts, the N terminal and C terminal domains. The N terminal domain is essential for DNA binding and multimerization, while the C-terminus contains the catalytic domains. The objective of this dissertation is to elucidate the mechanism by which the N terminal domain communicates with the catalytic domain to facilitate helicase activity. To address

this question, two approaches were taken. One approach identified conserved residues found in the N terminus and investigated their properties using various biochemical and biophysical methods. By analyzing several proteins with mutations in the conserved residues, a loop that is essential for MCM helicase activity was identified. The study suggests that the loop is involved in coupling the N-terminal DNA binding function and the catalytic activity of the AAA+ domain. Some other conserved residues, however, did not directly affect the MCM helicase activity but showed differences in biochemical properties suggesting that they may play a role in maintaining the structural integrity of the MCM helicase. Another approach determined the differences in thermal stability of the MCM protein in the presence of various cofactors and DNA substrates. The study shows that the protein has two unfolding transitions when ATP and the DNA are present, while non-hydrolyzable ATP results in one transition. This study suggests possible conformational changes arising from decoupling of two domains that occur during the ATP hydrolysis in the presence of DNA. Furthermore, both DNA binding function by the N terminal domain and ATP binding by the catalytic domain are essential for the change.

BIOCHEMICAL CHARACTERIZATION OF THE
MINICHROMOSOME MAINTENANCE (MCM) HELICASE FROM
METHANOTHERMOBACTER THERMAUTOTROPHICUS

By

NOZOMI SAKAKIBARA

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:

Professor Douglas Julin, Chair
Professor Jeffery DeStefano
Professor Barbara Gerratana
Professor Jason Kahn
Professor Zvi Kelman, Co-chair

© Copyright by
Nozomi Sakakibara
2009

To the memory of my mother,

Acknowledgements

I would like to express deep gratitude to my advisor, Dr. Zvi Kelman who has been a terrific mentor throughout my Ph.D. studies. His wisdom and excellent attitude have taught me very important lessons not only in science but also in life. I feel very fortunate to have a mentor who is so passionate about the science and teaching students.

I am especially grateful to my committee members who have exhibited much kindness and have given valuable suggestions in the course of my program. Foremost, Dr. Douglas Julin, for being my official advisor and guiding me through many obstacles. The “Nucleic acids” course he taught was crucial in sparking my interest in DNA replication. I thank Dr. Barbara Gerratana and Dr. Jason Kahn for their time and advice. I learned a great deal from their expertise and insights. Also Dr. Jeffrey DeStefano, for giving his time to become involved as the Dean’s representative.

I would also like to thank the members of the Kelman lab - in particular –Rajesh, a talented scientist to whom I will always aspire; Jae-Ho and Gunyoung who nurtured the scientist in me; and Gyri for her encouragement and friendship. I would like to thank current members, Roxane, Zhou and Miao for their support in recent times. Also, helpful undergraduate and high school students, Mimi, Claire, Becky, and Evan, for their help. And, I would like to thank the pillar of the Kelman lab, Dr. Lori Kelman who took the time to proofread the manuscripts and always kept us in the circle.

I have been also fortunate to work with the fantastic faculty at CARB, and would like to thank especially Dr. Fred Schwarz, Dr. John Marino, Dr. Phil Bryan, Dr. Dan Nelson, and Dr. Illarion Turko for their help and suggestions during the course of my

Ph.D. work. I am also thankful to be in a great company of the members of the “3rd floor group meetings”, from the past to present; they have made me feel at home.

I have made many friends during graduate school. The presence of wonderful friends such as Madhura, Anni, Niya, Manish, Nobuko, Jayanthi, Ravi, Eugene, Al, and Quiyue made the past 5 years very enjoyable and unforgettable. Their friendship always made me feel supported and appreciated.

I would like to also thank Dr. Richard Highton, who encouraged me to go on to graduate school during my time in his lab as a volunteer researcher.

I would like to take this opportunity to thank my parents, Rumiko and Masahiro Sakakibara. You are to me what ATP is to MCM - the driving force. Thank you for your constant love and support, the examples you have shown will keep guiding me throughout my life.

Lastly, I would like to thank my family, my dear son, Zeus for a valuable relaxing time spent together, and my husband Askar, for encouraging me to always go the extra mile and strive for excellence. You are to me, what DNA and Mg^{2+} is to MCM - the essence of my life.

TABLE OF CONTENTS

Dedication.....	ii
Acknowledgements.....	iii
TABLE OF CONTENTS.....	v
LIST OF TABLES.....	xiii
LIST OF FIGURES.....	xiv
LIST OF ABBREVIATIONS.....	xviii
Chapter 1 Introduction.....	1
1.1 Overview of DNA replication.....	1
1.2 Archaea.....	2
1.2.1 Methanothermobacter thermoautotrophicus.....	2
1.2.2 Halobacterium sp. NRC-1.....	3
1.3 DNA replication initiation.....	3
1.4 Archaeal replication initiation.....	4
1.5 Origin.....	4
1.6 Cdc6 proteins.....	4
1.7 MCM helicase.....	5
1.7.1 What defines MCM helicase? Structural organization.....	7
1.7.2 Primary sequence analysis.....	8
1.7.3 Oligomeric state in solution.....	9
1.7.4 X-ray crystallographic structure.....	9
1.7.5 Electron microscopy (EM) structure.....	11

1.8	Biochemical properties of archaeal MCM helicases	12
1.8.1	Inherent MCM properties	13
1.8.2	Structure-Function studies of MCM	17
1.8.3	Active structure of MCM.....	24
Chapter 2	Coupling of DNA binding and helicase activity is mediated by a conserved loop in the MCM protein	36
2.1	Abstract.....	36
2.2	Introduction.....	36
2.3	Materials and Methods.....	38
2.1.1	Materials	38
2.1.2	Multiple alignment.....	38
2.1.3	Expression and purification of MCM mutant proteins	39
2.1.4	ATPase assay	40
2.1.5	Nitrocellulose filter DNA binding assay.....	41
2.1.6	Fluorescence polarization anisotropy (FPA) measurement	42
2.1.7	DNA helicase assay	43
2.1.8	Gel filtration analysis.....	45
2.1.9	Differential scanning calorimetry (DSC) measurements.....	45
2.1.10	Circular dichroism	46
2.1.11	UV crosslinking	46
2.4	Results.....	47

2.1.12	The archaeal MCM helicase contains highly conserved residues in the N-terminal portion.....	47
2.1.13	Mutations in the loop region between $\beta 7$ and $\beta 8$ affect MCM helicase activity.....	47
2.1.14	Mutation of conserved loop residues does not affect MCM structural integrity.....	49
2.1.15	Mutations in the loop do not affect DNA binding by MCM	50
2.1.16	Mutations in the loop region of the N-terminal domain alter the ATPase activity of MCM	51
2.5	Discussion.....	53
Chapter 3 Mutations at Asp 202 and 203 affect structural integrity of the		
	<i>Methanothermobacter thermautotrophicus</i> MCM helicase.....	73
3.1	Abstract.....	73
3.2	Introduction.....	73
3.3	Materials and Methods.....	75
3.3.1	Materials	75
3.3.2	Expression and purification of MCM mutant proteins	75
3.3.3	ATPase assay	75
3.3.4	Gel filtration.....	75
3.3.5	Nitrocellulose filter DNA binding assay.....	76
3.3.6	DNA helicase assay	76

3.3.7	Differential scanning calorimetry (DSC) measurements:	76
3.3.8	Circular dichroism (CD)	76
3.4	Results.....	76
3.4.1	Mutations at D202 and D203 residues reduce DNA binding of MCM	76
3.4.2	Mutations affect structural integrity of MCM protein	77
3.4.3	Mutations have no effect on helicase or ATPase activity	78
3.4.4	ATP stimulates DNA binding function of the mutants.....	79
3.5	Summary and Discussion.....	79
Chapter 4 Biochemical characterization of conserved residues found in N-terminal		
region of MCM helicase from <i>Methanothermobacter thermautotrophicus</i>		89
4.1	Abstract.....	89
4.1	Introduction.....	89
4.2	Materials and Methods.....	90
4.2.1	Materials	90
4.2.2	Expression and purification of MCM mutant proteins	90
4.2.3	Gel Filtration.....	90
4.2.4	ATPase assay	91
4.2.5	Nitrocellulose filter DNA binding assay.....	91
4.2.6	DNA helicase assay	91
4.2.7	Differential scanning calorimetry (DSC) measurements:	91
4.3	Results.....	91

4.3.1	The mutant proteins show diverse effect on ATPase activity.....	91
4.3.2	Mutations do not affect MCM helicase structure	92
4.3.3	Mutations in protein do not affect helicase activity.....	93
4.3.4	Mutations in MCM protein alter the DNA binding	94
4.3.5	ATP increases binding of MCM proteins to ssDNA	95
4.3.6	The Cdc6-2 protein does not inhibit the G211A, P210G, PG210,211GA mutant proteins.....	95
4.4	Summary and Discussion.....	96
Chapter 5 Effect of ATP hydrolysis and DNA binding on the thermostability of the MCM helicase.....		
		107
5.1	Abstract.....	107
5.1	Introduction.....	107
5.2	Materials and Methods.....	109
5.2.1	Materials	109
5.2.2	Differential scanning calorimetry (DSC) measurements	109
5.2.3	ATPase assay	111
5.2.4	Helicase assay	112
5.2.5	Isothermal titration calorimetry (ITC)	113
5.2.6	Tryptophan fluorescence study	113
5.3	Results.....	113
5.3.1	ATP increases the thermal stability of MCM in the presence of DNA	113

5.3.2	ATP γ S and ADP-AlF $_4^-$ are efficient competitors of ATP Binding.....	117
5.3.3	Binding to DNA ends required for MCM stabilization.....	118
5.4	Discussion.....	120
Chapter 6 Cloning, purification, and partial characterization of the <i>Halobacterium sp.</i>		
NRC-1 minichromosome maintenance (MCM) helicase.....		
6.1	Abstract.....	135
6.2	Introduction.....	135
6.3	Materials and Methods.....	136
6.3.1	Sequence alignment of MCM proteins.....	136
6.3.2	Cloning the MCM gene.....	136
6.3.3	Purification of the MCM proteins.....	138
6.3.4	Mass spectrometry analysis of the <i>E. coli</i> expressed proteins.....	139
6.3.5	Generation of antibodies.....	139
6.3.6	Preparation of <i>Halobacterium</i> cell extract.....	140
6.3.7	Western blotting analysis.....	140
6.4	Results and Discussion.....	141
6.4.1	Sequence analysis of <i>Halobacterium</i> MCM protein.....	141
6.4.2	Purification and identification of <i>Halobacterium</i> MCM helicase.....	141
6.4.3	Detection of <i>Halobacterium</i> MCM protein <i>in vivo</i>	143
Chapter 7 How is the archaeal MCM helicase loaded at the origin? Possible mechanisms.....		
		151

7.1	Abstract.....	151
7.2	Introduction.....	151
7.3	MCM structure.....	151
7.4	Helicase loaders	152
7.5	Formation of the replication bubble.....	153
7.6	Possible mechanisms of archaeal MCM helicase loading at the origin.....	154
7.6.1	Helicase loader-dependent assembly	154
7.6.2	Helicase loader-independent assembly	155
7.7	What is the role of duplex DNA translocation by the MCM helicase?	157
7.8	Do all archaea utilize the same loading mechanism?	158
Chapter 8	Concluding Remarks.....	164
Chapter 9	Appendices.....	169
9.1	Protein expression and purification	169
9.1.1	<i>M. thermoautotrophicus</i> MCM protein.....	169
9.1.2	<i>Halobacterium sp.</i> NRC-1 MCM protein	170
9.2	Fluorescence polarization anisotropy (FPA) measurements.....	172
9.1.3	General principle.....	172
9.1.4	Experimental procedure	178
9.3	Determination of kinetic parameter for ATPase activity	185
9.3.1	General principle.....	185
9.3.2	Experimental procedure	188
9.4	Differential scanning calorimetry (DSC).....	190

9.4.1 General principle.....	190
9.4.2 Experimental procedure.....	194
9.4.3 Rough estimation of bound species at higher temperature.....	198
9.5 Circular Dichroism (CD).....	200
9.5.1 General principle.....	200
9.5.2 Experimental procedure.....	201
9.6 Tryptophan fluorescence.....	202
9.6.1 General principle.....	202
9.6.2 Experimental procedure.....	204
BIBLIOGRAPHY.....	218

LIST OF TABLES

Table 2-1 Oligonucleotides used to generate the mutant clones.....	57
Table 2-2. Differential Scanning Calorimetry (DSC) analysis of wild type and mutant MCM proteins.....	57
Table 3-1 List of primers used to generate mutants.....	82
Table 3-2 Thermostability of wild-type and mutants from DSC results	82
Table 4-1 List of primers used to generate mutants.....	99
Table 4-2 List of mutants and their location and polar contacts.....	102
Table 4-3 DSC analyses of wild-type and mutant proteins.	102
Table 5-1 Oligonucleotides used in this study.....	123
Table 5-2 Thermostabilities of wild type MCM from DSC Results.....	123
Table 5-3 Thermostabilities of mutant MCM proteins from DSC Results.....	124
Table 5-4 Thermostabilities of wild type MCM from DSC Results.....	125
Table 6-1 Amino acid analysis of MCM proteins.	146
Table 9-1 Example of how the anisotropy value is recorded as protein is titrated into the fluorescently labeled DNA.	206
Table 9-2 Rough estimation of MCM bound by DNA based on K_d obtained from room temperature, using quadratic equation.	214

LIST OF FIGURES

Figure 1-1 Schematic diagram for initiation of DNA replication.....	26
Figure 1-2 Schematic representation of the structural motifs of the archaeal MCM helicase.....	27
Figure 1-3 Organization of AAA+ proteins.....	29
Figure 1-4 Alignment of archaeal MCM helicase that have been characterized <i>in vitro</i>	30
Figure 1-5 Structures of N-terminal portion of the MCM helicase.	32
Figure 1-6 Structures of monomeric forms of the MCM helicases.	34
Figure 1-7 Models of hexameric structures of the MkaMCM2 and SsoMCM proteins..	35
Figure 2-1 Alignments of the N-terminal portion of MCM and amino acid sequences of the loop between $\beta 7$ and $\beta 8$	58
Figure 2-2 Monomeric structure of the N-terminal part of MCM.....	60
Figure 2-3 Loop is located in the opposite side of the dimer interface.....	61
Figure 2-4 EM structure of full length MCM.....	62
Figure 2-5 Phylogenetic tree of N-terminal archaeal MCM.....	63
Figure 2-6 Conserved residues in the loop between $\beta 7$ and $\beta 8$ affect MCM helicase activity.....	64
Figure 2-7 Mutations in the loop between $\beta 7$ and $\beta 8$ alter the rate of DNA unwinding.	65
Figure 2-8 Mutations in the loop do not change the overall structure of MCM.....	66
Figure 2-9 Mutations in the loop between $\beta 7$ and $\beta 8$ do not affect MCM DNA binding	68
Figure 2-10 Mutations in the loop between $\beta 7$ and $\beta 8$ alter the ATPase activity of MCM.	70
Figure 2-11 MCM proteins with mutations in the loop can bind ATP	71

Figure 2-12 A model for the roles played by the N-terminal loop in MCM function	72
Figure 3-1 Residues, Asp 202 and 203 are highly conserved among archaea.....	83
Figure 3-2 Mutations affect ss and dsDNA binding	84
Figure 3-3 The effect of mutations on structural integrity.....	85
Figure 3-4 Mutants have helicase and ATPase activities	86
Figure 3-5 DNA binding affinity is increased in the presence of ATP.	87
Figure 3-6 Asp 202 and 203 are located in the interface of domain C and domain A	88
Figure 4-1 Locations of conserved residues in monomeric MCM.	100
Figure 4-2 Mutations cause various levels of ATPase activity.....	101
Figure 4-3 Mutations alter the oligomeric state of MCM.....	103
Figure 4-4 Mutations do not substantially affect helicase activity	104
Figure 4-5 Mutations affect ss and dsDNA binding	105
Figure 4-6 Cdc6 proteins do not interact with the conserved residues	106
Figure 5-1 Effect of DNA and nucleotides on MCM thermostability.....	126
Figure 5-2 Effects of nucleotide analogues on MCM ATPase activity	128
Figure 5-3 The effects of nucleotide analogues on MCM helicase activity	129
Figure 5-4 The effects of DNA substrates on MCM thermostability	130
Figure 5-5 MCM interacts with DNA substrates.	131
Figure 5-6 Effect of DNA on MCM ATPase activity.....	132
Figure 5-7 Summary of MCM protein thermostability in the presence of different co-factors.....	133
Figure 6-1 General scheme of how the intein is expressed and excised from the <i>Halobacterium</i> MCM helicase.....	144

Figure 6-2 Alignment of full length <i>Halobacterium</i> MCM protein	145
Figure 6-3 The purification of the <i>Halobacterium</i> MCM helicase.....	147
Figure 6-4 The <i>E. coli</i> expressed proteins are the <i>Halobacterium</i> MCM proteins.....	148
Figure 6-5 Only the MCM protein without the intein can be detected in <i>Halobacterium</i> cells	150
Figure 7-1 Possible mechanisms of replication bubble formation.....	160
Figure 7-2 Possible mechanisms of loader-dependent helicase assembly.....	162
Figure 7-3 Helicase loader-independent assembly	163
Figure 9-1 Example of anisotropy vs. protein concentration plot	207
Figure 9-2 Normalized anisotropy.....	207
Figure 9-3 Analysis of MCM mutant (F172I) binding to DNA using GraFit 5.0.	208
Figure 9-4 Analysis of MCM mutant (F172I) binding to DNA data fitting to cooperative binding equation.....	209
Figure 9-5 Example of ATPase data fitting to Michaelis-Menten equation using GraFit program.....	210
Figure 9-6 Example of a DSC scan.....	211
Figure 9-7 Example of subtraction of baseline (second scan) from the first scan.....	211
Figure 9-8 BSA standard plot	212
Figure 9-9 Standard curve of Bradford assay and spectroscopic (Abs ₂₈₀ nm) determination	212
Figure 9-10 Example of DSC analysis using EXAM program.....	213
Figure 9-11 Example of DSC analysis to obtain ΔH_{cal}	213

Figure 9-12 CD spectra of wild type MCM and increasing amount of guanidium hydrochloride (GuHCl).....	215
Figure 9-13 Tryptophan fluorescence signal of MCM protein over time using 5 nm slit width	216
Figure 9-14 Tryptophan fluorescence signal of MCM protein over time using 4 nm slit width	217

LIST OF ABBREVIATIONS

AAA+	ATPases associated with various cellular activities
Afu	<i>Archaeoglobus fulgidus</i>
ASCE	Additional strand conserved E family
bp	Base pair
BSA	Bovine serum albumin
CD	Circular dichroism
Cdc6	Cell division cycle 6
DSC	Differential scanning calorimetry
dsDNA	Double-stranded DNA
EM	Electron microscope
H2ins	Helix 2 insertion
Hal	<i>Halobacterium sp.</i> NRC-1
HTH	Helix turn helix
IR	Inverted repeat
ITC	Isothermal titration calorimetry
kDa	Kilodalton
Mbu	<i>Methanococoides burtonii</i>
MCM	Minichromosome maintenance
Mka	<i>Methanococcus kandleri</i>
Mth	<i>Methanothermobacter thermautotrophicus</i>
OB	Oligonucleotide/oligosaccharide binding
OBP	Origin binding protein

ORB	Origin recognition box
ORC	Origin recognition complex
ORP	Origin recognition protein
Pfu	<i>Pyrococcus furiosus</i>
PS1BH	Pre-sensor I β hairpin
PS-II	Pre-sensor II insertion
R-F	Arginine finger
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-Polyacrylamide gel electrophoresis
S-I	Sensor I
S-II	Sensor II
ssDNA	Single-stranded DNA
Sso	<i>Sulfolobus solfataricus</i>
Tac	<i>Thermoplasma acidophilum</i>
T _m	Melting temperature
WA	Walker A
WB	Walker B
WH	Winged-helix

Chapter 1 Introduction

1.1 Overview of DNA replication

DNA replication is an essential process and thus conserved in all life forms (reviewed in (1,2)). DNA replication starts at a particular region of the chromosome called the origin (3). Origin binding proteins (OBP) bind specifically to the origin and distort the duplex DNA, creating a replication bubble. Single stranded (ss) DNA temporarily becomes available for the replicative helicase to be recruited and to unwind the double stranded (ds) DNA in a bidirectional manner. The ssDNA binding proteins (SSB) bind to ssDNA to prevent the ssDNA from reannealing and to protect the DNA. Primase is recruited onto the ssDNA region and generates the primers required by the DNA polymerase to initiate DNA synthesis. The primers are subsequently removed by endonucleases, and polymerase fills in the resulting gap. The nicks created between newly synthesized DNA fragments are ligated with DNA ligase. The termination stage of DNA replication in prokaryotes takes place at the point where the progressing replication forks meet in the circular chromosome. In eukaryotes, however, it remains unclear how the replication is terminated (4,5).

All these factors and events are tightly coordinated. One of the key events in licensing the initiation of replication is the helicase assembly at the origin. Many studies point to the minichromosome maintenance (MCM) helicase protein as the replicative helicase in eukaryotes and archaea. While eukaryotic Mcm helicase consists of six subunits (Mcm2-7p), most archaea contain only one MCM homologue.

1.2 Archaea

The archaea, along with bacteria and eukarya, have been established as one of three domains of life (6-8). The archaea, due to their prokaryotic nature, share similarity to bacteria in terms of cell structure and metabolism. However, archaeal proteins involved in informational processing (replication, transcription, and translation) are more similar to those from eukaryotes. Archaea often inhabit extreme conditions and are classified into five major kingdoms; Euryarchaeota, Crenarchaeota, Korearchaeota, Nanoarchaeota, and Thaumarchaeota (9-11). Euryarchaeota consist of methanobacteriales, halobacteriales, and many other species. Crenarchaeota consist of solfolobales, thermoproteales, and desulflurococcales. Among the 48 archaeal genomes that are sequenced to date, 33 belong to the Euryarchaeota kingdom, 14 to the Crenarchaeota, and 1 to the Nanoarchaeota (NCBI website).

1.2.1 Methanothermobacter thermautotrophicus

Methanothermobacter thermautotrophicus is a strict anaerobic organism with an optimal growth temperature between 62 and 65°C. It metabolizes H₂ and CO₂ producing CH₄ and H₂O, with a 5 hour generation time (12). The complete genome of the *M. thermautotrophicus* strain ΔH was sequenced and published in 1997, revealing 1.7 Mb with about 1900 open reading frames, and a GC content of 49.5% (13). Flow cytometry analysis and DAPI staining suggest that *M. thermautotrophicus* cells each contain 2-8 nucleoids and form filaments with different lengths depending on the growth phase (14). When the cells are growing exponentially, they tend to form longer filaments, but once the stationary phase is reached, the length of the filaments is limited to two cells. The *M. thermautotrophicus* genome contains one origin of replication, one homologue of the

eukaryotic MCM helicase, and two homologues of the putative helicase loader and OBP, Cdc6.

1.2.2 Halobacterium sp. NRC-1

Halobacterium sp. NRC-1 is a halophilic archaeon with an optimal growth temperature of 42°C and a 2 hr generation time (15). The genome of *Halobacterium* sp. NRC-1 consists of 2.5 Mb and encodes for about 2,600 proteins. The entire genome consists of one large chromosome (2 Mb, 67.9 % GC) and two extrachromosomal replicons, pNRC100 (191 kb, 59.9% GC) and pNRC200 (365 kb, 57.2 % GC) (16). Both extrachromosomal replicons contain essential genes and thus have been suggested to be vital for cell survival. There were two predicted origins in *Halobacterium* sp. NRC-1, however only one origin has been identified to contain an autonomously replicating sequence (ARS) *in vitro* (17,18). One MCM helicase homologue and three Cdc6 encoding genes are located in the large chromosome, and six *cdc6* genes are mapped in the pNRC200 and one *cdc6* gene is located in the pNRC100. MCM protein, as well as Cdc6-2 and Cdc6-10 were determined to be essential for cell survival (19).

1.3 DNA replication initiation

DNA replication initiation requires critical timing and regulation of origin firing by initiation proteins such as OBP, helicase loader, and replicative helicase. The initiation stage involves the binding of OBPs to the origin recognition boxes (ORBs) in a sequence specific manner and the distortion of the neighboring A/T rich region known as the DNA unwinding elements (DUE), forming partially unwound DNA (Figure 1-1). Onto this temporarily separated DNA region (also known as the replication bubble) the

helicase loaders together with OBPs load the helicase (1,20,21). This ensemble at the origin is known as the pre-replicative complex (pre-RC) in eukaryotes. The formation of the pre-RC licenses replication by recruiting proteins involved in the elongation phase, which proceed to synthesize DNA (22,23). Origin architecture, as well as these proteins are essential for cell viability and are thus conserved in all three domains of life (reviewed in (2)).

1.4 Archaeal replication initiation

1.5 Origin

The typical origin of replication is located in an intergenic region and consists of a DUE, and several ORBs. The ORBs have direct and inverted repeats to facilitate the binding of OBPs. This region is often flanked by the *cdc6* gene which encodes the archaeal OBP, Cdc6 (21,24-27). *In vitro*, a single origin has been identified in the archaeons, *M. thermautotrophicus*, *Halobacterium sp.* NRC-1, *Pyrococcus abyssi*, and *Archaeoglobus fulgidus* (18,28-30). However, multiple origins have been found in several other archaea, including *Sulfolobus solfataricus*, *Aeropyrum pernix*, and *Haloferax volcanii* (26,27,31).

1.6 Cdc6 proteins

Many archaea carry two homologues of eukaryotic helicase loaders, Cdc6 proteins (1). The eukaryotic Cdc6 protein has a sequence homology to the eukaryotic OBP, origin recognition complex 1 (Orc1) (32). Thus the archaeal homologue, Cdc6 is also referred to as Cdc6/Orc1, but for simplicity it is referred to as Cdc6 throughout this text. Archaeal Cdc6 proteins belong to a large family of AAA+ (ATPase associated with

various cellular activities) proteins. Archaeal Cdc6 homologues have been classified into two classes by sequence analysis (33). Class I type Cdc6 bind specifically to the origin sequence, while class II type proteins have non-specific binding to dsDNA (29,34-36). Crystal structures of both classes of Cdc6 proteins reveal that Cdc6 comprises three domains, two of which have a canonical AAA+ type fold (described later) and a C-terminal winged helix (WH) domain (33,36-38). The structures of Class I Cdc6 proteins in complex with short ORB sequence revealed that the binding of Cdc6 proteins causes distortion of dsDNA by bending it by $\sim 15\text{-}30^\circ$ (36,38).

1.7 MCM helicase

The MCM proteins were first identified in a screen to identify proteins required for minichromosome maintenance in *Saccharomyces cerevisiae* (39). Additional studies showed that six of these proteins, Mcm2-7, form a family that found in all eukaryotic species and are essential for the initiation and elongation phases of DNA replication in eukarya (reviewed in (40-42)). In eukaryotes the MCM complex is a family of six essential polypeptides (Mcm2-7) stable hexamer in solution (43). Biochemical studies have shown that a dimeric complex of the Mcm4,6,7 heterotrimer possesses weak 3'-5' DNA helicase activity, can translocate on single and double stranded DNA, can bind DNA and RNA, and has the ability to unwind DNA-RNA duplexes while translocating on the DNA strand (40,41,44). A recent study indicates that the Mcm2-7p activity *in vitro* is strongly anion dependent (45). It was suggested that that Mcm2,5 plays a role in regulation of Mcm4,6,7, possibly as a gate of the ring opening in order to be loaded onto the DNA in the Mcm2-7p complex (45,46). Mcm proteins are distributed randomly in eukaryotic genome and exist in excess number over the number of origin present in

eukaryotic genome, and a recent study suggests that the Mcm proteins play a role in rescuing the replication process in stress induced cells (47-49). All archaeal species sequenced to date contain at least one homologue of the eukaryotic Mcm2-7 family (50,51). Based on structural and biochemical features it is thought that the archaeal MCM proteins function as the archaeal replicative helicases.

Although a limited number of *in vivo* studies have been reported on archaeal MCM proteins, the MCM helicase is thought to be the replicative helicase. Studies with the *Halobacterium* enzyme suggested that the protein is essential for cell viability (19). MCM expression level is constant throughout the exponential and stationary phases of the cell cycle in *Sulfolobus acidocaldarius* and *Pyrococcus abyssi* (52,53). In both archaeons, MCM is specifically bound to the origin during early exponential phase but moves away from the origin as the cell cycle progresses (52,53).

Archaea usually contain one MCM homologue in their genomes, although some archaea contains up to 4 MCM homologues (1,2,54,55). Although MCM protein is expressed in all archaea, the level of expression varies. *Pyrococcus abyssi* contains about 400 molecules of MCM per cell (29) and *M. thermautotrophicus* has about 300 molecules. *T. acidophilum* cells contain more than 1000 molecules per cell (56). *P. abyssi* cells contain up to 15 genome equivalents in rapidly dividing cells (29) and thus there are about 12 MCM molecules per replication fork. Similar numbers of MCM molecules are present in *M. thermautotrophicus*, which contains up to 12 genome equivalents (14).

1.7.1 What defines MCM helicase? Structural organization.

MCM helicase belongs to a large group of AAA+ proteins. Primary structure analysis of the MCM helicase suggests that it is divided into three domains: N-terminal, AAA+ core, and C-terminal helix-turn-helix (HTH) domains (Figure 1-3) (57).

The N-terminal portion of MCM helicase is less conserved than the AAA+ region across species. However, this domain contains a highly conserved accessory motif, the Zinc (Zn) binding motif, that plays a role in DNA binding and in a protein-protein interaction (58-60). The Zn binding motif is usually characterized by four conserved cysteine (C) residues that coordinate a Zn atom. The C4 type has a common consensus sequence of C-X₂-C-X_n-C-X₂-C (found in Euryarchaeota) and the C3H type contains a histidine with a consensus sequence of H-X₄-C-X_n-CX₂-C (found in Crenarchaea) (61). Another conserved motif found in archaeal MCM is the β hairpin finger motif that contains positively charged residues in a sequence that is involved in DNA binding (59).

The AAA+ core domain is highly conserved and contains many of the signature motifs of MCM helicase. The AAA+ proteins contain motifs that are important in ATP binding and hydrolysis and are predicted to have a common fold that has a propensity to oligomerize (62-66). This common fold among a diverse group of NTPases is known as ASCE (additional strand conserved E family) core fold (Figure 1-2 A). The typical ASCE core fold consists of an $\alpha\beta\alpha$ fold that contains a Walker-A (WA) and a Walker-B (WB) motifs, an arginine (R) finger, and a sensor I (S-I) that are all essential for ATP binding and hydrolysis (Reviewed in (67,68)) (Figure 1-2 A). AAA+ proteins often contain a lid domain in the C-terminal portion of the AAA+ domain that consists of α bundle (Figure 1-2 A). This lid domain contains a highly conserved motif known as

sensor II (S-II) that is located near the active site of the ASCE fold and also plays a role in ATP binding and hydrolysis.

In addition to those AAA+ key motifs, the MCM helicase has several features that distinguishes it from the rest of the AAA+ proteins (Figure 1-2 B) (63). These features include a pre-sensor 1 β hairpin (PS1BH), a helix-2-insertion (H2ins), and a pre-sensor 2 insertion (PS-II). The PS1BH and H2ins are insertions of β hairpins in the typical ASCE core fold (Figure 1-2 B) (63,67). The PS-II is an insertion of a long α helix in the lid domain before the S-II. This insertion repositions the S-II motif in a distant location, closer to R finger that assists ATP hydrolysis of a neighboring subunit, thus acting in *trans* (Figure 1-2 B).

The archaeal MCM C-terminal domain has been suggested to fold into a common DNA binding motif, helix-turn-helix (HTH). Since this homology is only found in archaeal MCM, the domain is called either the C-terminal or HTH domain (57). Here it will be referred to as the C-terminal HTH domain. The HTH motif is usually found in transcription factors and regulators that have specific interactions with DNA (57).

Based on primary sequence, structure, and biochemical analysis, a large group of helicases and translocases were classified into six different superfamilies (SF) (58). While most of the helicases that belong to AAA+ proteins belong to SF3 helicases, MCM proteins belong to SF6 along with RuvB helicases that play a role in homologous recombination in prokaryotes (58).

1.7.2 Primary sequence analysis

MCM helicases from *M. thermautotrophicus* (Mth), *Sulfolobus solfataricus* (Sso), *Thermoplasma acidophilum* (Tac), *Archaeoglobus fulgidus* (Afu), *Methanopyrus*

kandleri (Mka), and *Pyrococcus furiosus* (Pfu) *Methanococcus burtonii* (Mbu), *Halobacterium sp.* NRC-1 (Hal) have been isolated and characterized. As shown in Figure 1-4, the alignment of all of these archaeal MCM helicases contain the key motifs WA, WB, S-I, R-finger, S-II, PS1BH, H2-ins, and PS-II in the AAA+ core domain, except for MkaMCM2. *M. kandleri* has two MCM homologues, MkaMCM1 encodes for active protein but MkaMCM2 encodes for an inactive truncated form, as it is missing several essential motifs (Figure 1-4). HalMCM and AfuMCM contain an intein, an inserted peptide which excises itself at the level of the peptide, after translation (69). HalMCM was determined to be free of inteins *in vivo*, but that could not be demonstrated *in vitro* ((70); also see Chapter 6).

1.7.3 Oligomeric state in solution

Original reports of MthMCM forming double hexamers in solution were demonstrated independently by three laboratories using gel filtration and glycerol gradient analysis (71-73). MCM helicases from *S. solfataricus*, *T. acidophilum*, *A. fulgidus*, and *M. burtonii* have been reported to form single hexamer from gel filtration analysis (74-77). MkaMCM2 and PfuMCM were determined to exist as monomers (78,79). Monomers, in addition to multimers were also observed in both SsoMCM and MthMCM. The latter two proteins seem to be in equilibrium between different oligomeric states in solution as will be discussed later (71,74).

1.7.4 X-ray crystallographic structure

X-ray crystallographic structures of the N-terminal portion of *M. thermautotrophicus* and *S. Solfataricus* have been solved (Figure 1-5) (59,80). The

structure of MthMCM shows a dumb-bell like double hexamer, and SsoMCM exists as a single hexamer. Overall the structures of the two proteins resemble each other, with a reported root mean square deviation (rmsd) value of 1.3 Å based on C α atom superimposition (60). Each monomer from both proteins folds into three distinct sub-domains (Figure 1-6 B and C). Domain A, at the N-terminus, is mostly α -helical. Domain B has three β -strands and contains a Zn motif that mediates double hexamer formation in MthMCM. Domain C contains five β -strands that form an oligonucleotide/oligosaccharide binding (OB) fold, and a β hairpin finger motif that protrudes into the central cavity of hexameric molecule (Figure 1-6 B and C). Detailed analysis of the two structures of the N-terminal MCM revealed that there are differences. While MthMCM has a central cavity large enough to accommodate dsDNA, SsoMCM has a smaller size (with the smallest diameter of 23 Å and 17 Å respectively) (Figure 1-5). Another difference between two structures is found in the relative position of the Zn motif, where the distance between zinc atom of MthMCM and SsoMCM is reported to differ by 7.2 Å when two structures were overlapped and compared (60). It has been suggested that this difference might explain the observed single and double hexamers between two almost identical crystals. The overall charges of these crystal structures revealed that residues facing central cavity are positively charged (59,60,81).

Recently, the crystal structure of monomeric SsoMCM was solved with resolution of 4.1 Å (82). The structure revealed the N-terminal and AAA+ regions of SsoMCM. The N-terminal portion resembles the N-terminal structures solved from MthMCM and SsoMCM as described above (Figure 1-5 D). The structure of the AAA+ domain consists of a canonical ASCE core fold, an α/β domain consisting of 5 main helices and 5

main β strands connected to an α bundle domain with three long α helices. A loop that contains an α helix connects the latter two domains (Figure 1-6 D). In the SsoMCM structure, this α/β domain is interrupted by three β hairpins, PS1-hairpin, H2-ins, and a third β hairpin, known as EXT-hp since it is located on the exterior surface of modeled hexameric helicase (Figure 1-2). The α bundle domain also known as the lid domain consists of three long helices. One of the three, $\alpha 6$ is also known as PS-II that positions S-II in a close proximity to the R-finger (Figure 1-2 B). Modeling of monomeric SsoMCM to the N-terminal MthMCM hexameric structure allowed depicting a putative active hexameric helicase (Figure 1-7 B).

The crystal structure of monomeric full length MCM2 from *M. kandleri* was determined to a resolution of 1.9 Å (79). *M. kandleri* has two homologues of MCM, 1 and 2. MkaMCM2 lacks the several essential motifs as well as C-terminal HTH domain (Figure 1-4). Due to the lack of several domains, the crystal structure reveals that the N-terminal portion has poor similarity to the MthMCM or SsoMCM N-terminal crystal structures (Figure 1-6 E). Modeling of the AAA+ domain structure of the MkaMCM2 into a cryo-EM structure of hexameric MthMCM allowed reconstruction of putative active sites in the hexameric structure (Figure 1-7 A). This reconstruction revealed that the ATPase active site of MCM involves the conserved motifs WA, WB, and S-I acting in *cis* while the R-finger and S-II act in *trans* consistent with the structure of SsoMCM.

1.7.5 Electron microscopy (EM) structure

Until recently the high resolution structures of the full length MCM helicases were not available. Thus in the past several years, EM studies have shed light on the

overall MthMCM structure as well as its unique properties. Consistent with the structure in solution, the first image of MthMCM obtained by EM revealed a cylindrical structure showing that it exists as a double hexamer (72). However, subsequent EM studies on MthMCM revealed heptamer, hexamer, filaments and open forms (83-86). These structures may be dependent on the treatment of the samples prepared for the EM studies. Nonetheless, the flexibility to form various oligomers may be a mechanism of regulation as will be discussed later.

Conformational and oligomeric state changes in MthMCM protein in the presence of DNA substrate and nucleotide were observed using EM. According to a set of studies, the presence of ADP-AlFx and AMP-PNP stabilize the double heptamer of MCM, otherwise heterogeneous but mostly single hexamers under the same condition without these nucleotides (87,88). The presence of dsDNA resulted in a stable double hexamer, and the addition of nucleotide analogues did not change the double hexameric structure (88). Surprisingly, the two rings that comprise the double heptamer or hexamer are not symmetrical (87).

1.8 Biochemical properties of archaeal MCM helicases

Among reported studies on MCM helicase, most fall into one of two general types of investigation. In order to simplify the large amount of information, this section is divided into two parts. The first part summarizes the studies that characterize the function of wild-type MCM by observing its inherent activity. The second part reviews the studies that characterize the structure-function relationships by means of mutagenesis.

1.8.1 Inherent MCM properties

1.8.1.1 Helicase activity

All archaeal MCM proteins characterized unwind substrate in the 3'-5' direction since helicase activity can be detected only with substrates with 3' overhang available (71,72,74-76,89,90). The opposite strand of a forked substrate, the 5' overhang, has been determined to be important for helicase activity in TacMCM and SsoMCM. All archaeal MCM proteins have better activity when a 5' overhang is present (44,74-76,89,91). TacMCM requires about 16 nt on the 3' end and 8 nt on the opposite strand while SsoMCM requires 9 nt in the 5' end (74,75). Moreover, FRET analysis of SsoMCM suggests that there is a transient interaction of the 5' tail of forked substrate with the outer surface of the protein (91). The preferred specificity for the forked substrate could be due to inherent affinity of MCM for the replication bubble like structure. To support this, SsoMCM and AfuMCM showed higher affinity for the DNA substrate mimicking the replication bubble than for ss or dsDNA (76,92). AfuMCM has even higher affinity for those bubbled substrates with longer ssDNA portion. DNA foot printing assay supported that AfuMCM specifically binds to the bubble, not to the duplex region (76). MthMCM was also shown to prefer the forked like substrate over a flat substrates with experiments with various substrates (44). These evidences suggest that MCM preferentially binds to the replication fork.

In vivo, DNA is coated by histones and other proteins. The MthMCM helicase can displace proteins such as transcription factor and histones bound on DNA as it translocates along the DNA (93). It is likely that MCM helicase would encounter these proteins while it is progressing on the replication fork, and may have to remove them so

that the fork is not stalled. Other than replication fork progression, it has been suggested that MCM proteins have diverse functions other than as replicative helicases, such as in transcription (94). MthMCM can displace DNA-RNA duplex substrate while translocating along the DNA, suggesting multiple role for archaeal MCM (93).

1.8.1.2 ATPase activity

Helicase unwinding requires the coupling of ATP hydrolysis to a translocation mechanical movement. All archaeal MCM helicases have ATPase activity, some of which are stimulated in the presence of DNA. In MthMCM, it was shown that ATPase activity is strongly dependent on the presence of either ss or dsDNA (71-73). AfuMCM also has ATPase activity that is stimulated in the presence of a ss oligonucleotide (76). In addition to these observations, a cryo-EM structural study mentioned earlier suggests that DNA may provide structural order that allows better hydrolysis by MCM (87). In the case of SsoMCM, TacMCM, and PfuMCM, the ATPase activity is not stimulated in the presence of DNA (74,78,95).

Based on the crystal structures, several models have been proposed for ATP hydrolysis by the six subunits of the hexameric replicative helicases (58,96-99). These include a sequential model in which ATP hydrolysis is coordinated between adjacent subunits, a stochastic model in which ATP hydrolysis by different subunits is independent of one another, and a concerted model in which all six subunits hydrolyze ATP at the same time. Studies with the SsoMCM revealed that the enzyme hydrolyzes ATP in a stochastic manner while helicase activity uses a sequential mode of unwinding (100). In the case of SsoMCM ATPase activity is not coupled with DNA binding

function, thus, it is not clear whether the model of ATP hydrolysis can be generalized to other archaeal MCM proteins.

Analysis of the thermostability of MthMCM revealed that the protein is stabilized in the presence of both ATP and DNA while ATP or DNA alone did not have such an effect (see Chapter 5). Among several nucleotide analogues used, the ATP hydrolysis transition state analogue revealed the same effect as the ATP suggesting that hydrolysis is required for thermostability of MthMCM. This implies that the conformational change occurs during ATP hydrolysis in the presence of DNA.

1.8.1.3 DNA binding

Archaeal MCM helicases must interact with DNA to have helicase activity. Binding constants (K_d) of archaeal MCM to different DNA substrates have been reported to vary from 10 to 300 nM (76,90,95,101-103). DNA binding is affected by several factors. ATP also stimulates MthMCM binding to DNA (61,71), while ATP γ S had only limited effect on DNA binding. The stimulatory effect of ATP on DNA binding can only be observed when the reaction is incubated at 50-60°C. When the DNA binding reactions were incubated at 22°C no change in DNA binding could be detected in the presence of ATP (72,76,102,103).

The orientation in which the MCM helicase is loaded onto the DNA was studied using FRET analysis (90). The study demonstrated that the C-terminal portion of SsoMCM enters the 3' end of DNA first and slides toward the 5' end. This study was performed at 25 °C without ATP thus it is likely that the C-terminal portion of SsoMCM has an inherent affinity for 3' end and slides without having to translocate by ATP hydrolysis (90).

1.8.1.4 Regulators of MCM activity

As mentioned above, ATPase activity and DNA binding are stimulated in the presence of each other in MthMCM. In some archaea, instead of DNA, other factors stimulate ATPase and helicase activities. In the case of PfuMCM, both activities were significantly enhanced in the presence of GINS, another initiation factor found in eukaryotes and archaea, and a similar effect was observed for TacMCM in the presence of Cdc6 proteins (78,95). In contrast to these factors that stimulate activity, some factors inhibit the MCM helicase. In MthMCM and SsoMCM, helicase and ATPase activities were inhibited in the presence of Cdc6 proteins (104,105). This inhibition of MthMCM by Cdc6 was determined to be via protein-protein interaction, rather than by indirect inhibition by binding to the DNA substrate (35). Interestingly SsoMCM had increased affinity for DNA in the presence of Cdc6 proteins, suggesting that the tight interactions between SsoMCM and DNA may prevent translocation (106).

1.8.1.5 Optimal MCM activity

Many archaea inhabit under harsh environmental conditions, including extremes of salt concentration, pH, and temperature. The archaeal MCM proteins characterized *in vitro* have been from thermophilic organisms. Thus the helicase and ATPase assays are usually carried out between 55 and 70°C, temperatures at which activity is detectable and substrates are stable. Detailed analysis of TacMCM helicase activity dependence on temperature, ATP concentration, salt concentration, pH, and buffer conditions has been reported (95). Striking differences can be seen in the results. TacMCM helicase has three fold higher activity when Tris buffer is used instead of Hepes buffer, at the optimal pH of 8.5. AfuMCM has optimal ATPase activity at temperature of about 75°C, while

helicase assays were carried out between 50 and 60°C (76). DNA binding assays were carried out at 60°C, room temperature or on ice. It is clear that the various conditions alter the enzyme activity. However, it is not well understood how the heat may activate the enzyme activity.

1.8.2 Structure-Function studies of MCM

In order to understand the mechanism of archaeal MCM helicase, structure-function relationships were investigated by means of mutagenesis. MCM helicases from two archaea, Mth and Sso have been at the forefront of investigation in this section. When appropriate, these two helicases are compared and contrasted. Three major domains, N-terminal, AAA+ and C-terminal HTH domains are addressed separately.

1.8.2.1 N-terminal domain

When expressed alone, the N-terminal domain forms double hexamer in MthMCM (59,81). The N-terminal domain of SsoMCM forms monomers at lower concentration, but forms multimer at higher concentration, and as shown in the crystal structure, it is sufficient to form a hexameric ring (89,101). The SsoMCM AAA+ core domain was isolated and had helicase activity on otherwise “non-active” substrates, such as blunt end duplex and 5' overhang DNA substrate. Reconstitution of this AAA+ core and N-terminal domains *in vitro* resulted in a formation of an active complex with increased processivity and substrate specificity (89). The reconstitution of the AAA+ core domain and the N-terminal domain with β hairpin mutation completely abolished processivity. Thus it was suggested that the N-terminal domain may be acting as a clamp that stabilizes the AAA+ C-terminal domains on DNA substrates, and also playing a role

in discriminating the substrate thus acting as a regulator (89). However the deletion of 112 amino acids in AfuMCM (Δ N112) causes AfuMCM to have robust ATPase activity and increased processivity (76). Interestingly, on the forked substrate with a shorter duplex region, the activity of Δ N112 was much weaker compared to the wild-type AfuMCM, but as the length of duplex region was increased, Δ N112 activity increased better than the wild-type. In all archaeal MCMs tested, N-terminal domain alone binds to DNA with less affinity than the full length wild-type protein (72,81,89,90,101). In MthMCM, the deletion of the N-terminal domain (Δ N) abolishes helicase and ATPase activity, suggesting that it is essential for the activity (72).

The structure of the N-terminal domain revealed three distinct domains in MthMCM and SsoMCM. Since it is the least conserved, it has been suggested that domain A may have a regulatory role. In some archaea, such as *A. fulgidus*, domain A does not exist (see Figure 1-4). However, the deletion of domain A results in decreased DNA binding, ATPase and helicase activity in MthMCM, while the equivalent mutant in SsoMCM has similar level of ATPase and helicase activity compared to the wild-type (81,89). Cryo-EM revealed that long dsDNA wraps around MthMCM in the outer surface in the clefts of domain A and domain C (107). Deletion of domain A abolished binding of MCM to long dsDNA (200mer) suggesting that domain A may be responsible for binding to long dsDNA. The binding of the wild-type protein to the long dsDNA caused bending with an angle of $\sim 100^\circ$, possibly causing the dsDNA to be distorted. Thus an analogy of a possible evolutionary link was made between the archaeal MCM and viral helicases that have origin binding and melting functions as well as an unwinding function (108).

There are several residues that are characterized in domain A. The mutation of Pro 62 to Leu (P62L) in MthMCM was characterized due to an analogous mutant found in *Saccharomyces cerevisiae* Mcm5-*BOB1* mutant (P82L). Mcm5-*BOB1* mutant bypasses the activation of Dbf4 dependent Cdc7 kinase (DDK) that allows only one replication per cell cycle during the initiation stage of DNA replication (59,109,110). According to the crystal structure of the N-terminal portion of MthMCM, Pro62 is located at the interface of domains A and C and closely interacts with F109 found in domain C. The ATPase activity of P62L did not show a significant difference, but helicase and DNA binding activity resulted in a slight decrease as compared to the wild-type (110). Structural analysis of P62L mutant revealed only a slight increase in the distances between domain A and C, causing domain A to be pushed out (59). Using MthMCM structure as analogy, it was suggested that the mutation of P62 to any bulky side chain imposes a steric hindrance causing “domain push” allowing DDK independent activation. The *in vivo* assay using *S. cerevisiae* supports this hypothesis, since mutating P82 to smaller hydrophobic residues did not allow activation of MCM5 in the absence of DDK, but mutating it to larger hydrophobic residues allowed the bypass (59). Another mutation of a conserved residue found in the linker between domain A and domain C, Arg 99 to Ala (R99A) was characterized (102). This mutation alone did not cause a significant difference in the ATPase nor helicase activity, but caused slightly decreased DNA binding.

Deletion of domain B (Δ B) abolishes ssDNA binding and ATPase activity but retains helicase activity for MthMCM (81). Domain B contains a Zn binding motif conserved in all MCM proteins. Mutation of one of the four Cys to a Ser results in wild-

type level of binding to ssDNA in the presence of ATP in both gel shift and nitrocellulose filter binding assays (61). In the absence of ATP, however, the Zn mutant shows a drastic decrease in binding to ssDNA. Despite the observed binding in the presence of ATP, the Zn mutant abolished helicase activity or stimulation of ATPase activity in the presence of DNA. Although the Zn motif may not be the sole region in which MCM binds DNA, the coordination of the AAA+ catalytic activity and the DNA binding function may be disrupted. EM images of the ΔB proteins revealed mostly filaments, compared to the wild-type or ΔA (85). When the structure of N-terminal MthMCM is viewed from the top, the Zn motif is twisted such that it is aligned to the domain C of the next neighboring subunit. It is possible that domain B plays an important role in orienting the DNA contacting regions along the helical twist as it moves along the DNA.

Domain C is most likely involved in the multimerization of MthMCM, since the deletion of domain C (ΔC) results in a monomer according to the gel filtration study (81). Yeast two hybrid analysis also revealed that the deletion of domain C prevents self-interaction of MthMCM, suggesting that it does not form oligomers (81). The ΔC mutant abolishes helicase and ATPase activity as well as DNA binding activity. Domain C is likely to support the structural integrity of oligomeric state of MCM helicase which may be crucial for the helicase mechanism. Domain C contains a β -hairpin finger, which protrudes into the central cavity of the MCM helicase ring. The mutation of the positively charged Arg and Lys residues to Ala abolishes MthMCM binding to DNA, but decreases the affinity for DNA in SsoMCM by half (81,90,92,101). As expected, this β -hairpin mutant protein does not have the helicase and ATPase activity in MthMCM, but the activity is observed in SsoMCM.

Domain C contains highly conserved residues not only among archaea, but also in eukaryotes. In particular, the conserved residues found in the loop between $\beta 7$ and $\beta 8$ mapped in MthMCM have been characterized ((103); see Chapter 2). When this loop is mapped in the EM structure, it is located at the interface between the N-terminal domain and the neighboring AAA+ domains of the neighboring molecule. Mutations in the loop cause a loss of ATPase and helicase activities while DNA binding function is retained. Thus it was suggested that the conserved loop plays an essential role in coupling the catalytic activity of the AAA+ domain and N-terminal DNA binding function.

1.8.2.2 AAA+ domain

The AAA+ domain alone has been successfully isolated, purified and characterized for SsoMCM (89). This AAA+ core domain exists as a monomer at lower concentration but forms multimers at higher concentration, similar to the N-terminal domain (89). The AAA+ domain has ATPase activity and robust helicase activity with less selectivity when various substrates, including blunt end, 3' overhang DNA substrates were used.

As described earlier, AAA+ domain of MCM contains several motifs that are highly conserved. The PS1BH motif is essential for DNA translocating mechanical motion in other helicases as well as it is in MCM helicase (90,99). Mutation of a positively charged residue found in the PS1BH motif to Ala (K430A) caused decreased DNA binding and ATPase activity and abolished helicase activity in SsoMCM (90). The role of another conserved β hairpin motif found in PS-II clade, H2-ins has been characterized in MthMCM (102). The deletion of 15 amino acids that comprise the H2-ins ($\Delta h2-i$) resulted in significantly enhanced ATPase activity in the presence of dsDNA

while ssDNA had no effect, suggesting possible changes in the diameter of the central cavity (102). The $\Delta h2-i$ mutant protein was also completely devoid of helicase activity, but exhibited enhanced binding to both ssDNA (about 60 fold higher) and dsDNA compared to the wild-type. Thus it was suggested that H2-ins may serve as a plough that turns over one strand from duplex DNA.

It has been reported that WA, WB, S-I and S-II coordinate in *cis*, while the R-finger coordinates to the neighboring subunit in *trans* to assist ATP binding and hydrolysis in typical AAA+ proteins (Figure 1-2 A) (67). The crystal structures of MkaMCM2 and SsoMCM predict that the S-II positions in *trans* in MCM proteins (Figure 1-2 B). To support this biochemically, *in vitro* complementation study on SsoMCM with mutants of WA, WB, R-finger, S-I, S-II was performed (100). The complementation study revealed that the S-II in SsoMCM is acting in *trans*. The physical interaction of S-II to neighboring subunits was confirmed by the cross linking experiments.

1.8.2.3 C-terminal HTH domain

The last ~100 amino acid residues that give rise to the C-terminal domain have been suggested to fold into a HTH (helix-turn-helix) motif (57). Similar to the N-terminal and AAA+ core domains, the C-terminal HTH domain also exists as monomer at lower concentration, but seems to form multimers at higher concentration for SsoMCM (89). This HTH domain by itself does not bind to DNA, nor does it have ATPase activity in SsoMCM. When it is expressed with AAA+ core domain, helicase and ATPase activity is detected (89,101). Deletion of HTH domain (-HTH) in SsoMCM results in significant increase in ATPase and helicase activity (89). Similar observation was made

in MthMCM, where the deletion of the last 69 residues (ΔC) results in an increased processivity and ATPase activity (~ 2 fold) in the presence of dsDNA (102). Interestingly, the level of helicase activity was compatible with the wild-type, suggesting that the C-terminal HTH domain may be involved in allosteric regulation of MCM activity, as seen in the AfuMCM $\Delta N112$ mutant. The $-HTH$ and ΔC mutation were combined with mutations in the N-terminal domain, ΔA and R99A in MthMCM and SsoMCM respectively ($-A/HTH$ and $RA\Delta C$). The double mutant, $-A/HTH$ exhibits similar level of activity (slightly less active) as seen in $-HTH$ mutant in SsoMCM (111). However, in MthMCM, the $RA\Delta C$ mutant causes a significant increase in helicase activity and even more increase in processivity compared to the ΔC mutant. However, the ATPase activity of this mutant protein is approximately the same as the wild-type level (102). Interpretations of these results can be complicated, but it appears that MCM proteins are regulated in multiple levels, including via a long range allosteric regulation.

A location of the C-terminal HTH domain in the structure has been proposed (84,87,90). Originally, the C-terminal HTH domain was suggested to be located in the cleft created by N-terminal and AAA+ domains, though there was no clear electron density in the actual EM data (84). More recently, electron density of C-terminal HTH domain has been observed for MthMCM by cryo-EM. The C-terminal HTH domain was shown to be located adjacent to the AAA+ core domain, opposite of the N-terminal double hexamer interface (87). Interestingly, the electron density map of the HTH shows obvious differences in dsDNA bound or ADP-AlF₄⁻ bound states, and two different conformations (of opposite molecule) in the same double hexameric unit.

1.8.3 Active structure of MCM

An unusual number of oligomeric states is observed in MthMCM according to EM study. The active form of the MCM helicase remains in question. In the original study, the monomer fraction from glycerol gradient had stronger ATPase and weaker helicase activities compared to the double hexameric fraction in MthMCM (71). Double hexamer, rather than single hexamer has been suggested to be the active form in MthMCM (112).

As polymorphism observed in the EM structures, the oligomeric states of MthMCM seem to be dynamic also in solution as described. Oligomeric states of MthMCM protein vary depending on the protein concentration, temperature, and salt concentrations (113). The results indicate that MthMCM is active as a single hexamer rather than the double hexamer, contrary to previously published results.

Other than experimental conditions, other factors may influence the oligomeric state of MCM protein. The inhibition of MthMCM by Cdc6-2 is possibly caused by a dissociation of MCM oligomers (114). Gel filtration analysis of MthMCM in the absence and presence of Cdc6 revealed that the MCM form multimer and monomer respectively. According to the EM structure, the MCM form stable double heptamer in the presence of nucleotide, but forms stable double hexamer in the presence of dsDNA (87). This suggests that other than protein, cofactors and substrates may play a role in regulating the oligomeric state of MCM. Moreover, some archaeal MCM require other proteins for better helicase or ATPase activity. As described earlier, PfuMCM is a stable monomer even in the presence of ATP or MgCl₂ but has weak helicase activity (78). Since the helicase and ATPase activity is enhanced in the presence of GINS, it is possible that the

GINs may regulate the oligomeric state of PfuMCM.

A large excess archaeal MCM protein, relative to the amount needed for two replication forks has been detected in the cell. A euryarchaeon, *Pyrococcus abyssi*, has been estimated to contain about 200-400 molecules per cell, *M. thermautotrophicus* and *T. acidophilum* were estimated to have 300 and 1000 molecules per rapidly growing cell respectively (29,75). This excess amount of MCM molecules per cell is also observed in Eukaryotes, and has long been known as a MCM paradox since the role of excess MCM is unclear (40,48). It is not yet clear whether the number of MCM proteins present in archaeal cells plays a role in regulating oligomeric structure.

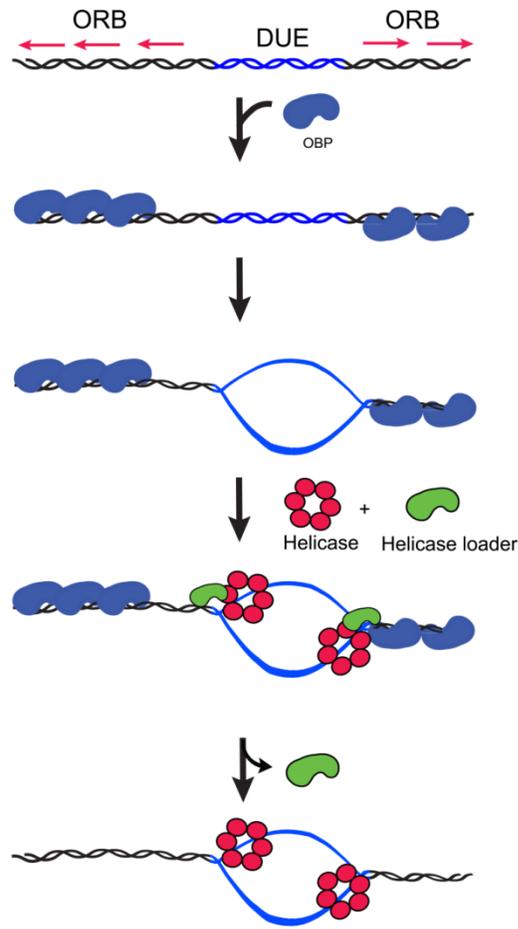


Figure 1-1 Schematic diagram for initiation of DNA replication.

Archaeal origins consist of inverted repeats known as origin recognition boxes (ORB) shown in red arrows, and A/T rich region known as DNA unwinding elements (DUE) shown in blue. Multiple origin binding proteins (OBP) shown in blue bind to ORB, and distort DNA and create a replication bubble. Helicase loader (green) loads helicase (red ring) onto the replication bubble. Helicase loader leaves the origin and helicase, and helicase continues to unwind the duplex DNA.

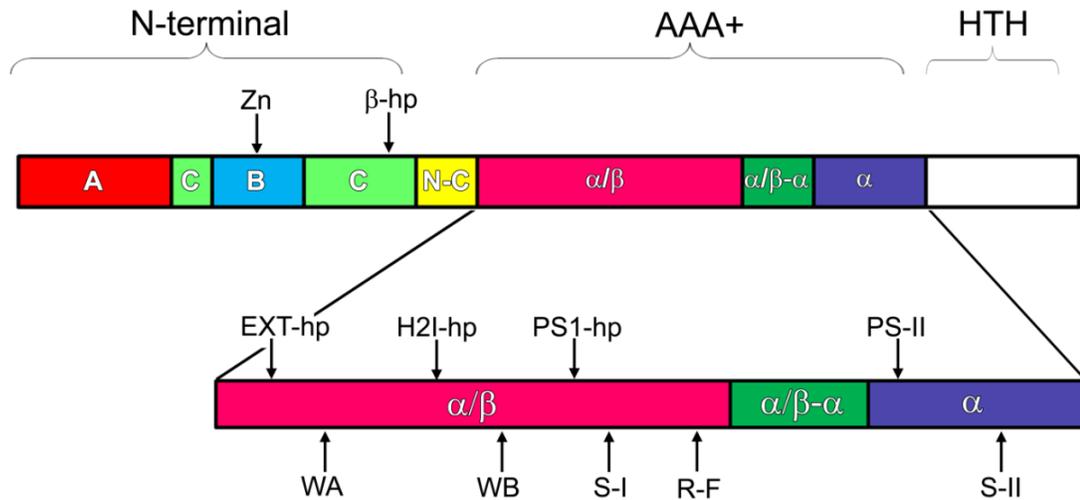
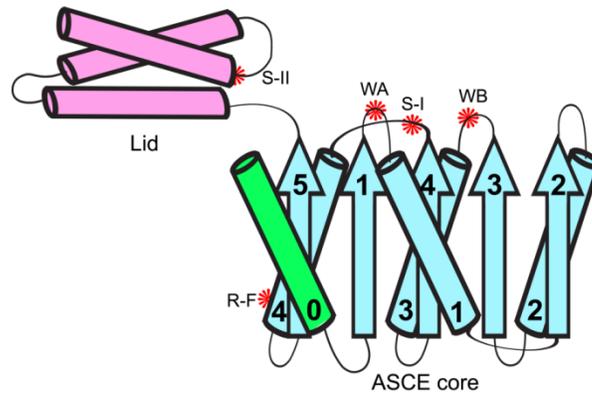


Figure 1-2 Schematic representation of the structural motifs of the archaeal MCM helicase.

Archaeal MCM helicase is divided into three major portions, N-terminal, AAA+, and HTH regions. The N-terminal portion consists of three domains, A (red), B (blue), and C (light green), and contains a Zinc finger motif (Zn) and beta-hairpin finger motif (beta-hp). The N-terminal and AAA+ portions are connected via the N-C linker (yellow). The AAA+ region consists of an alpha/beta domain (pink) connected to an alpha bundle domain (dark blue) via the alpha/beta-alpha linker (dark green). The AAA+ region contains common motifs that AAA+ proteins share, and also are unique to fewer proteins including MCM helicase (see text for details). Common motifs include WA (Walker A), WB (Walker B), S-I (Sensor 1), R-F (Arg finger), and S-II (Sensor 2), shown in the bottom of enlarged AAA+ domain. MCM contains several motifs that are in addition to the core motifs, such as EXT-hp (external hairpin), H2-I (helix-2 insert), PS1-hp (pre-sensor 1 beta-hairpin), and PS-

II (pre-sensor 2 helix) shown above enlarged AAA+ region. The C-terminal HTH domain is shown in white.

A) Typical AAA+ organization



B) MCM helicase organization

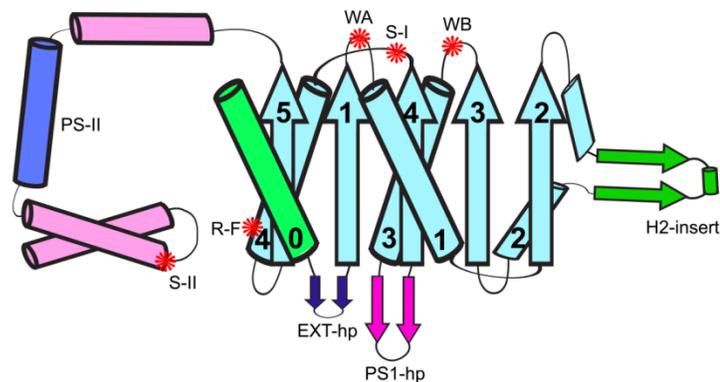


Figure 1-3 Organization of AAA+ proteins.

A typical AAA+ protein contain an ASCE core (shown in cyan) and a lid or α bundle domain (shown in light pink) (A). Key motifs identified in AAA+ proteins are marked with red stars. WA=Walker A, WB=Walker B, S-I=Sensor 1, S-II=Sensor 2, R-F=Arg finger. MCM helicase contain additional motifs as shown in B, EXT-hp=external hairpin (dark blue), PS1-hp=pre-sensor 1 β -hairpin (hot pink), H2-ins=Helix 2 insertion (green), and PS-II=pre-sensor 2 helix (light blue). The figure has been modified from (67).

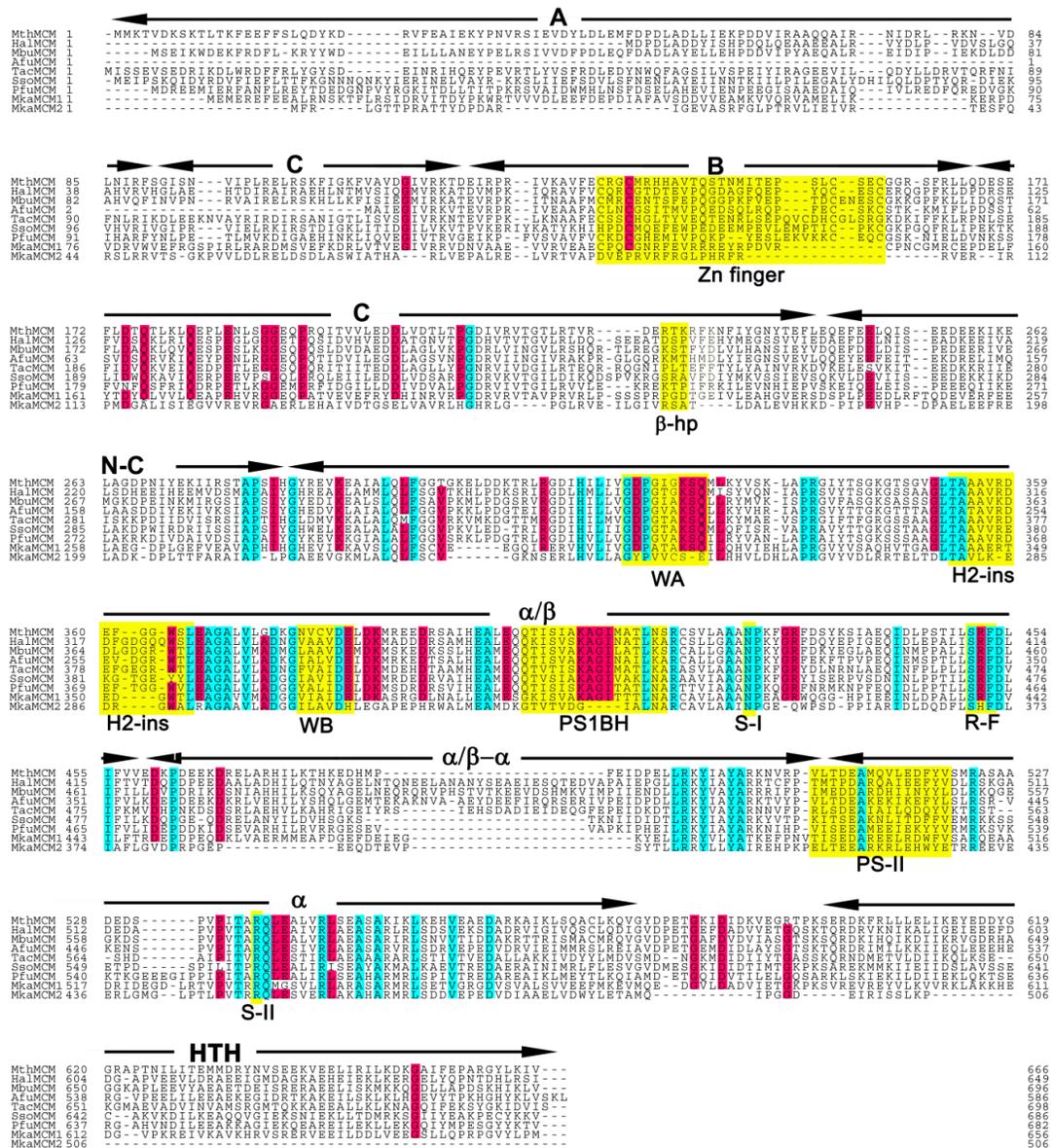
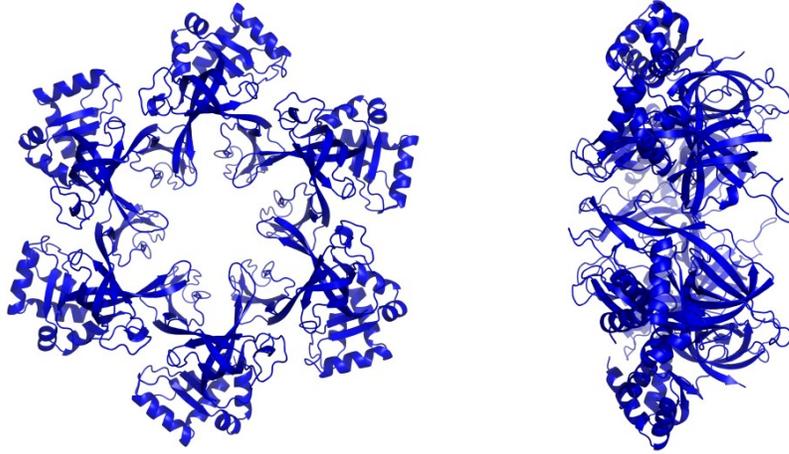


Figure 1-4 Alignment of archaeal MCM helicase that have been characterized *in vitro*.

Nine archaeal MCM helicases that have been studied to date are pooled and aligned using MUSCLE program (EBI website). Residues with 100% identity are represented in cyan and >80% are in pink. Identified domains as well as key motifs described in the text are labeled. For clarity, intins from HalMCM and AfuMCM are omitted. Domains are

represented in the space above the sequence. Common motifs are highlighted in yellow and labeled in the space below.

A) *M. thermautotrophicus* N-terminal



B) *S. solfataricus* N-terminal

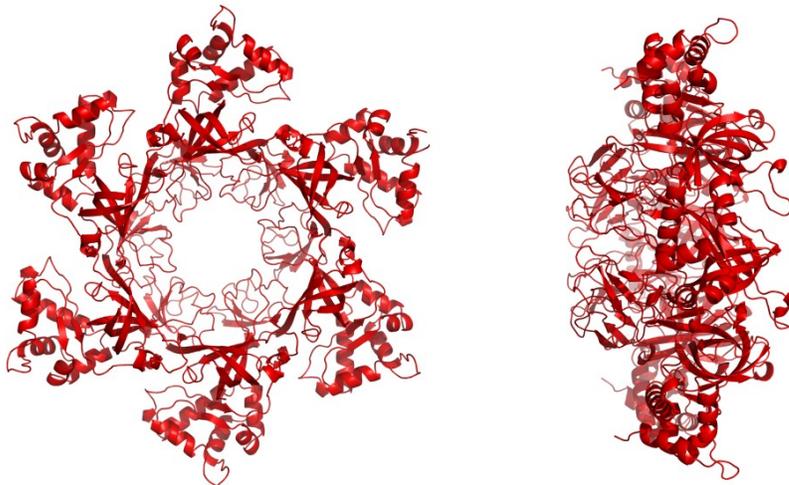
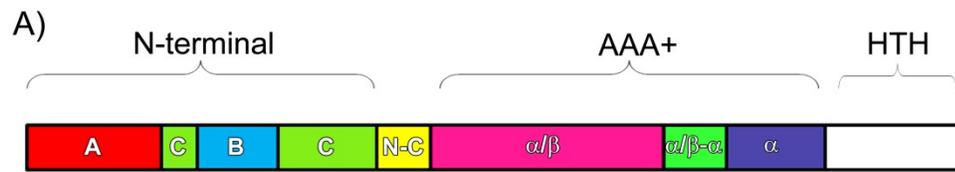
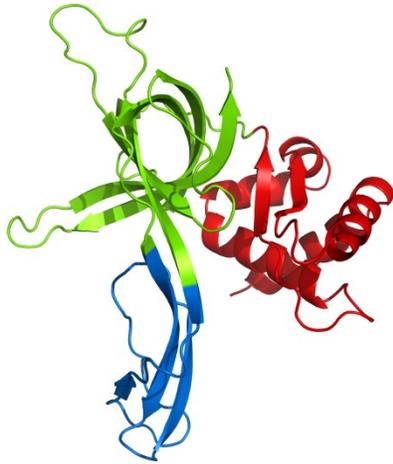


Figure 1-5 Structures of N-terminal portions of MCM helicases.

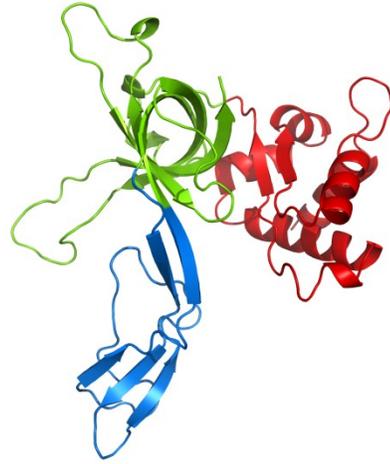
Crystal structure of N-terminal portion of MCM from *M. thermautotrophicus* is shown in Blue (top) (PDB ID: 1LTL) (59), and from *S. solfataricus* is shown in red (bottom) (PDB ID: 2VL6) (60). PyMOL was used to generate the structures.



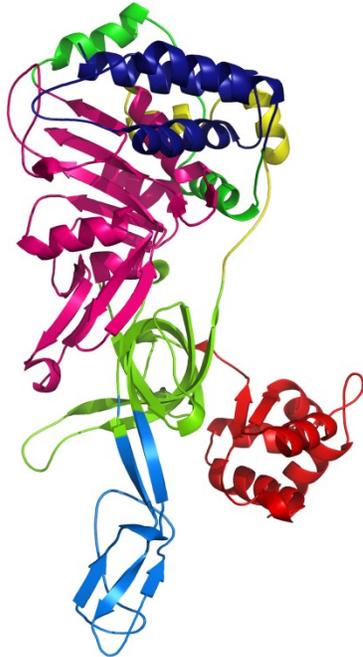
B) *M. thermoautotrophicus*
N-terminal MCM protein



C) *S. solfataricus*
N-terminal MCM protein



D) *S. solfataricus* full-length
MCM protein



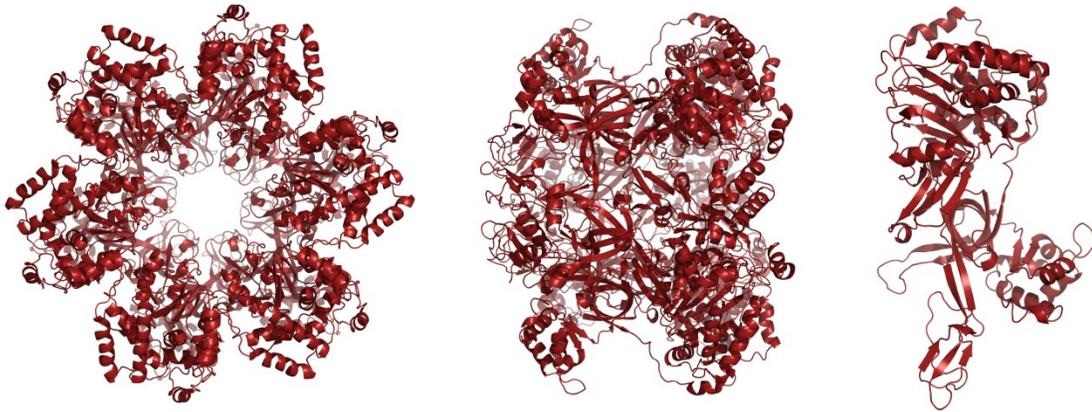
E) *M. kandleri* full-length
MCM protein



Figure 1-6 Structures of monomeric forms of MCM helicases.

A) Schematic representation of the structural motifs of MCM helicase. Domains are highlighted with different colors and are also represented in the structures. N-terminal portions of MthMCM and SsoMCM are shown in B and C respectively and the structures of the monomeric SsoMCM and MkaMCM2 are shown in D and E. The domains of MkaMCM2 were defined based on the alignment with SsoMCM. N-terminal domain A is in red, domain B is in blue, and domain C is in light green. The N-C linker that connects N-terminal and AAA+ region is shown in yellow, α/β domain is in pink, α/β - α linker is in green, and α bundle domain is in dark blue.

A) *S. solfataricus*



B) *M. Kandleri*

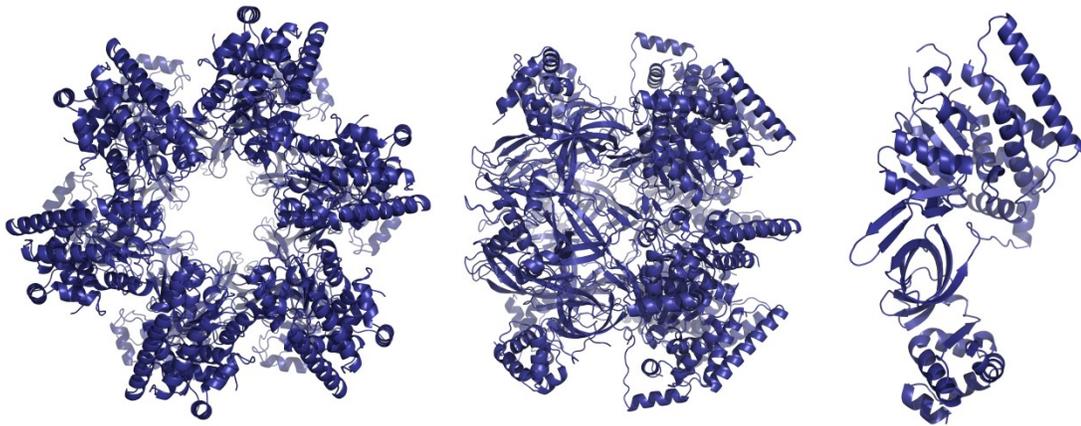


Figure 1-7 Models of hexameric structures of MkaMCM2 and SsoMCM proteins

Hexameric models of SsoMCM (A) and MkaMCM2 (B) are shown in the left (top view) and the center (side view). The modeled MkaMCM2 consists of the N-terminal portion of MthMCM and the AAA+ domain from MkaMCM2. Monomeric structures without any modification are shown in the right.

Chapter 2 Coupling of DNA binding and helicase activity is mediated by a conserved loop in the MCM protein

2.1 Abstract

Minichromosome maintenance (MCM) helicases are the presumptive replicative helicases, thought to separate the two strands of chromosomal DNA during replication. In archaea, the catalytic activity resides within the C-terminal region of the MCM protein. In *Methanothermobacter thermautotrophicus* the N-terminal portion of the protein was shown to be involved in protein multimerization and binding to single and double stranded DNA. MCM homologues from many archaeal species have highly conserved predicted amino acid similarity in a loop located between $\beta 7$ and $\beta 8$ in the N-terminal part of the molecule. This high degree of conservation suggests a functional role for the loop. Mutational analysis and biochemical characterization of the conserved residues suggests that the loop participates in communication between the N-terminal portion of the helicase and the C-terminal catalytic domain. Since similar residues are also conserved in the eukaryotic MCM proteins, the data presented here suggest a similar coupling between the N-terminal and catalytic domain of the eukaryotic enzyme.

2.2 Introduction

The minichromosome maintenance (MCM) complex is thought to function as the replicative helicase of archaea and eukarya (115,116). In eukaryotes the MCM complex is a family of six essential polypeptides (Mcm2–7) with highly conserved amino acid sequences. Biochemical studies have shown that a dimeric complex of the Mcm4,6,7 heterotrimer possesses 3'-5' DNA helicase activity, can translocate on single and double stranded DNA, can bind DNA and RNA, and has the ability to unwind DNA-RNA

duplexes while translocating on the DNA strand (40,41). The interactions of Mcm4,6,7 with either Mcm2 or Mcm3,5 were shown to inhibit helicase activity and therefore were suggested to play regulatory roles (40,41).

Most archaeal species examined contain a single MCM homologue (115,116) with biochemical properties similar to the eukaryotic Mcm4,6,7 complex. The archaeal MCM proteins were shown to contain 3'-5' DNA helicase activity, bind and translocate along ss and dsDNA, unwind DNA-RNA duplex substrate while translocating along the DNA, and to displace proteins from DNA ((93), and references therein).

The MCM helicases are divided into a C-terminal portion, which contains the helicase catalytic domain, and a N-terminal region (59,81,84). To date, a high-resolution structure has been determined only for the N-terminal portion of the MCM protein from the archaeon *Methanothermobacter thermautotrophicus* (59). That structure revealed a dumbbell-shaped double hexamer. Each monomer folds into three distinct domains. Domain A, at the N terminus, is mostly α -helical. Domain B has three β -strands and contains a zinc-finger motif shown to be needed for DNA binding (61,117). Domain C contains five β -strands that form an oligonucleotide/oligosaccharide binding (OB) fold. This domain connects the N-terminal portion of the enzyme to the C-terminal catalytic region. Domain C contains a β -finger motif shown to be essential for ss and dsDNA binding (59,117). The domain was also shown to be necessary for MCM multimerization (81). Sequence alignment of MCM proteins from many archaeal species has revealed highly conserved residues in a loop between β 7 and β 8 in domain C (Figure 2-1 and Figure 2-2). Based on the crystal structure of the N-terminal part of the molecule, the loop is located in the opposite side of the dimer interface between the two hexamers

(Figure 2-3).

Electron micrograph (EM) reconstruction of the full-length *M. thermautotrophicus* MCM helicase (84,88) also suggest that the loop is in close proximity to the catalytic domain located at the C-terminal part of the molecule (Figure 2-4). Loop regions are known to be less conserved than other secondary structures unless they play an important functional role.

The high conservation suggests that the loop between $\beta 7$ and $\beta 8$ may play a role in MCM function. Biochemical characterization of proteins harboring mutations in this region suggest that the loop region is likely to be involved in coupling the N-terminal multimerization and DNA binding domains with the C-terminal catalytic domain. Thus the loop may function as a bridge, allowing a movement between the two domains to transmit a signal.

2.3 Materials and Methods

2.1.1 Materials

ATP, [γ - 32 P]ATP and [α - 32 P]ATP were obtained from GE Healthcare, and oligonucleotides were synthesized by the CARB DNA synthesis facility. All proteins used in the study were purified as previously described (81) and are derived from the full-length MCM gene.

2.1.2 Multiple alignment

The *M. thermautotrophicus* MCM protein sequence (MTH1770) was aligned using BLAST against 41 archaeal genomes (National Center for Biotechnology

Information, NCBI). Full length MCM sequence relatives with expectation scores $<10^{-5}$ from the 41 archaeal genomes were pooled and aligned using the MUSCLE program and default parameters. Aligned sequences were loaded onto Jalview 2.2.1, and the N-terminal portion (MTH1770 residues 1 – 244) was kept for the following analysis. PHYLIP promlk (version 3.6) was used to build the maximum likelihood phylogenetic tree, which resulted in four subgroups (Group I – IV) as shown in Figure 2-5. From the tree, a total of 21 MCM proteins from subgroup I (List of subgroup I with accession number: *Methanothermobacter thermautotrophicus*, NP_276876; *Thermoplasma acidophilum*, NP_394261; *Thermoplasma volcanium*, NP_111551; *Picrophilus torridus*, YP_023995; *Ferroplasma acidarmanus*, ZP_01708764; *Haloquadratum walsbyi*, YP_659323; *Haloarcula marismortui*, YP_137231; *Natronomonas pharaonis*, CAI50644; *Methanocorpusculum labreanum*, YP_001030268; *Methanospirillum hungatei*, YP_502454; *Methanoculleus marisnigri*, YP_001047160; *Methanosarcina acetivorans*, NP_618700; *Methanosarcina barkeri*, YP_305120; *Methanosarcina acetivorans*, NP_615641; *Methanosarcina mazei*, NP_633860; *Methanococcoides burtonii*, YP_567033; *Methanosaeta thermophila*, YP_842696; *Methanosphaera stadtmanae*, YP_447408; *Sulfolobus tokodaii*, NP_376352; *Archaeoglobus fulgidus*, NP_069353; *Halobacterium*, NP_280836) that contain the *M. thermautotrophicus* MCM sequence (MTH1770) were selected to view as an alignment (Figure 2-1).

2.1.3 Expression and purification of MCM mutant proteins

All *M. thermautotrophicus* MCM mutant proteins used in this study are derivatives of the full-length enzyme and were generated using PCR-based site-directed

mutagenesis as previously described (81). All constructs contain a C-terminal His₆-tag and were cloned into the NdeI and XhoI sites of the pET-21a vector (Novagene). The oligonucleotides used for the mutagenesis are shown in Table 1. The wild-type and mutant proteins were overexpressed in codon plus cells (Stratagene) at 37°C and purified on a Ni-column as previously described (104).

2.1.4 ATPase assay

ATPase activity was measured in reaction mixtures (15 µl) containing 25 mM Hepes-NaOH (pH 7.5), 1 mM dithiothreitol, 5 mM MgCl₂, 100 µg/ml bovine serum albumin (BSA), and 1500 pmol of unlabeled ATP and [γ -³²P]ATP (3000 Ci/mmol; GE Healthcare) and 10 or 30 nM MCM protein (as monomer) in the presence or absence of 50 ng ssDNA (5'-GGCAGATAACAGTTGTCCTGGAGAACGACCTGGTTGACACCCTCACACCC - 3'). After incubation at 60°C for 60 min, samples were placed on ice, then an aliquot (1 µl) was spotted onto a polyethyleneimine cellulose thin layer plate, and ATP and P_i were separated by chromatography in 1 M formic acid and 0.5 M lithium chloride. The extent of ATP hydrolysis was quantitated by phosphorimager analysis. All ATPase assays were repeated three times, and their averages with standard deviations are shown in Figure 2-2.

For the kinetics measurement time-based ATPase assays were performed. ATPase activities were measured in reaction mixtures (50 µl) containing 25 mM Hepes-NaOH (pH 7.5), 1 mM dithiothreitol, 5 mM MgCl₂, 100 µg/ml BSA, and 50 nM of MCM protein (as monomer) in the presence or absence of 50 ng ssM13 (New England Biolabs)

with various ATP concentrations (0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 1 or 1.4 mM) containing 50 nM of [γ - 32 P]ATP (3000 Ci/mmol; GE Healthcare) at 60°C.

At 0, 20, 40, and 80 minute time points, 3 μ l aliquots were taken out and quenched on ice. At all ATP concentrations, the reactions were determined to be in steady state at about 40 min (data not shown). Thus, a single time point measurement at 30 min was performed to determine the K_M and V_{max} with various ATP concentrations (0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 1 and 1.4 mM). The kinetic parameters were calculated using GraFit version 5.0.1. (Erithacus software) and the Michaelis-Menten equation, where graphs were fit to $v = \frac{V_{max} \cdot [S]}{K_m + [S]}$; where v is the rate of reaction in $\mu\text{M} \cdot \text{min}^{-1}$, S is the substrate concentration in μM , V_{max} is the maximum reaction rate of the enzyme, and K_M is the Michaelis-Menten constant. The turnover rate, k_{cat} , was determined by dividing the V_{max} by the total amount of MCM protein in the reaction (50 nM). Several data resulted in non-informative values for K_M and V_{max} due to the fact that the standard errors were greater than 90% of the values calculated. All experiments repeated three times and the averages are shown in Figure 2-10 B-D.

2.1.5 Nitrocellulose filter DNA binding assay

Single stranded and dsDNA substrates for filter binding assays were prepared by labeling the oligonucleotide using [γ - 32 P]ATP and T4 polynucleotide kinase. Unincorporated [γ - 32 P]ATP was removed from the DNA substrate by extraction from polyacrylamide gel as previously described (118). For ssDNA binding a 49mer oligonucleotide with the sequence 5' - GCAGATAACAGTTGTCCTGGAGAACGACCTGGTTGACACCCTCACACCC-3'

was used. For dsDNA the same oligonucleotide was hybridized to its complementary sequence.

Filter binding assays were carried out in a reaction mixture (20 μ l) containing 25 mM Hepes-NaOH (pH 7.5), 2 mM dithiothreitol, 10 mM $MgCl_2$, 100 μ g/ml BSA, 50 fmol of ^{32}P -labeled ss or dsDNA substrates and 10, 30, or 90 nM of MCM protein (as monomer).

After incubation at 60°C for 10 min the mixture was filtered through an alkaline-washed nitrocellulose filter (Millipore, HA 0.45 μ m) (119) which was subsequently washed with 20 mM Tris-HCl (pH 7.5). The radioactivity adsorbed to the filter was measured by liquid scintillation counting. All DNA binding experiments were repeated three times, and their averages with standard deviations are shown in Figure 2-9 A and B.

2.1.6 Fluorescence polarization anisotropy (FPA) measurement

Fluorescence anisotropy measurements were performed at 25°C using a Fluoromax-3 spectrofluorimeter equipped with an autopolarizer (Jobin Yvon Inc.), using a 3 mm path length cuvette with starting volume of 150 μ l. A 50-mer ssDNA (5'-CGCAGATAACAGTTGTCCTGGAGAACGACCTGGTTGACACCCTCACACCC -3') was 5'-labeled with Cy3 and purified with a HPLC C18 column. The concentration of DNA was calculated using an absorbance of 260 nm with an extinction coefficient of 477300 $M^{-1}\cdot cm^{-1}$ and absorbance at 546 nm using extinction coefficient 136,000 $M^{-1}\cdot cm^{-1}$ for Cy3 dye. DNA concentrations calculated by both measurements were different by less than 10%. The measurements using absorbances at 260 nm were used to calculate the concentrations for the experiments. The initial reaction mixture contained 25 mM Hepes-NaOH (pH 7.5), 2 mM DTT, 5 mM $MgCl_2$ and 10 nM DNA, with or without 1

DNA helicase activity assays were measured in reaction mixtures (15 μ l) containing 20 mM Tris-HCl (pH 8.5), 2 mM DTT, 10 mM MgCl₂, 100 μ g/ml BSA, 3.33 mM ATP, 10 fmol (0.66 nM) of ³²P-labeled substrate and 10, 30 or 90 nM MCM protein (as monomer) for flat substrate or 3, 9, or 27 nM MCM protein (as monomer) for forked substrate. Following incubation at 60°C for 1 hr the reactions were stopped by adding 5 μ l of loading buffer containing 1% SDS, 100 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue, and 50% glycerol and chilling on ice. Aliquots (10 μ l) were loaded onto an 8% native polyacrylamide gel in 0.5X TBE and electrophoresed for 40 min at 180V. Gels were visualized and quantitated by phosphorimaging. All helicase experiments were repeated three times, and their averages with standard deviations are shown in Figure 2-6.

The relative rates of unwinding were measured in reaction mixtures (140 μ l) containing 20 mM Tris-HCl (pH 8.5), 2 mM DTT, 10 mM MgCl₂, 100 μ g/ml BSA, 3.33 mM ATP, 93.33 fmol (0.66 nM) of ³²P-labeled substrate and 30 nM proteins for flat substrate and 10 nM proteins for forked substrate. Following incubation at 60°C samples (10 μ l) were removed at 0, 5, 10, 15, 30, 45 and 90 min, quenched with 5 μ l of loading buffer containing 1% SDS, 100 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue, and 50% glycerol, and put on ice. Aliquots (10 μ l) were loaded onto an 8% native polyacrylamide gel in 0.5X TBE and electrophoresed for 40 min at 180V. Gels were visualized and quantitated by phosphorimaging. All helicase experiments were repeated three times, and their averages with standard deviations are shown in Figure 2-7. The slope of initial rate was used to calculate the relative rate of unwinding and it is presented as μ moles of DNA unwound per moles of protein per minute (μ mol \cdot mol⁻¹ \cdot min⁻¹).

2.1.8 Gel filtration analysis

One hundred micrograms of wild-type and mutant MCM proteins were applied to a superose-6 gel filtration column (HR10/30, GE Healthcare) pre-equilibrated with buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 10% glycerol. The column was run at a flow rate of 0.2 ml/min at room temperature and 250 μ l fractions were collected. The presence of protein was determined by ultraviolet absorbance at 280 nm.

2.1.9 Differential scanning calorimetry (DSC) measurements

The purified proteins were dialyzed at room temperature against buffer containing 20 mM potassium phosphate (pH 7.5), 100 mM NaCl and 10% glycerol. The protein concentrations were determined using an absorbance at 280 nm and a calculated extinction coefficient of ϵ_{280} of $28730 \text{ cm}^{-1} \cdot \text{M}^{-1}$. The solution outside the dialysis bag was retained and used as reference for the experiments.

DSC measurements were performed using a VP-DSC Microcalorimeter from Microcal Inc. (Northampton, MA). The volume of the solution and the reference vessels were 0.511 ml and the scan was either at slow rate (15 K hr^{-1}) or medium rate (60 K hr^{-1}), with the temperature ranging from 25 – 85°C. Since the scans of protein samples were irreversible, the second scan for each sample was used as the baseline. After subtraction of the baseline from the protein scan, the resulting raw data of differential powers as a function of time were divided by the scan rate to convert into a heat capacity as a function of temperature. Utilizing the EXAM program (122), a two state, $A \rightleftharpoons B$, transition model was then fitted to the heat capacity as a function of temperature scan to determine the van't Hoff enthalpy (ΔH_{VH}) for the scan from the shape of the transition peak; a transition temperature (T_m), and a calorimetric transition enthalpy (ΔH_{cal}) was

calculated from the area under the transition peak (mJ) divided by the moles of protein in the sample cell (concentration of protein x 0.511 ml). DSC scans on the samples were repeated twice.

2.1.10 Circular dichroism

Circular dichroism (CD) measurements were performed with a Model J-720 Jasco Spectropolarimeter using quartz cells with a 0.005 cm path length at room temperature. The proteins used for CD were prepared as described above for the DSC experiments, with concentrations of proteins ranging from 10 – 30 μ M. Far-UV wavelength scans were recorded at room temperature from 250 to 190 nm. Averages for three CD spectra were presented. Ellipticity results were expressed as mean residue ellipticity $[\theta] = \text{degrees} \times \text{centimeter}^2 \times \text{decimole}^{-1}$.

2.1.11 UV crosslinking

The UV crosslinking experiments were performed using 250 ng of wild-type, K325A and the various loop mutant proteins as previously described (61) in a 20 μ l reaction mixture containing 3.3 pmol of [α -³²P]ATP, 20 mM Hepes-NaOH (pH 7.5), 5 mM MgCl₂, 2 mM DTT, and 250 ng BSA. The samples were incubated for 10 min at 60°C followed by exposure to 2.5 J/cm² UV irradiation in a model 2400 Stratalink (Stratagene) and analyzed by SDS-PAGE followed by Coomassie Blue staining and autoradiography. The percent of ATP cross linked was calculated using the following equation: $\% \text{ crosslinked} = \frac{I_A}{I_{A_{WT}}} \times \frac{I_C}{I_{C_{WT}}} \times 100$; where I_A and $I_{A_{WT}}$ are the intensity of radiolabeled ATP cross linked to the mutant or wild-type proteins, respectively, detected

by autoradiography and I_c and I_{cWT} are the intensity of Coomassie stained mutant and wild-type MCM proteins, respectively.

2.4 Results

2.1.12 The archaeal MCM helicase contains highly conserved residues in the N-terminal portion

Comparison of the primary amino acid sequences of the MCM helicases from a number of archaea identified highly conserved residues in the N-terminal region of all species (data not shown). Not only have most of the highly conserved residues been identified in domain C of the N-terminal portion of the molecule, but when the conserved residues are mapped on the three dimensional structure of the *M. thermautotrophicus* MCM many of them are located at a loop located between $\beta 7$ and $\beta 8$ in domain C (Figure 2-1 to Figure 2-3).

2.1.13 Mutations in the loop region between $\beta 7$ and $\beta 8$ affect MCM helicase activity

As shown in Figure 2-1, there are several highly conserved amino acids in the loop between $\beta 7$ and $\beta 8$ in the N-terminal part of MCM. As the loop is on the surface of the N-terminal part of the molecule and loops are usually not highly conserved unless they have a functional role, one may hypothesize that the conserved residues play an important role for MCM function. To evaluate whether the hypothesis is correct, conserved residues were mutated in the background of the wild-type MCM from the

archaeon *M. thermautotrophicus*. This MCM was chosen as, to date, this is the only MCM homologue with high-resolution structural information.

First, the effect of the mutation on helicase activity was determined. It was previously shown that helicase activity on forked substrate is higher than on flat substrate (104). Therefore, both substrates were used in the study but higher protein concentrations were used with the flat substrate. As shown in Figure 2-6 A and B, three mutant forms of the protein have lost the ability to unwind flat DNA substrate (E182R, E185R and P193G). When forked DNA substrate was tested (Figure 2-6 C and D) the P193G mutant protein remained inactive while E182R showed limited activity. However, on forked DNA the E185R mutant protein showed activity that is as good if not better than that of wild-type MCM at low enzyme concentrations (3 nM). Additionally, the E185R mutant protein cannot unwind more than ~70% of the substrate even at higher protein concentrations up to 270 nM (data not shown).

The experiments described above were performed using a single time point and thus cannot provide information regarding the differences in unwinding rates between the various mutant proteins. To determine the effect of the mutations on the rate of DNA unwinding, a time course experiment was performed (Figure 2-7). As was the case with the experiments described in Figure 2-3, no appreciable activity could be detected with the E182R, E185R and P193G mutant proteins when flat substrate was used. On forked substrate the E185R mutant protein exhibits helicase activity which is similar to the wild-type MCM (Figure 2-7). The L184A mutant protein has a slightly higher rate of unwinding on flat substrate compared to forked substrate (Figure 2-7 C).

Why some mutant enzymes are not active on a flat substrate is not yet known. One

possibility is that the central hole created by the N-terminal part and the catalytic domain is larger and thus more of the enzyme moves along the duplex when provided with only a 3'-overhang region as has previously been reported for the archaeal and eukaryal MCM helicases (44,123).

2.1.14 Mutation of conserved loop residues does not affect MCM structural integrity

It was previously shown that the N-terminal portion of MCM plays a major role in MCM multimerization and that domain C, in particular, is needed for MCM hexamerization (81). It was also shown that protein hexamerization is needed for helicase activity (113) and the deletion of domain C results in an inactive enzyme (81). Thus, it is possible that the substitutions in the loop, which is located in domain C, alter the structure of the MCM molecule, leading to the reduced helicase activity observed in the mutant proteins. Therefore, the oligomeric state, structural integrity and thermostability of the mutant proteins were determined using gel filtration, differential scanning calorimetry (DSC) and circular dichroism (CD).

As shown in Figure 2-8 A, the wild-type and all mutant proteins form dodecamers in solution as it was previously reported for the wild-type enzyme (71-73). Next, the overall structural integrity of the mutant proteins was evaluated using CD spectra. The results indicate that all mutant proteins are folded, retaining a secondary structure similar to the wild-type enzyme (Figure 2-8 B).

However, it is possible that the mutations do not affect the oligomeric state of the helicase but do affect its conformational stability. Because the helicase assays are performed at high temperature (60°C), it is important to determine whether the mutations

affect the stability of the enzyme at high temperature. The DSC study showed that all mutant proteins, except for P193G, have melting temperatures similar to the wild-type MCM (Table 2-2). For P193G melting is about 7°C higher than the wild-type protein suggesting a more stable structure. DSC scans of several of the mutants were performed at the slower scan rate of 15 K hr⁻¹ and it was observed that the transition temperature and enthalpies were within experimental uncertainty of the results at 60 K hr⁻¹. Manly et al. (124) have justified that those proteins that exhibit transition properties independent of scan rate could be analyzed in terms of a thermodynamic two-state transition model although the transitions do not re-appear upon a re-scan of the solutions. Thus, all mutant proteins are dodecameric in solution and appear to form intact secondary structures with melting temperatures in the range of the wild-type enzyme. Thus, gross structural alterations are not likely to be the cause for the decrease in helicase activity observed with the E182R, E185R and P193G mutant MCM proteins.

2.1.15 Mutations in the loop do not affect DNA binding by MCM

The N-terminal portion of the *M. thermautotrophicus* MCM protein was shown to contain two motifs required for DNA binding. Domain B contains a zinc-finger motif shown to participate in DNA binding, as substitution of one of the Cys residues in the motif to Ser (C158S) substantially reduced ss and dsDNA binding (61,117), and deletion of domain B also resulted in reduced DNA binding by MCM (81). Domain C contains a β -finger motif shown to be essential for DNA binding because substitution of key Arg and Lys residues in the motif to Ala (R227A and K229A) completely abolished both ss and dsDNA binding (59,117). Although the mutations in the loop between β 7 and β 8 are not in close proximity to either of the DNA binding motifs, it is possible that the changes

slightly alter the structure of the molecule and thus affect DNA binding. As proper DNA binding is needed for helicase activity, loss of DNA binding may result in the effects observed in the mutant forms of the protein. Therefore, filter binding assays were performed to evaluate the ability of the various mutants to bind ss and dsDNA (Figure 2-9). As shown in Figure 2-9, all mutant proteins bind ssDNA (Figure 2-9 A) and dsDNA (Figure 2-9 B).

To obtain a more quantitative thermodynamic comparison on the binding of the various mutant proteins to DNA and to determine whether ATP has an effect on the DNA binding, a fluorescence polarization anisotropy (FPA) analysis were performed (Figure 2-9 C-E). As a negative control a β -finger mutant protein that cannot bind DNA (59,117) was used. As was the case in previous studies, the mutant protein did not bind DNA in the FPA assay (Figure 2-9 C and D). The results show that all mutant proteins bind ssDNA with similar affinity as the wild-type MCM (Figure 2-9 E). The data also show that ATP does not affect DNA binding by MCM as the K_d values do not substantially change whether ATP is present or not (Figure 2-9 E). These results suggest that the reduced helicase activity observed with two of the mutant MCMs proteins, E182R and P193G, is not due to their inability to bind DNA.

2.1.16 Mutations in the loop region of the N-terminal domain alter the ATPase activity of MCM

The data presented in Figure 2-9 show that mutations in the loop between β 7 and β 8 do not substantially affect the ability of the proteins to interact with DNA. This suggests that DNA binding is not the reason for the impaired helicase activity observed. ATP hydrolysis is required for helicase activity, as it fuels the unwinding reaction. A

mutant protein in which a key Lys residue needed for ATPase activity is replaced by Ala (K325A) does not have ATPase activity (72,73) and is not active in helicase assays (72,73). It has also been shown that the MCM ATPase activity is stimulated in the presence of DNA (71,72). Thus, the effect of the mutations on the ATPase activity and the stimulation by DNA was determined (Figure 2-10). As a negative control the K325A mutant protein was used (Figure 2-10 A). The data show that the two mutant proteins with reduced helicase activity, E182R and P193G, also have very limited, if any, ATPase activity with no stimulation by DNA (Figure 2-10 A) and no activity was observed when three fold higher protein concentrations were used (data not shown). In addition, several mutant proteins (Q181A, L184A, E185R and G190A) showed higher ATPase activity than the wild-type enzyme (Figure 2-10 A).

To get a more quantitative picture of the ATPase activity of the mutant MCMs in comparison to the wild-type enzyme the K_M and k_{cat} were determined in the presence and absence of ssDNA (Figure 2-10 B-D). When the two mutant proteins with limited helicase activity, E182R and P193G, were analyzed, the K_M and k_{cat} values could not be determined regardless of whether ssDNA was present or not (Figure 2-10 B-D). All other proteins show an increase of k_{cat} in the presence of ssDNA but not to the same extent. Even with those differences between the mutant proteins, the ability of ssDNA to stimulate ATPase activity suggests that they are able to translocate along ssDNA. These results are consistent with the ability of these mutant proteins to unwind DNA (Figure 2-6). The observation that several mutations at the N-terminal, non-catalytic, portion of the enzyme either stimulate or inhibit ATPase activity suggests that the N-terminal part regulates the ATPase and helicase activities of the catalytic domain.

However, the lack of ATPase activity by the E182R and P193G mutant proteins could be due to their inability to bind ATP. Therefore, as a control, and in order to demonstrate the ability of the mutant proteins to bind ATP a UV-cross linking experiment was performed (Figure 2-11). This method was previously used to demonstrate ATP binding to the wild-type enzyme (61,125). As shown in Figure 2-11 all mutant proteins, including E182R and P193G, can bind ATP. The only mutant protein that did not bind ATP is K325A, as has previously been reported (61,125).

2.5 Discussion

The archaeal MCM and probably the eukaryotic enzymes are composed of two main parts in which the C-terminal region contains the catalytic activity and the N-terminal portion participates in DNA binding. However, the mutational analysis presented here suggests that the loop between $\beta 7$ and $\beta 8$ in the N-terminal part is involved in coupling the DNA binding of the N-terminal portion of the molecule to the ATPase and helicase activity of the C-terminal AAA+ catalytic region. This is based on the observation that some of the mutant proteins have DNA binding and basal ATPase activity similar to the wild-type enzyme but do not show DNA stimulation of the ATPase activity and/or helicase activity.

This is not the first example of a N-terminal mutation in the *M. thermautotrophicus* MCM that affects catalytic activity. It was shown that a Pro to Leu substitution at position 62 (P62L) affects helicase activity but not DNA binding or basal ATPase activity (110). A similar mutation in the eukaryotic Mcm5 protein (P83L) from *Saccharomyces cerevisiae*, known as the *BOB1* mutation, was shown to bypass the

requirement for the activity of the cell-cycle regulator Dbf4-dependent kinase Cdc7 required for the initiation of DNA replication (109). It was suggested that this mutation causes a conformational change in the MCM structure (59).

Another residue in the N-terminal part of the *M. thermautotrophicus* MCM protein important for the biochemical properties of the enzyme is Arg at position 99. It was shown that although substitution of Arg to Ala at that position (R99A) has no effect on basal ATPase activity it reduces the stimulation of the ATPase activity by ssDNA (102).

The data suggest that the loop between $\beta 7$ and $\beta 8$ in the N-terminal region may interact with a region on the C-terminal part of the helicase and that these interactions mediate communication between the two regions of the enzyme. Such an interaction between the loop and the catalytic domain may also be postulated from the EM studies of the full-length *M. thermautotrophicus* MCM protein (Figure 2-4) (84,88). The structure suggests that the loop between $\beta 7$ and $\beta 8$ is in close proximity to the C-terminal part of the molecule and seems to be the only interacting region between the N and C part of the molecules of MCM.

The mechanism by which the loop in the N-terminal part of MCM regulates the activity of the catalytic domain is not yet clear. However, a suggested model is shown in Figure 2-12. In this model the loop mediates communication between the two parts of the enzyme and transmits signals from the N-terminal part to the AAA+ catalytic region in the C-terminal part. These signals may include DNA binding by the N-terminal region and nucleotide binding and hydrolysis by the catalytic domain. Since a high resolution three-dimensional structure of the full length MCM protein is not yet available one must

rely on the structure of other helicases to suggest conformational changes occurring upon ATP binding and hydrolysis. Even with those similarities one cannot clearly know at which states, apo, ATP-, and ADP-bound state, the enzyme binds to DNA, releases and moves along it. Thus, in the model the three different states are shown without commitment to a specific nucleotide binding state (the nucleotide is therefore not shown, Figure 2-12 A, states I-III). The data presented here suggest that the mutations in the loop disrupt the proper interactions between the N- and C-terminal parts and that these interactions are needed for helicase activity. The mutation(s) in the loop may have one of two effects (each mutation may have a different one). They may either stabilize the interactions between two parts of the molecule and thus prevent the movement needed for activity (Figure 2-12 B, no interactions) or the mutations may have the opposite effect by preventing interactions between the two regions of the molecule, again resulting in impaired activity (Figure 2-12 B, stable interactions).

Interactions between the N-terminal part and the C-terminal AAA+ catalytic domains and the regulation of the catalytic domain by the N-terminal part have also been suggested by a study with the MCM homologue from the archaeon *Sulfolobus solfataricus* (89). This study used purified N-terminal and C-terminal portions of MCM to determine their effect on the biochemical properties of the enzyme. The study demonstrated an interaction between the two parts of the molecule and suggested that the N-terminal portion functions to tether the catalytic domain to the DNA. Similar experiments, however, cannot be performed with the *M. thermautotrophicus* enzyme, as we and others have failed to obtain soluble and properly folded intact catalytic domain. However, as both *S. solfataricus* and *M. thermautotrophicus* MCM proteins are similar in

primary amino acid sequence and in their biochemical properties one would expect that in *M. thermautotrophicus* similar direct interactions between the C- and N-terminal regions do exist. The data presented in this paper suggest that the loop between $\beta 7$ and $\beta 8$ of the N-terminal part may participate in interactions and regulation of the C-terminal catalytic region. As similar residues are also conserved in the *S. solfataricus* MCM one may expect that the same loop is involved in interactions between the N- and C-terminal regions of the *S. solfataricus* protein.

The N-terminal part of MCM was shown to interact with the two Cdc6 proteins of *M. thermautotrophicus* (Cdc6-1 and -2) (35,104). It was also shown that the interactions between MCM and Cdc6 proteins regulate MCM activity (35). Thus, it is possible that the loop between $\beta 7$ and $\beta 8$ of the N-terminal part of MCM also transmits the signal of Cdc6 binding, and perhaps other MCM interacting proteins (e.g. GINS (55)) to the catalytic part of the molecule.

To date, no high-resolution structure of the eukaryotic MCM is available. However, the similarity in primary amino acid sequence of the N-terminal region between the archaeal and eukaryal MCM proteins (in particular in domain C (81)) suggests that the eukaryotic MCM proteins will fold like the archaeal enzyme (59,115). As shown in Figure 2-1, the eukaryotic Mcm2-7 proteins contain similar conserved residues in the same location as the loop in the archaeal enzyme. As the structure of the eukaryotic enzyme is predicted to be the same as the archaeal, one can speculate that the loop will play a similar role in the eukaryotic MCM proteins. The loop may participate in interacting with the catalytic domain and, like the archaeal enzyme, couple the N-terminal region and the catalytic domain

Table 2-1 Oligonucleotides used to generate the mutant clones.

Primer Name	Sequence (5'-3')
MCM F	GGC TCC CAT ATG ATG AAA ACC GTG GAT AAG AGC
MCM R	CCG CTC GAG TCA TCA GTG GTG GTG GTG GTG GAC TAT CTT AAG GTA TCC CCT TGC
Q181AF	ACA CTG AAA CTC GCG GAG CCC CTG GAG
Q181AR	CTC CAG GGG CTC CGC GAG TTT CAG TGT
E182RF	CTG AAA CTC CAG CGG CCC CTG GAG AAC
E182RR	GTT CTC CAG GGG CCG CTG GAG TTT CAG
L184AF	CTC CAG GAG CCC GCG GAG AAC CTT TCC
L184AR	GGA AAG GTT CTC CGC GGG CTC CTG GAG
E185RF	CAG GAG CCC CTG CGG AAC CTT TCC GGT
E185RR	ACC GGA AAG GTT CCG CAG GGG CTC CTG
G190AF	AAC CTT TCC GGT GCG GAA CAG CCC CGG
G190AR	CCG GGG CTG TTC CGC ACC GGA AAG GTT
P193GF	GGT GGG GAA CAG GGC CGG CAG ATA ACA
P193GR	TGT TAT CTG CCG GCC CTG TTC CCC ACC

Table 2-2. Differential Scanning Calorimetry (DSC) analysis of wild type and mutant MCM proteins.

<u>Protein</u>	<u>T_m (°C)</u>	<u>ΔH_{VH} (kJ/mole)</u>	<u>ΔH_{cal} (kJ/mole)</u>
Wild type	66.2 ± 1.7 ¹	1246.8 ± 159.0	530.6 ± 3.1
Q181A	62.5 ± 2.4 ¹	834 ± 56.3	296.7 ± 96.3
E182R	67.7 ± 3.9 ¹	1022.5 ± 334.4	384.1 ± 197.4
L184A	66.0 ± 0.2	1365.5 ± 342.9	189.4 ± 98.3
E185R	68.7 ± 0.2	693.5 ± 177.6	335.8 ± 208.4
G190A	65.5 ± 1.9 ¹	1377.5 ± 126.6	122.4 ± 19.9
P193G	73.0 ± 1.2 ¹	951.5 ± 337.3	331.6 ± 62.0

¹Average of medium (60 K hr⁻¹) and slow (15 K hr⁻¹) scan.

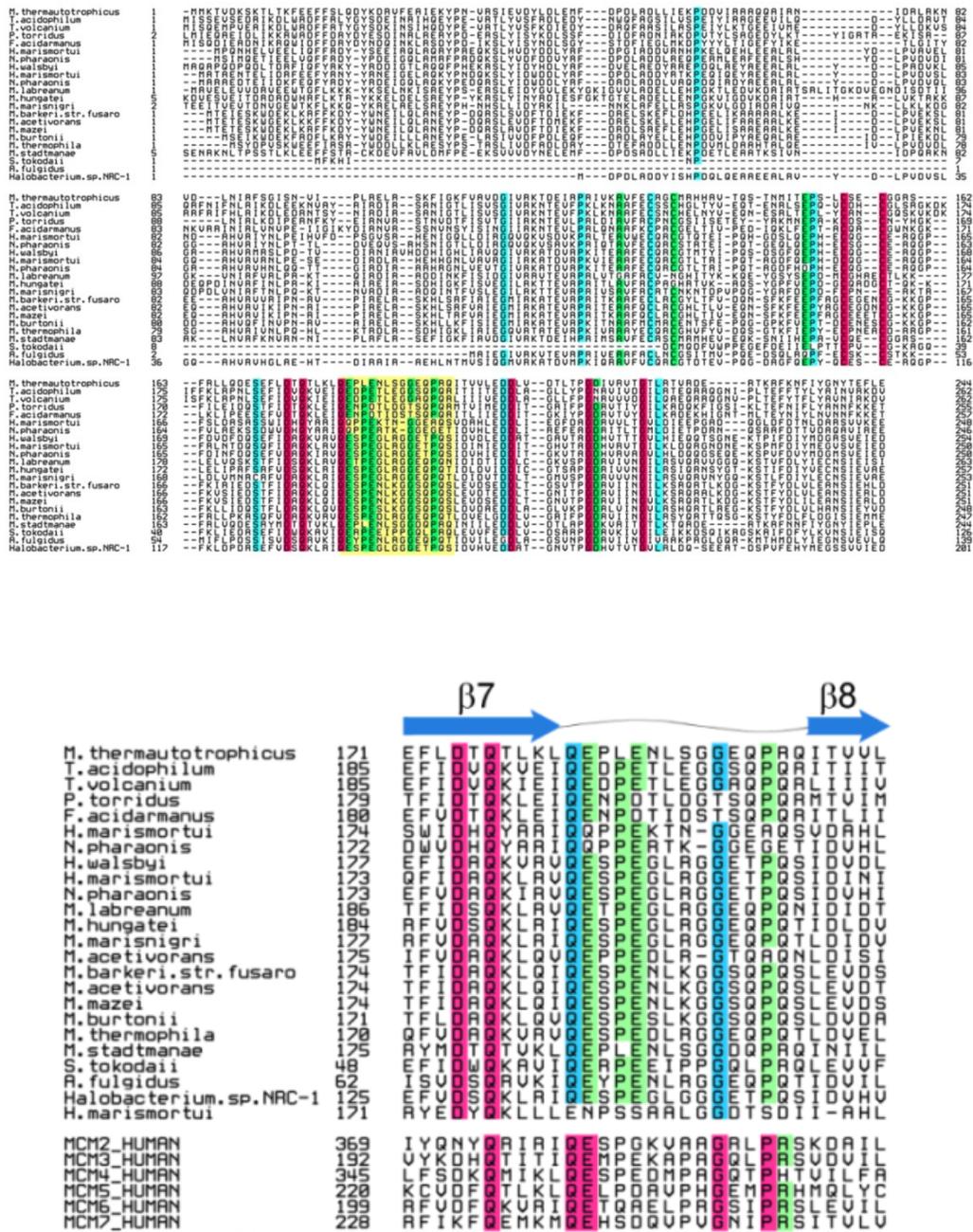


Figure 2-1. Alignments of the N-terminal portion of MCM and amino acid sequences of the loop between $\beta 7$ and $\beta 8$.

Highlighted colors represent residues with >95% identity in pink, >90% identity in blue, and >85% identity in green. The loop between $\beta 7$ and $\beta 8$ is highlighted in yellow (Top).

An enlargement of the alignment of the region between $\beta 7$ and $\beta 8$ is shown in lower panel. An alignment of human Mcm2-7 proteins is shown below archaeal proteins. $\beta 7$ and $\beta 8$ are marked at the top of the alignment (bottom of lower panel).

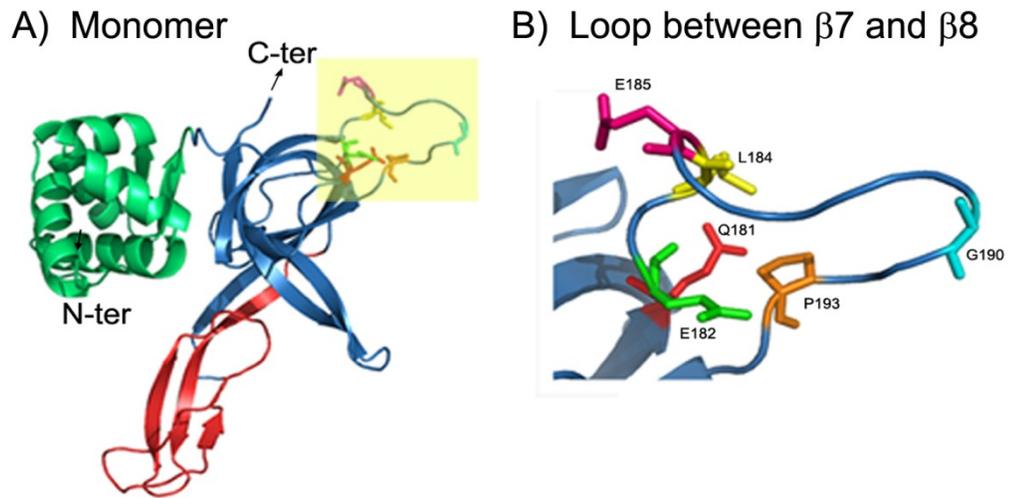
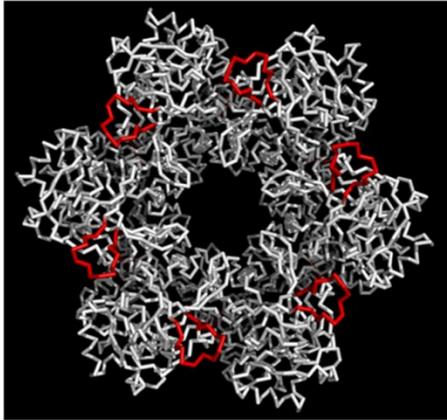


Figure 2-2 Structure of the N-terminal part of MCM.

Domain A shown in green, domain B in red and domain C in blue. The structure was constructed using PyMOL and PDB ID: 1LTL. The loop between $\beta 7$ and $\beta 8$ is highlighted in yellow and the conserved residues are also shown. The loop and the conserved residues are enlarged in panel B.

A) Top view



B) Side view

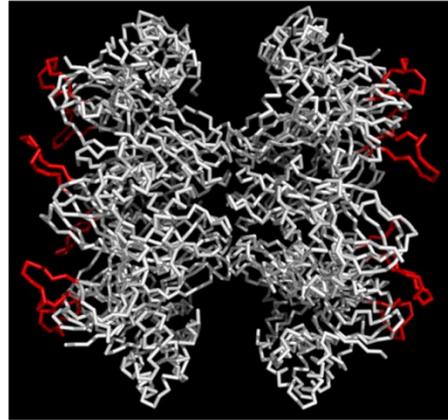


Figure 2-3 Loop is located in the opposite side from dimer interface.

Three dimensional structure of the dodecameric N-terminal portion of the *M. thermautotrophicus* MCM protein (PDB ID: 1LTL (59)) is shown with the loop between $\beta 7$ and $\beta 8$ highlighted in red. A, top view; B, side view.

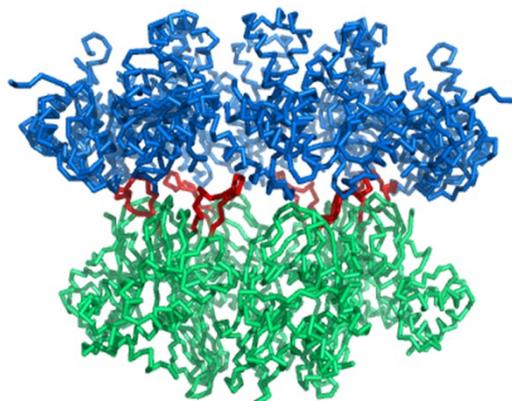


Figure 2-4. EM structure of full length MCM.

Electron micrograph (EM) reconstruction of the intact *M. thermautotrophicus* MCM was generated using PyMol and the coordinates that previously been reported (11). The N-terminal part of the molecule is shown in green, the C-terminal part is shown in blue and the loop between $\beta 7$ and $\beta 8$ is highlighted in red.

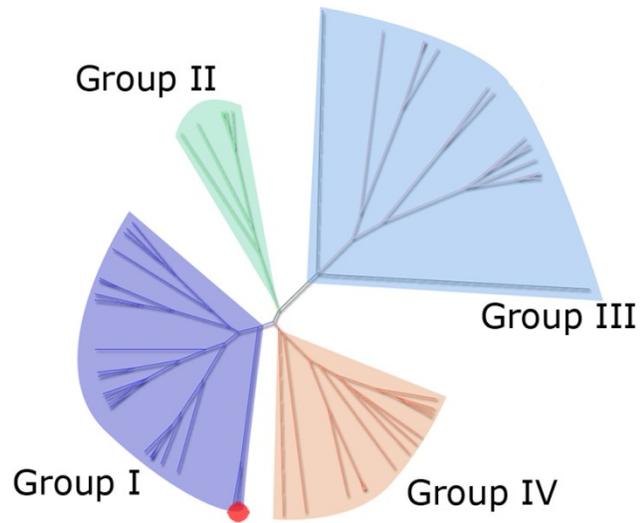


Figure 2-5. Phylogenetic tree of N-terminal archaeal MCM.

Phylogenetic tree based on alignment of the N-terminal part of archaeal MCM proteins constructed by the MUSCLE program. The four groups in the tree are shown in different colors. The position of the archaeon *M. thermautotrophicus* MCM in group I is marked by a red circle.

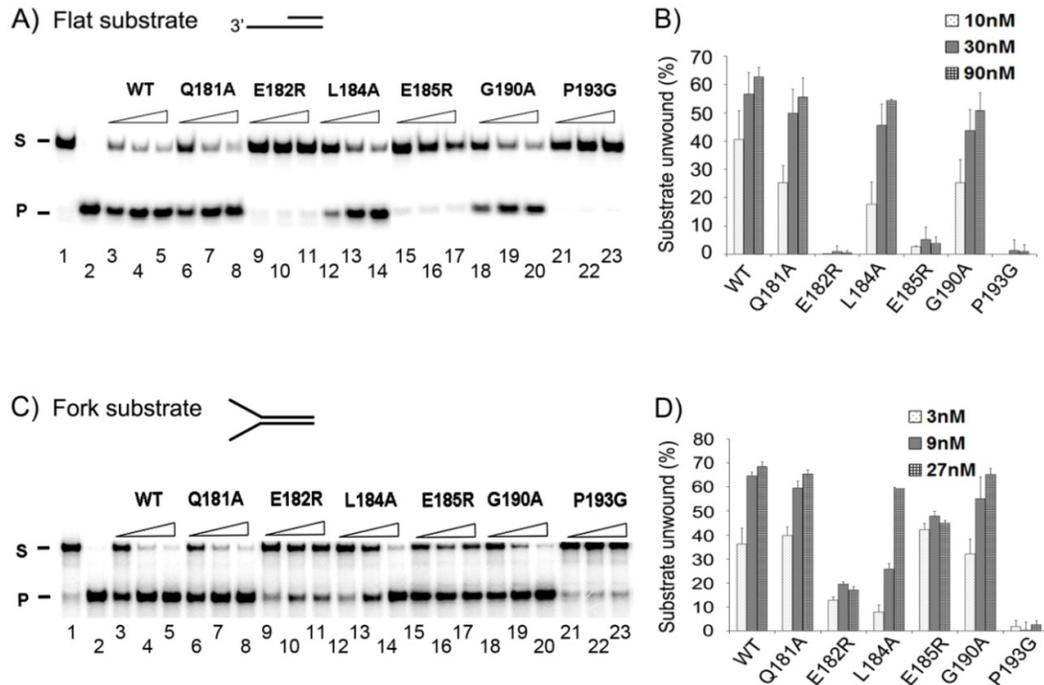


Figure 2-6 Conserved residues in the loop between $\beta 7$ and $\beta 8$ affect MCM helicase activity.

Helicase assays of wild type and mutant MCM proteins were performed as described in “Materials and Methods” with 10 fmol (0.66 nM) of flat (A and C) or forked (B and D) substrates. In panel A the protein concentrations used were 10, 30, and 90 nM (as monomer) while in panel C they were 3, 9 and 27 nM (as monomer). The average result of three experiments are shown in panes C and D. Panels A and B are representative gels. Panel A lane 1, substrate only; lane 2, boiled substrate; lanes 3, 6, 9, 12, 15, 18, and 21 are 10 nM; lanes 4, 7, 10, 13, 16, 19, and 22 are 30 nM; lanes 5, 8, 11, 14, 17, 20, and 23 are 90 nM of MCM proteins. Panel B lane 1, substrate only; lane 2, boiled substrate; lanes 3, 6, 9, 12, 15, 18, and 21 are 3 nM; lanes 4, 7, 10, 13, 16, 19, and 22 are 9 nM; lanes 5, 8, 11, 14, 17, 20, and 23 are 27 nM of MCM proteins.

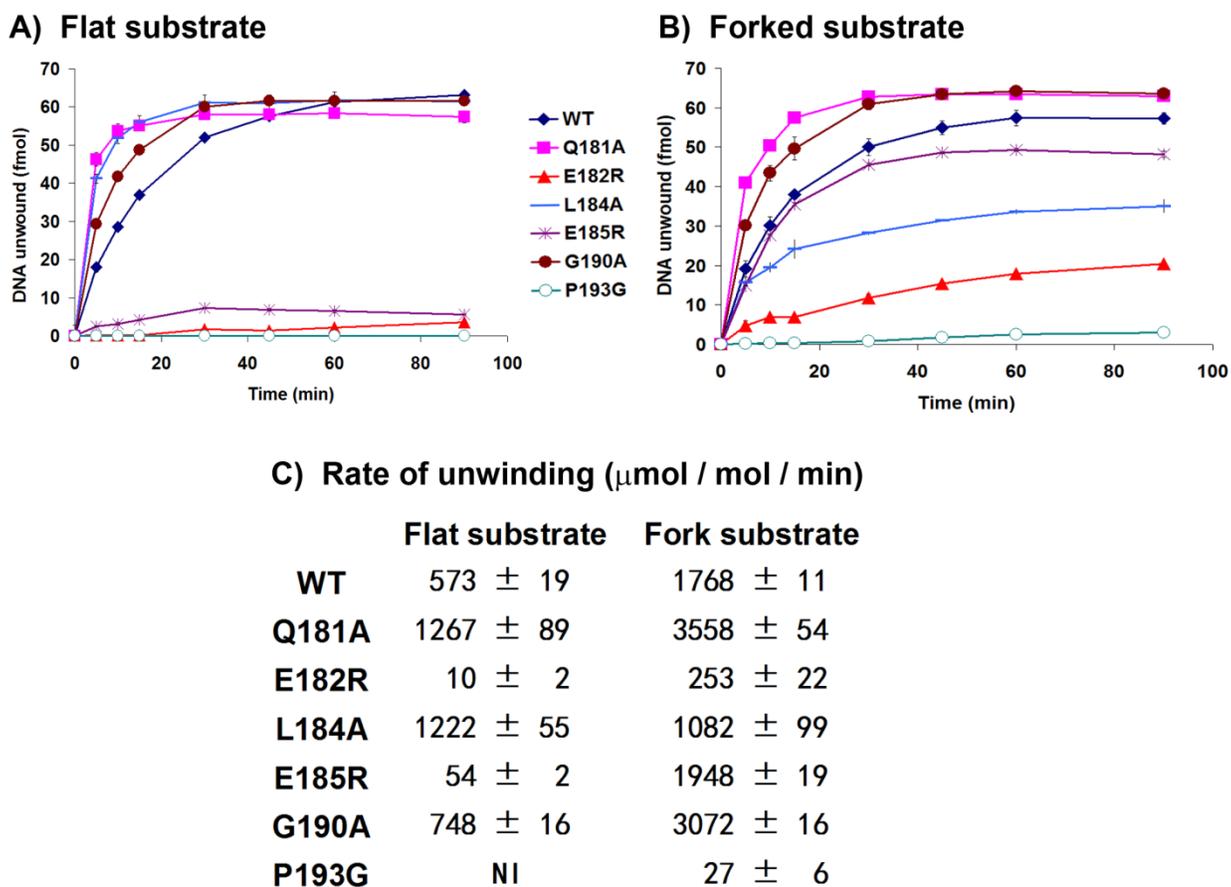


Figure 2-7. Mutations in the loop between $\beta 7$ and $\beta 8$ alter the rate of DNA unwinding.

Helicase assays were performed as described in “Materials and Methods” with flat (A) or forked (B) DNA substrates and 10 nM MCM (A) or 30 nM MCM (B). Following incubation at 60°C aliquots were analyzed at 0, 5, 10, 15, 30, 45, and 90 min. The rate of unwinding was calculated as described in “Materials and Methods” and shown in panel C. NI, not informative due to no observable activity.

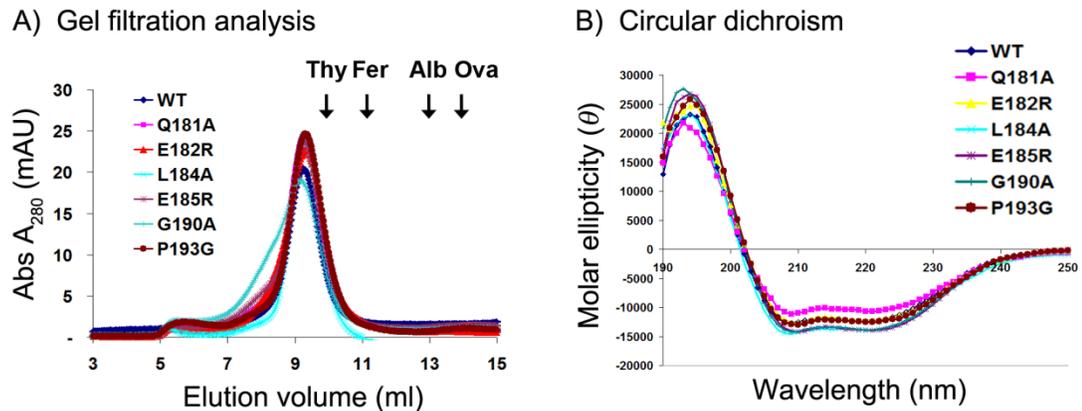
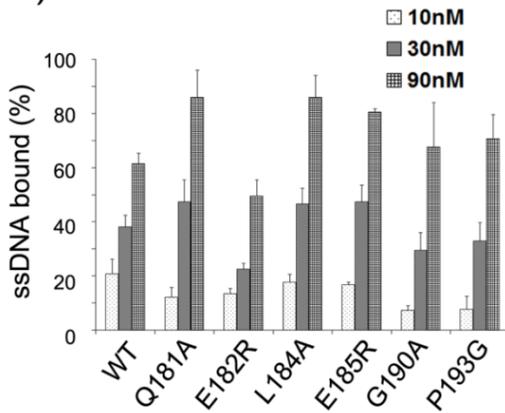


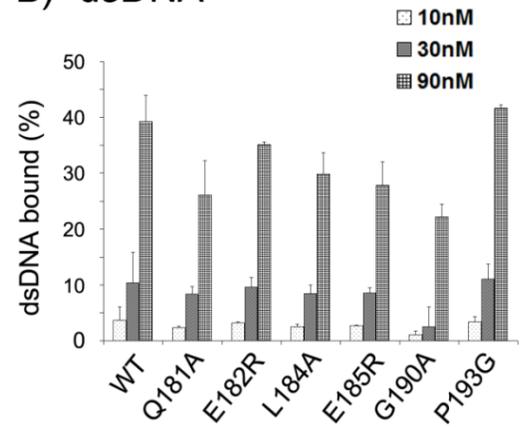
Figure 2-8. Mutations in the loop do not change the overall structure of MCM.

A. Gel-filtration analysis of the wild-type and mutant MCM proteins. One hundred micrograms of protein were loaded onto a superose-6 gel filtration column and analyzed as described in “Materials and Methods”. The peak elutions of several molecular weight standards are shown at the top of the figure. Thyroglobulin (Thy, 670 kDa), ferritin (Fer, 440 kDa), albumin (Alb, 67 kDa) and ovalbumin (Ova, 45 kDa). B. Circular dichroism (CD) spectra of wild-type and mutant MCM proteins. Each spectrum was normalized to mean residue ellipticity $[\theta] = \text{degrees} \times \text{centimeter}^2 \times \text{decimole}^{-1}$ according to the concentrations of protein determined as described in “Materials and Methods”. The spectra were taken at 23°C.

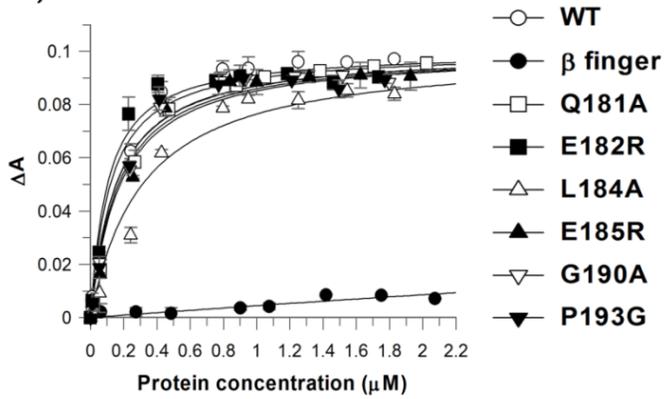
A) ssDNA



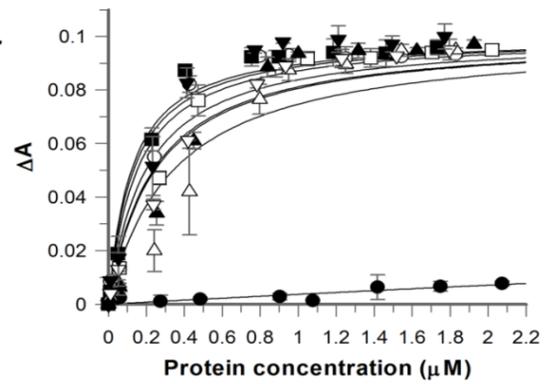
B) dsDNA



C) Without ATP



D) With ATP



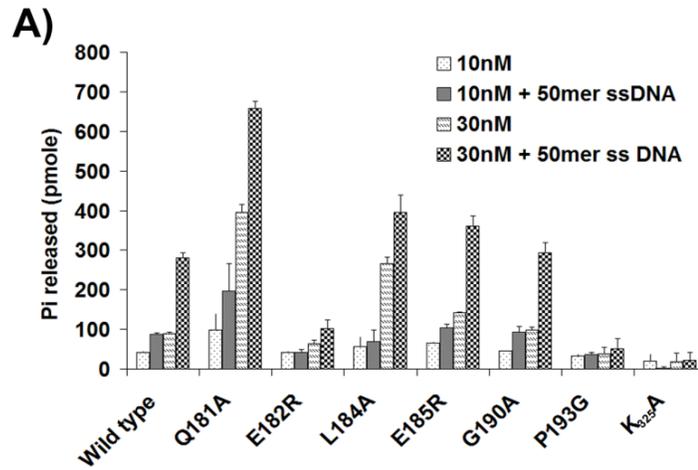
E) Dissociation constants (μM)

	Without ATP	With ATP
WT	0.132 \pm 0.002	0.189 \pm 0.028
Q181A	0.200 \pm 0.009	0.245 \pm 0.040
E182R	0.117 \pm 0.005	0.150 \pm 0.025
L184A	0.376 \pm 0.130	0.517 \pm 0.135
E185R	0.206 \pm 0.000	0.356 \pm 0.027
G190A	0.163 \pm 0.028	0.276 \pm 0.030
P193G	0.174 \pm 0.059	0.162 \pm 0.029

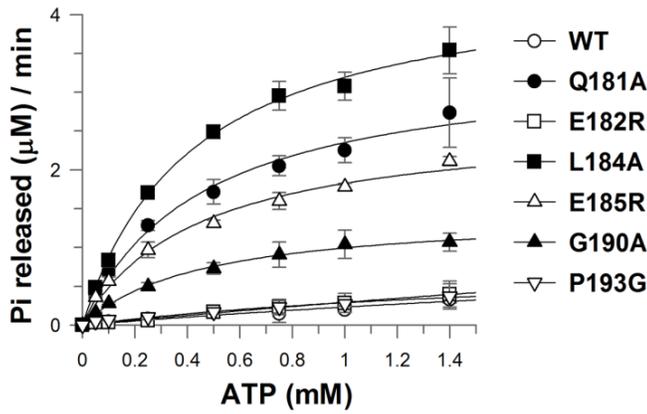
Figure 2-9 Mutations in the loop between $\beta 7$ and $\beta 8$ do not affect MCM DNA binding.

A and B. Filter binding assays were performed as described under “Materials and Methods” using 50 fmol of ^{32}P -labeled single strand (A) or double strand (B) DNA in the presence of 10, 30, and 90 nM of proteins (as monomers). The average result of three experiments is shown.

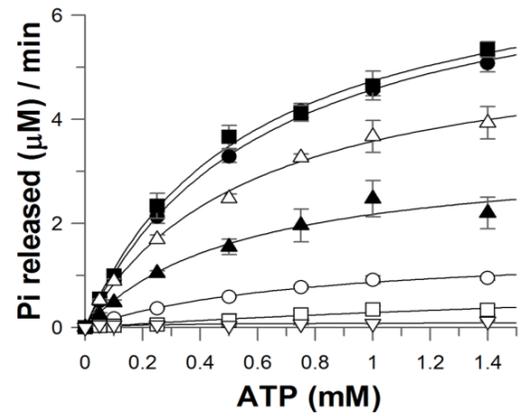
C-E. The K_d values of the interactions between the MCM mutant proteins and DNA were measured using Fluorescence polarization anisotropy (FPA) analysis as described under “Materials and Methods” in the absence (C) and presence (D) of 1 mM ATP and 10 nM Cy3-labeled ssDNA. The anisotropy change was measured as the proteins were titrated into the reaction mixture. Data were analyzed as described in “Materials and Methods” and the K_d values are shown in panel C.



B) Without DNA



C) With DNA



D)

	Without DNA			With DNA		
	Km (mM)	Kcat (sec ⁻¹)	Kcat/Km (sec ⁻¹ ·mM ⁻¹)	Km (mM)	Kcat (sec ⁻¹)	Kcat/Km (sec ⁻¹ ·mM ⁻¹)
WT	NI	NI	NI	0.85 ± 0.13	0.54 ± 0.05	0.64
Q181A	0.47 ± 0.14	1.20 ± 0.22	2.55	0.62 ± 0.07	2.50 ± 0.14	4.03
E182R	NI	NI	NI	NI	NI	NI
L184A	0.43 ± 0.06	1.50 ± 0.00	3.49	0.56 ± 0.02	2.50 ± 0.06	4.46
E185R	0.45 ± 0.06	0.88 ± 0.03	1.96	0.58 ± 0.12	1.90 ± 0.25	3.28
G190A	0.45 ± 0.04	0.48 ± 0.08	1.07	0.54 ± 0.10	1.10 ± 0.22	2.04
P193G	NI	NI	NI	NI	NI	NI

Figure 2-10. Mutations in the loop between $\beta 7$ and $\beta 8$ alter the ATPase activity of MCM.

A. ATPase activity of wild-type and mutant MCM proteins in the absence and presence of ssDNA was determined as described in “Materials and Methods” using 10 and 30 nM of MCM (as monomer) in the presence or absence of 50ng of ssDNA. B-D. K_M and k_{cat} ATP hydrolysis by the wild-type and mutant MCM proteins were determined using 50 nM of proteins in the absence (B) or presence (C) of 50 ng ssM13 and various concentrations of ATP (0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 1 and 1.4 mM) as described in “Materials and Methods”. Rate of reaction and k_{cat} were calculated as described in “Materials and Methods” (D). NI, not informative due to high standard error.

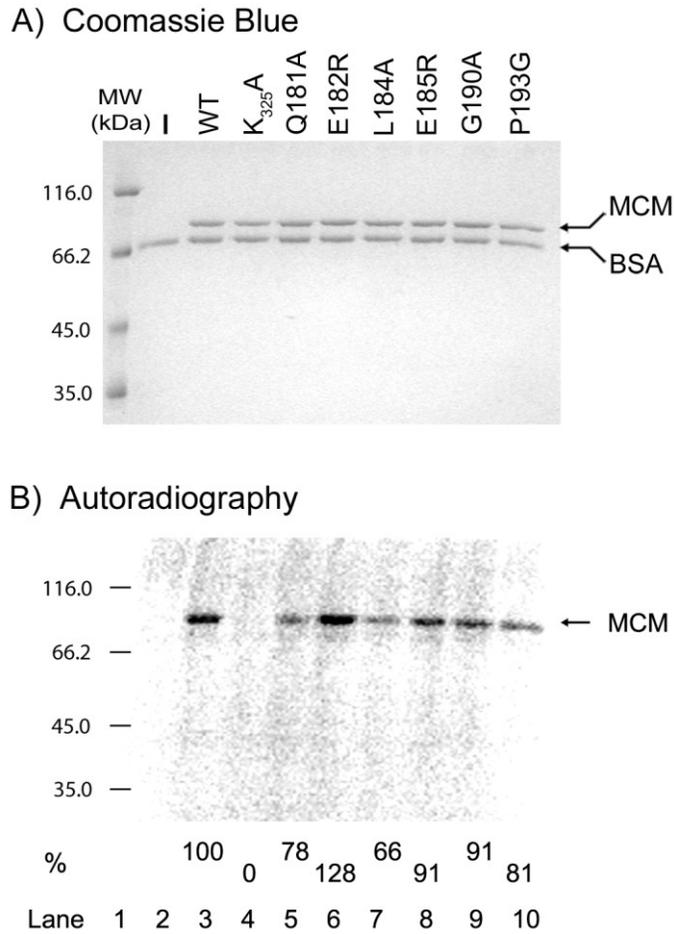


Figure 2-11. MCM proteins with mutations in the loop can bind ATP.

UV crosslinking of ATP to the various MCM proteins was performed as described in “Materials and Methods”. The proteins were separated on 10% SDS-PAGE and visualized by staining with Coomassie Blue (A) and phosphorimaging (B). Lane 1, molecular weight marker; lane 2, no MCM protein added; lane 3, wild-type MCM lane 4, K325A; lane 5, Q181A; lane 6, E182R; lane 7, L184A; lane 8, E185R; lane 9, G190A; lane 10, P193G. The ATP cross-linked to the wild-type and mutant MCMs is indicated (%).

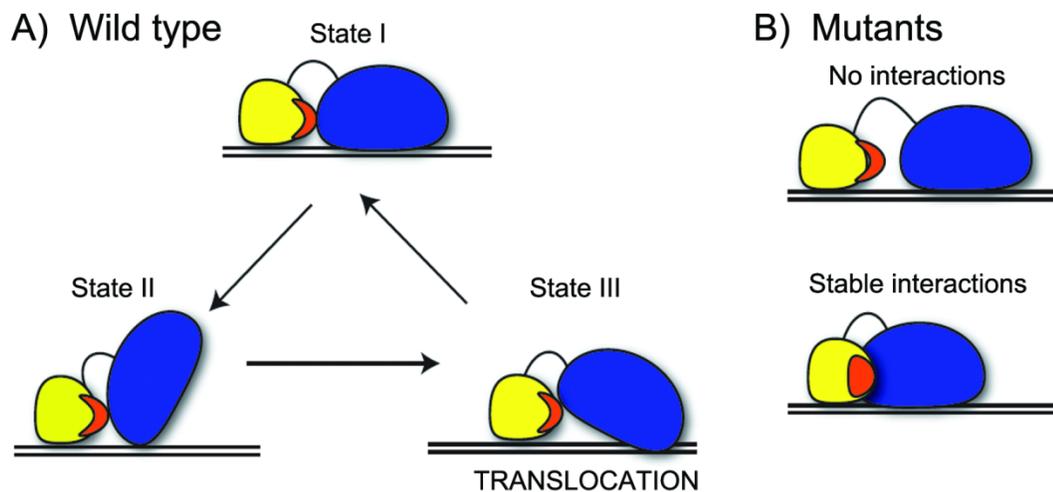


Figure 2-12. A model for the roles played by the N-terminal loop in MCM function.

In the model the C-terminal catalytic region is shown in blue, the N-terminal region in yellow and the loop between $\beta 7$ and $\beta 8$ in red. A. A proposed role for the loop region in the activity of the helicase. The loop connects the N- and C-parts of the molecule and serves as a pivot to allow regulated movements of the domains needed for activity. See text for details. B. Mutations in the loop may weaken the interaction between the N- and C-parts of the helicase (top) or stabilize the interaction (bottom), thus preventing the coupling (coordinate movement) between the two parts of the molecule. See text for details.

Chapter 3 Mutations at Asp 202 and 203 affect structural integrity of the *Methanothermobacter thermautotrophicus* MCM helicase

3.1 Abstract

Minichromosome maintenance (MCM) helicases are the replicative helicases in archaea. MCM unwinds the chromosomal DNA to provide a template for the polymerase during chromosomal replication. It is divided into three parts, the N-terminal, catalytic, and C-terminal regions. The N-terminal portion of the protein was shown to be involved in protein multimerization and binding to single and double stranded DNA. MCM homologues from many archaeal species contain highly conserved residues in the N-terminal part of the molecule. Two of those residues, Asp 202 and 203, are located in the loop between $\beta 8$ and $\beta 9$. When these residues are mapped onto the structure, they are located at the interface between domain A and domain C. Mutations of these residues to Asn result in reduced DNA binding and lower thermal stability, suggesting that these residues may play a role in stabilizing the protein structure.

3.2 Introduction

Alignment of the N-terminal portion of MthMCM helicase against 41 archaeal genomes results in ~60 MCM homologues. These MCM homologues can be further divided into four clades based on the phylogenetic tree ((103); and see Chapter 2). The clade in which MthMCM belongs consists of 21 archaeal MCM homologues. Careful analysis of this alignment revealed that there are two highly conserved residues, D202 and D203, in domain C. These residues are located in the loop between $\beta 8$ and $\beta 9$

(Figure 3-1 A). The conserved residues are located in the beginning of this loop, at the interface between domains A and C (Figure 3-1 B).

Previously, the role of a conserved residue, Pro 62, found in the interface between domain A and domain C have been characterized in *M. thermautotrophicus*. The Pro 62 is equivalent to the Pro 82 found in Mcm5 of *Saccharomyces cerevisiae*. Mutation of Pro 82 to Leu (also known as *Mcm5 -Bob1*) results in bypassing of the Dbf4 dependent Cdc7 kinase (DDK) that is essential in activation of initiation of DNA replication (59 ,109). According to the crystal structure of N-terminal MthMCM, the Pro 62 is located in the interface of domains A and C. Biochemical analyses of P62L mutant in MthMCM show that this mutant has normal DNA binding function as well as ATPase activity stimulated in the presence of DNA (110). However, only the limited helicase activity of P62L was observed, suggesting that this residue may regulate the helicase activity by altering the interactions between the domains A and C.

To understand the role of these conserved residues, D202 and 203 were mutated to Asn. The biochemical analysis of these mutants show that the mutation causes decreased DNA binding, but exhibits wild type level of helicase activity. The loss or reduction of DNA binding by these mutant proteins is rescued in the presence of ATP. Likewise, the ATPase activity of the mutant proteins was also stimulated in the presence of DNA suggesting proteins do interact with the DNA. Mutations of D202 and D203 to Asn also caused changes in the oligomeric state, as well as the thermostability of the protein, suggesting that these may play a role in structural stability.

3.3 Materials and Methods

3.3.1 Materials

ATP and [γ - 32 P]ATP were obtained from GE Healthcare, and oligonucleotides were synthesized by the CARB DNA synthesis facility.

3.3.2 Expression and purification of MCM mutant proteins

All *M. thermautotrophicus* MCM mutant proteins used in this study are derivatives of the full-length enzyme and were generated using PCR-based site-directed mutagenesis as previously described (81,103).

3.3.3 ATPase assay

ATPase assays were carried out as previously described ((103); also see chapter 2). ATPase assays were repeated three times, and their averages with standard deviations are shown in Figure 3-5 B.

3.3.4 Gel filtration

One hundred micrograms of wild-type and mutant MCM proteins were applied to a superose-6 gel filtration column (HR10/30, GE Healthcare) pre-equilibrated with buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 10% glycerol. The column was run at a flow rate of 0.2 ml/min at room temperature and 250 μ l fractions were collected. The presence of protein was determined by ultraviolet absorbance at 280 nm.

3.3.5 Nitrocellulose filter DNA binding assay

DNA binding assay was carried out as previously described ((103); also see chapter 2). DNA binding experiments were repeated three times, and their averages with standard deviations are shown in Figure 3-2.

3.3.6 DNA helicase assay

Helicase assay was carried out with the same substrate described in ((103); also see chapter 2). The helicase experiments were repeated three times, and their averages with standard deviations are shown in Figure 3-4 A.

3.3.7 Differential scanning calorimetry (DSC) measurements:

DSC experiments were carried out as described previously ((103); also see chapter 2). DSC scans on the samples were repeated twice.

3.3.8 Circular dichroism (CD)

Circular dichroism (CD) measurements were performed as previously described ((103); also see chapter 2).

3.4 Results

3.4.1 Mutations at D202 and D203 residues reduce DNA binding of MCM

Since D202 and 203 are highly conserved among different archaea, it is possible that they are involved in the function of helicase activity. One of the essential properties of helicases is their ability to bind to DNA. Thus nitrocellulose filter binding assay was performed to determine the effect of mutations on both ss and dsDNA binding. As

shown in Figure 3-2 A, the D203N and DD202,203NN mutant proteins show substantial decrease in ssDNA binding. The D202N mutant, however, has a limited effect on ssDNA binding compared to the wild-type protein. When dsDNA was used as a substrate, D202N and D203N mutants showed a similar level of binding but reduced binding compared to the wild-type protein (Figure 3-2 B). The double mutant shows almost no binding to dsDNA (Figure 3-2 B). Previously, the wild-type protein has been shown to bind ssDNA more tightly than dsDNA and it is also indicated here in the Figure 3-2 (102). This difference can be seen in the D202N and D202,203NN mutant proteins which show reduced binding to dsDNA compared to the ssDNA. However, in the case of D203N mutant, it has slightly greater affinity for dsDNA than to ssDNA. This suggests that D203 may be involved in discriminating between ss or dsDNA, though the mechanism of how it does so is unclear.

3.4.2 Mutations affect structural integrity of MCM protein

Since the loop between $\beta 8$ and $\beta 9$ is located at the surface of monomeric MCM it is possible that the oligomeric structure of MCM helicase is affected due to the mutation. Thus superose 6 size exclusion chromatography was used to determine the oligomeric state of MCM helicase wild-type and mutant proteins. As shown in Figure 3-3 A, D203N and DD202,203NN have broader elution profiles and appear to form larger oligomers compared to the wild type protein. The D202N mutant, however, appears to form a smaller oligomer compared to the wild-type protein. Thus, it is possible that the D202N mutant exists primarily as a single hexamer, unlike the wild-type protein that has a stable double hexamer structure in solution. The D203N and DD202,203NN mutants may have the tendencies to form filaments, as the filaments of wild-type MCM have been observed

in EM studies (85). The notable differences in the oligomeric structure may explain the differences in DNA binding between D202N mutant and D203N, and DD202,203NN mutants. It is possible that the higher oligomeric forms, such as filaments may not allow binding to the DNA. To ensure that the observed differences in oligomeric state are not caused by misfolding of the proteins, circular dichroism (CD) was used. As shown in Figure 3-3 B, there appears to be no gross changes in the secondary structure of mutant proteins.

Since the mutations cause variations in oligomeric states of the MCM, it may also have an effect on the stability of MCM. Thus, thermostability of MCM protein was investigated using differential scanning calorimetry (DSC). The DSC results show that the wild type protein has an average melting temperature (T_m) of 66.2°C (Table 3-2). The D202N, D203N and DD202,203NN mutants showed average T_m of 63.7, 62.5, and 57.2°C respectively. These values suggest that the mutation caused instability, especially for the DD202,203NN mutant that shows a significant decrease in T_m ($P < 0.005$). The DNA binding assay condition includes incubation at 60°C. It is possible that the DD202,203NN mutant may have denatured, causing lower affinity for DNA.

3.4.3 Mutations have no effect on helicase or ATPase activity

MCM helicase couples the ATP hydrolysis and translocation along ssDNA in order to unwind dsDNA. If the DNA binding property of helicase were to be impaired, helicase and stimulation of ATPase activity in the presence of DNA are expected to be affected. The helicase assay requires incubation of MCM helicase at 60°C thus instability of the protein may have effect on helicase activity. However, helicase assays demonstrate that the mutations do not have any effect on the helicase activity, showing

the same level of activities compared to the wild type (Figure 3-4 A). The ATPase activities of mutant proteins are also stimulated in the presence of ssDNA (Figure 3-4 B). The basal activities (without DNA) of mutants are slightly lower yet the fold stimulations are greater than the wild type. This is more apparent in the double mutant where ATPase activity in the presence of DNA is even greater than the wild type (Figure 3-4 B).

3.4.4 ATP stimulates DNA binding function of the mutants

The possible explanation for mutants' ability to unwind DNA and show DNA dependent ATPase activity is the effect of ATP in the reaction. ATP is present in both the helicase and ATPase assays, but absent in the DNA binding assay. Therefore, nitrocellulose filter binding assay was performed in the presence of ATP. As shown in the Figure 3-5, the ATP promotes the binding of wild type and mutant proteins to DNA. It has been reported that ATP stimulates the binding of MCM to ssDNA when incubated at 55°C or 66°C (61). This result suggests that the reduced DNA binding caused by the mutation is restored by the presence of ATP. It is possible that ATP stabilizes the structure of the mutants which in turn allow them to bind to the substrate and consequently translocate and unwind the DNA.

3.5 Summary and Discussion

According to the crystal structure of N-terminal portion of MCM, D202 and D203 are located on the interface between domains A and C (Figure 3-1). Recent structural study of MthMCM by EM in complex with long dsDNA showed that DNA binds to the outer surface of MCM in the cleft of domain A and B, where D202 and 203 are approximately located (107). Domain A was shown to be responsible for binding to the

long dsDNA, since deletion of domain A resulted in significant decrease in binding to long dsDNA. In this study, mutations caused decrease in DNA binding in the absence of ATP. Since these are acidic residues, it is not clear whether they are directly involved in interacting with DNA, but the assay clearly demonstrated that they indirectly affect the DNA binding. This effect is lost when ATP is present, suggesting that the catalytic activity in the AAA+ region influences the conformational changes in the N-terminal portion, including these residues.

Previously, the role of a conserved residue, Pro 62 (P62) found in the interface between domain A and domain C have been characterized in *M. thermautotrophicus*. The P62 is an equivalent residue of Pro 82 found in Mcm5 of *Saccharomyces cerevisiae*. Mutation of P82 to Leu (also known as *Mcm5 -Bob1*) results in bypassing of the Dbf4 dependent Cdc7 kinase (DDK) that is essential in activation of initiation of DNA replication (59,109). According to the crystal structure of N-terminal MthMCM, the P62 is located in the interface of domains A and C. Biochemical analyses of P62L mutant in MthMCM show that this mutant has normal DNA binding function as well as ATPase activity stimulated in the presence of DNA (110). However, only the limited helicase activity was observed in P62L, suggesting that this residue may be involved in regulation. When P62, D202 and D203 are mapped on the monomeric MCM, they are located in the cleft of domain B and A (Figure 3-6 A). It was suggested that P62 located in domain A is closely interacting with F109 of domain C. The crystal structure of N-terminal MCM with P62L mutation did not show significant difference in overall structure compared to the wild type, but showed slight increase in the distances between domain A and C, causing domain A to be pushed out (59). It was suggested that the mutation of P62 to

any bulky side chain such as Leu impose steric hindrance causing “domain push” allowing DDK independent activation. *In vivo* assay using *S. cerevisiae* proved this hypothesis that while mutating P82 to smaller hydrophobic residues did not allow activation of MCM5 in the absence of DDK, mutations to larger hydrophobic residues resulted in the bypass (59). In this study, it was suggested that this tight packing is centered solely on P62, making it a “sole anchoring point” between domain A and C (59). When the polar contacting residues of D202 and D203 were mapped using PyMOL, the side chains are making interactions with neighboring residues which are located in the domain A (Figure 3-6 B) (126). The side chain of D202 shows a hydrogen bond with amine group from Arg87 (R87). The one of the oxygen from the side chain of D203 makes hydrogen bond with Ser38 (S38) while other oxygen has a hydrogen bond with R87. S38 also makes a hydrogen bond with R87. Thus in addition to P62 as an anchor point between domains A and C, it is possible that D202 and 203 may stabilize inter-domain interaction. Yet the interaction may be ATP independent, since the presence of ATP restores the DNA binding possibly by affecting the conformations in the N-terminal region.

Table 3-1 List of primers used to generate mutants

Mutation	Sequence (5'-3')
D202N F	G CAG ATA ACA GTT GTC CTG GAG AAC GAC CTG GTT GAC ACC CTC ACA CCC
D202N R	GGG TGT GAG GGT GTC AAC CAG GTC GTT CTC CAG GAC AAC TGT TAT CTG C
D203N F	G CAG ATA ACA GTT GTC CTG GAG GAC AAC CTG GTT GAC ACC CTC ACA CCC
D203N R	GGG TGT GAG GGT GTC AAC CAG GTT GTC CTC CAG GAC AAC TGT TAT CTG C
DD202,203NN F	GTT GTC CTG GAG AAC AAC CTG GTT GAC ACC CTC
DD202,203NN R	GAG GGT GTC AAC CAG GTT GTT CTC CAG GAC AAC

Table 3-2 Thermostability of wild-type and mutants from DSC results

	T_m ave (°C)	ΔH_{VH} (kJ/mole)	ΔH_{cal} (kJ/mole) [§]
Wild type	66.2 ± 1.7*	1246.8 ± 159.0	530.6 ± 3.1
D202N	63.7 ± 0.1	1123.5 ± 149.2	276.6 ± 56.6
D203N	62.5 ± 0.2	1360.0 ± 157.0	284.5 ± 3.2
DD202,203NN	57.2 ± 1.7*	1277.5 ± 286.4	198.1 ± 48.3

*The average of slow and medium scan

§Calorimetric enthalpy is determined from measuring the absorbance of proteins at 280 nm and concentration was determined using the calculated extinction coefficients ($\epsilon=28730$)

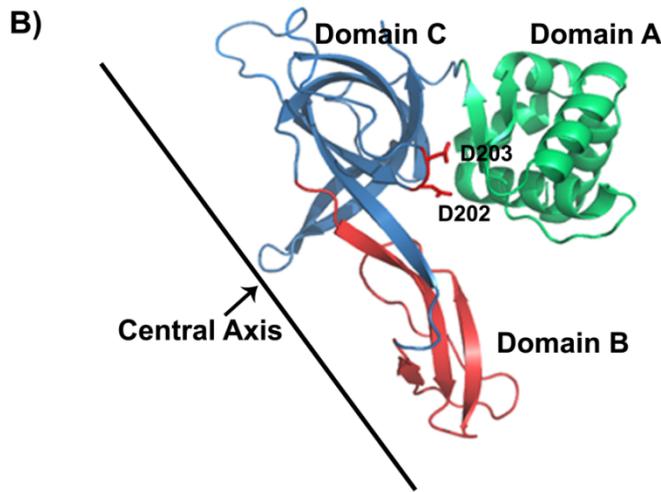
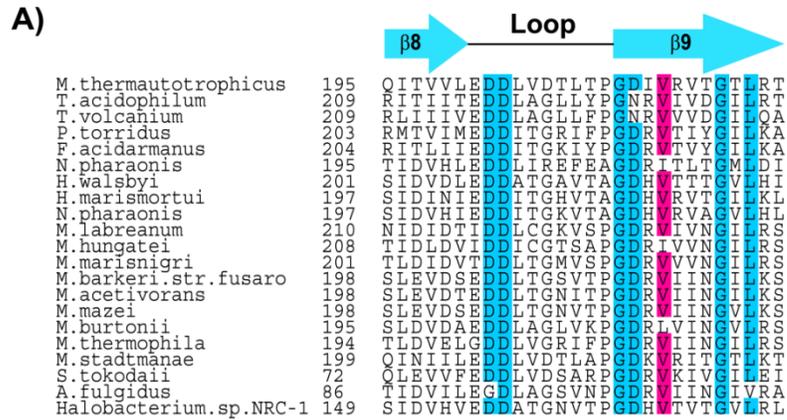


Figure 3-1 Residues, Asp 202 and 203 are highly conserved among archaea.

A) Monomeric structure of N-terminal MCM helicase with domain A colored in green, domain B in red, and domain C in blue. The Asp 202 and 203 are highlighted in red and labeled. The axis of the central cavity is drawn to clarify the relative location. PyMOL was used to construct the figure (PDBID: 1LTL). B) An alignment of the region of the loop between $\beta 8$ and $\beta 9$, where D202 and D203 are located. Highlighted in blue is >90% identity, >80% is in pink

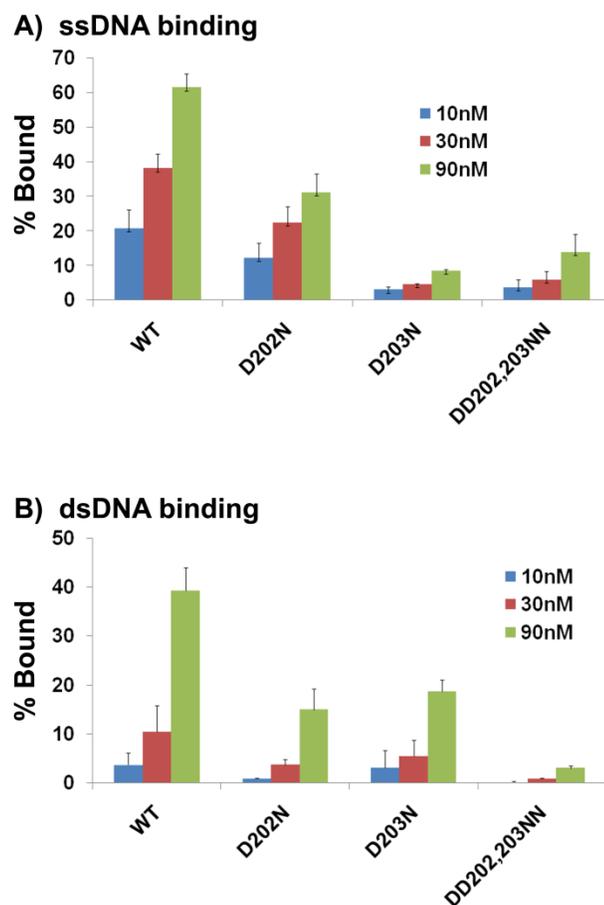


Figure 3-2 Mutations affect ss and dsDNA binding.

Nitrocellulose filter binding assays were performed as described under “Materials and Methods” using 50 fmol of ³²P-labeled single strand (A) or double strand (B) DNA (49-mer) in the presence of 10, 30, and 90 nM of proteins (as monomers). The average result of three experiments is shown.

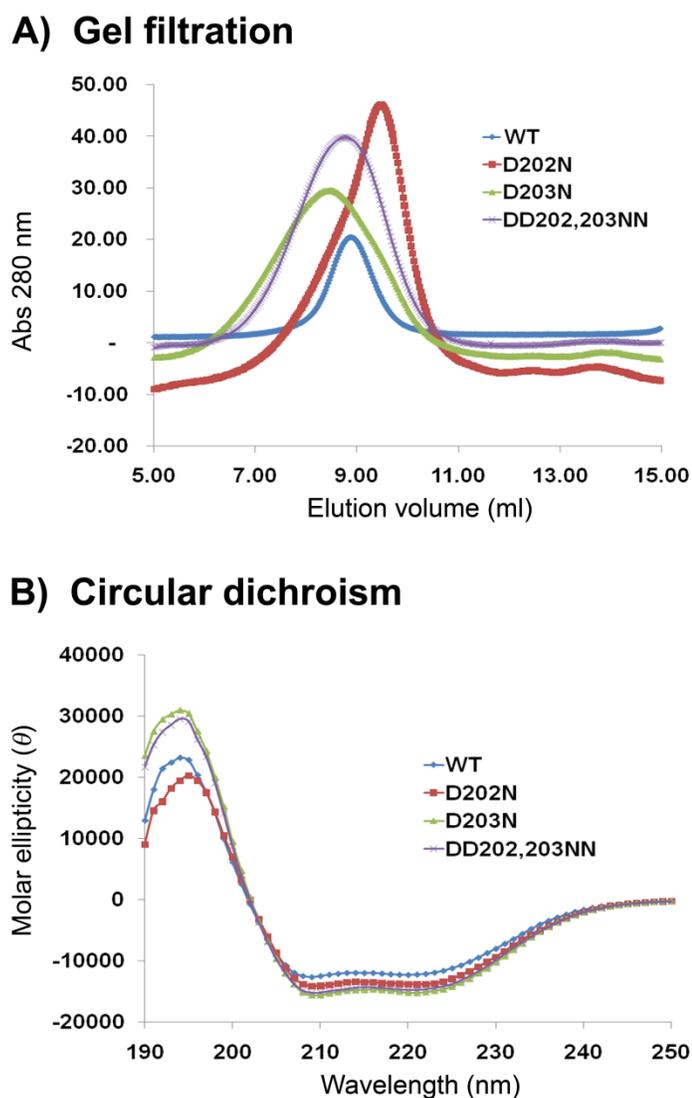


Figure 3-3 The effect of mutations on structural integrity.

The wild-type and mutant proteins were run on superpose 6 gel filtration column as described in “Materials and Methods” (A). The circular dichroism (CD) of wild-type and mutant proteins were measured at room temperature, as described in “Materials and Methods” (B).

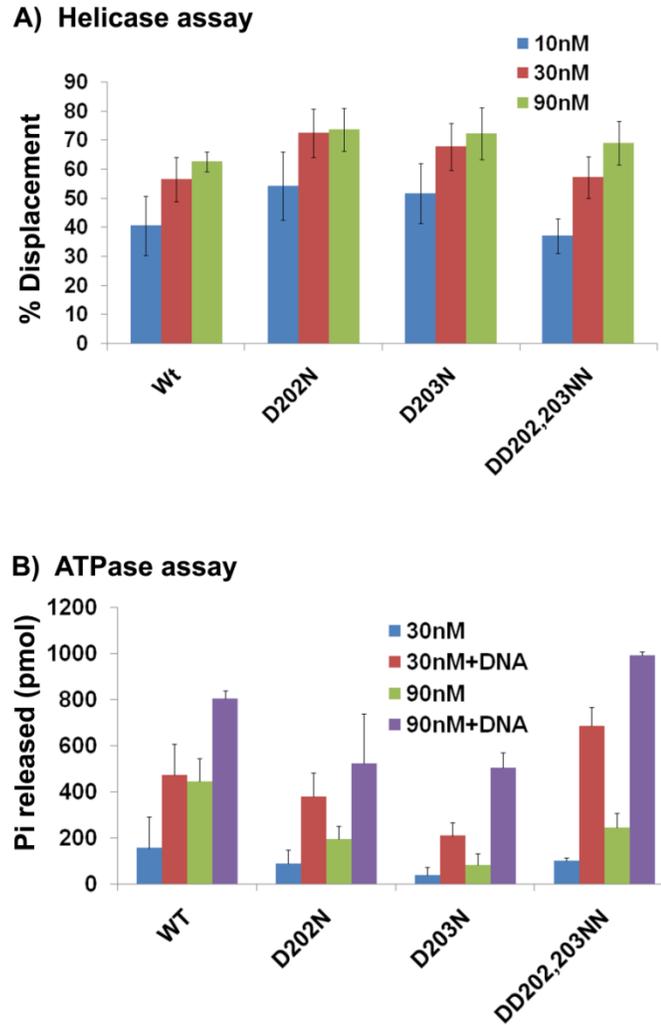


Figure 3-4 Mutant proteins have helicase and ATPase activities.

Helicase activity of wild type and mutants were determined using flat substrate as described in “Materials and Methods” using 10, 30, and 90 nM MCM as monomer (A). ATPase activity of wild-type and mutant MCM proteins in the absence and presence of ssDNA was determined as described in “Materials and Methods” using 30 and 90 nM of MCM (as monomer) in the presence or absence of 50 ng of ssDNA (49mer) (B).

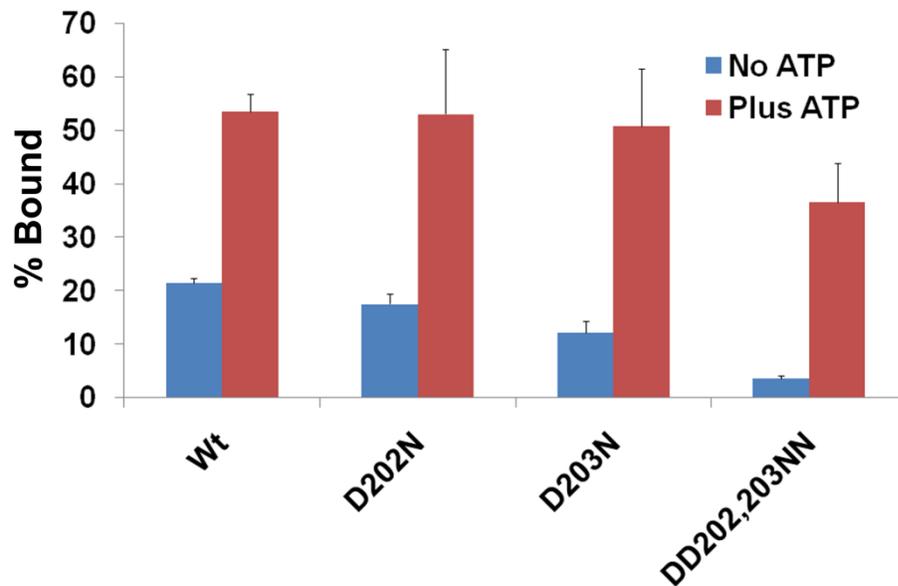


Figure 3-5 DNA binding affinity is increased in the presence of ATP.

Filter binding assays were performed as described under “Materials and Methods” using 50 fmol of ^{32}P -labeled ssDNA (49-mer) in the presence of 25 nM of proteins (as monomers) in the absence or presence of 1 mM ATP. The average of three experiments is shown.

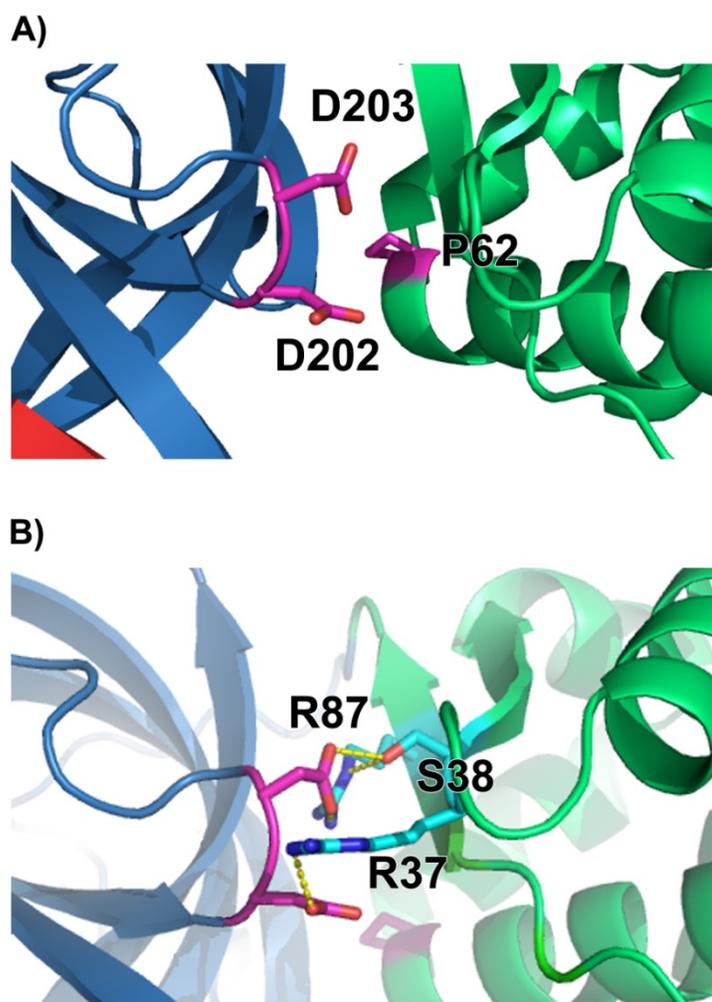


Figure 3-6 Asp 202 and 203 are located in the interface of domain C and domain A.

A. The close up structure of the N-terminal monomeric MCM is shown with domain A colored in red, domain B in red, and domain C in blue. The Asp 202 and 203 are highlighted in magenta. *Bob1* mutant, P62 is shown also in magenta and labeled with arrow. B. Yellow dots represent ionic interactions between D202 to R37, D203 to S38 and to R87, and between S38 and R87.

Chapter 4 Biochemical characterization of conserved residues found in N-terminal region of MCM helicase from *Methanothermobacter thermautotrophicus*

4.1 Abstract

Minichromosome maintenance (MCM) helicases are the archaeal replicative helicases that separate the two strands of chromosomal DNA during replication. The catalytic activity resides within the C-terminal region of the MCM protein while the N-terminal portion plays an important role in DNA binding and multimerization. The N-terminal portion is divided into three domains based on the crystal structure. Domain A is involved in regulation, domain B is involved in both ss and dsDNA binding, and domain C is involved in DNA binding and multimerization. Although the N-terminal portion of MCM is less conserved compared to the C-terminal catalytic region, an alignment of MCM homologues from many archaeal species revealed highly conserved amino acids. In this chapter, conserved residues found in domain C are characterized.

4.1 Introduction

Highly conserved residues found in the N-terminal region of MCM helicase among 41 archaea have been identified and characterized ((103); and see chapter 2 and chapter 3). They include residues located in the loop between $\beta 7$ and $\beta 8$ ((103); and chapter 2) and the loop between $\beta 8$ and $\beta 9$ (chapter 3). There are highly conserved residues scattered around within the domain C, and these residues include Gln 176 (Q176), Pro 210 (P210), Gly 211 (G211), Asp 212 (D212), Val 214 (V214), and Gly 218 (G218). Based on the crystal structure of the N-terminal MthMCM, these residues are located mostly in the surface of the monomeric MCM. In this chapter, the biochemical

properties of these conserved residues found in the domain C which have not been characterized are elucidated. Based on the crystal structure of the N-terminal MthMCM, these residues are located mostly in the surface of the monomeric MCM (Figure 4-1). The conserved residues, Q176, P210, G211, D212, V214, G218 were mutated to Ala (Q176A), Gly (P210G), Ala (G211A), Asn (D212N), Ala (V214A), and Ala (G218A) respectively. These mutants are characterized using various biochemical methods. The results are shown and the implications for their roles are discussed.

4.2 Materials and Methods

4.2.1 Materials

ATP and [γ -³²P]ATP were obtained from GE Healthcare, and oligonucleotides were synthesized by the CARB DNA synthesis facility.

4.2.2 Expression and purification of MCM mutant proteins

All *M. thermautotrophicus* MCM mutant proteins used in this study are derivatives of the full-length enzyme and were generated using PCR-based site-directed mutagenesis as previously described (81,103).

4.2.3 Gel Filtration

Gel filtration was performed as described in (103); and chapter 2 .

4.2.4 ATPase assay

ATPase assays were carried out as previously described ((103); also see chapter 2). ATPase assays were repeated three times, and their averages with standard deviations are shown in Figure 4-2.

4.2.5 Nitrocellulose filter DNA binding assay

DNA binding assay was carried out as previously described ((103); also see chapter 2). DNA binding experiments were repeated three times, and their averages with standard deviations are shown in Figure 4-5.

4.2.6 DNA helicase assay

Helicase assay was carried out with the same substrate described in ((103); also see chapter 2). The helicase experiments were repeated three times, and their averages with standard deviations are shown in Figure 4-3.

4.2.7 Differential scanning calorimetry (DSC) measurements:

DSC experiments were carried out as described previously ((103); also see chapter 2). DSC scans on the samples were repeated twice.

4.3 Results

4.3.1 The mutant proteins show diverse effect on ATPase activity

M. thermotrophicus MCM has been shown to have ATPase activity that is stimulated in the presence of DNA (71,72). To see whether the conserved residues affect ATPase activities of MCM helicase, the level of ATP hydrolysis in the absence or

presence of ssDNA was determined. As shown in Figure 4-2, ssDNA increases the wild-type ATPase activity by almost 7 fold at 30 nM concentration. The results also show that P210G, G211A, and PG210,211GA exhibit high basal activity in the absence of DNA. The basal activity of these mutants is almost saturated at 30 nM concentration thus the effect of DNA appears limited. The mutant proteins, Q176A, V214A, and G218A show relatively the same level of ATPase stimulation in the presence of DNA as the wild-type protein. The D212N mutant, however, shows slightly lower basal activity and stimulation by ssDNA.

4.3.2 Mutations do not affect MCM helicase structure

To understand how these residues may be involved in the regulation of DNA dependent ATPase activity, the residues were mapped in the crystal structure of N-terminal MthMCM. As shown in Figure 4-1, all the residues except for G218 are located on the domain C relatively on the surface of the monomeric protein. The G218 is located on the surface of domain A and C, similar to the conserved residues found in the loop between $\beta 8$ and $\beta 9$ (as described in Chapter 3). Detailed analysis of these residues indicates that they are involved in polar contacts with neighboring residues (summarized in Table 4-2). Except for P210, all these residues are involved hydrogen bonding with neighboring residues. The G218 is making a hydrogen bonding to support the anti-parallel β -sheets between $\beta 9$ and $\beta 3$, and located in close proximity to F109 that has been suggested to play a role in the hydrophobic interaction with P62 to stabilize the domain A and domain C (see Chapter 3).

We also investigated whether the mutations may cause instability in protein structure using gel filtration, CD and DSC measurements. Wild-type MCM protein has

been shown to have thermal transition peak at $66.2 \pm 1.7^\circ\text{C}$ (Chapter 2). To confirm that the mutants are stable in the reaction temperature range ($\sim 60^\circ\text{C}$), differential scanning calorimetry (DSC) was used. As shown in table 4-3, the wild-type protein has an average melting temperature (T_m) of 66.2°C . All of the mutant proteins show relatively the same T_m suggesting that the mutations do not interfere with the thermostability of the protein (Table 4-3). It is possible that the mutations may cause dissociation of MCM helicase oligomeric state. Superose 6 gel filtration analysis of the mutants revealed that D212N and V214A show broader peak starting from 6 ml to 10 ml elution volume. It is possible that elution profile to the wild-type, suggesting that it is a multimer in solution (Figure 4-3 bottom). Q176A also elutes slightly before the wild-type protein, suggesting that the population of higher molecular weight form is more prevalent in the sample. The secondary structural differences were observed by CD spectra. The CD results reveal that the proteins likely share the same secondary structure as the wild-type protein (Figure 4-3 top).

4.3.3 Mutations in protein do not affect helicase activity

MCM helicases couple the ATP hydrolysis and translocation along ssDNA in order to unwind dsDNA. Since mutants, P210G and G211A showed decoupling of DNA dependent ATPase activity, it is possible that the helicase activity may be affected. As shown in Figure 4-4 A, most of the mutants unwind forked substrate with almost the same strength as the wild-type, except for D212N. The helicase activity of D212N seems to saturate and inhibit after 27 nM of protein is present in the reaction when forked substrate was used. When the flat substrate was used, mutant proteins, Q176A, V214A, and G218A showed relatively similar level of activity as the wild-type protein (Figure 4-

4 B). The mutant proteins, P210G, G211A, PG210,211GA and D212N showed a decrease in activity compared to the wild-type protein, especially at lower concentration (10 nM). It is not clear why D212N has an inhibitory effect on forked substrate at concentration above 9 nM since this inhibition is not observed with flat substrate at 90 nM concentration.

4.3.4 Mutations in MCM protein alter the DNA binding

The MCM helicase has been shown to have stable interactions with ss or dsDNA. Since N-terminal portion of MCM helicase is involved in DNA binding, as well as multimerization of the protein, it is possible that the conserved residues may regulate the ATPase activity by altering the DNA binding function. Therefore, we performed the nitrocellulose filter binding assay to determine the effect of mutations on DNA binding. As shown in Figure 4-5 A, P210G, G211A or their double mutants PG210,211GA mutants show increased ssDNA binding compared to the wild-type, while these mutants show slightly decreased dsDNA binding compared to the wild-type (Figure 4-5 B). However, Q176A and D212N bind weaker to the ssDNA compared to the wild-type. Consistent with previous reports, the wild-type protein shows less affinity for dsDNA (102,117). Interestingly, mutants that showed increased ssDNA binding (P210G, G211A, and PG210,211GA) show almost 50% reduction in dsDNA binding at all concentrations compared to the wild-type protein. On the contrary, Q176A and D212N mutants show better binding to dsDNA than to ssDNA. It is possible that the Q176 and D212 are involved in ssDNA binding via charged side chains. That may explain the decreased ssDNA binding by Q176A and D212N mutants, but improved binding of dsDNA.

4.3.5 ATP increases binding of MCM proteins to ssDNA

The possible explanation for observed helicase activity for those mutants which had much lower ssDNA binding affinity could be due to the effect of ATP in the reaction. Therefore, a nitrocellulose filter binding assay was performed in the presence of ATP. As shown in Figure 5 C, the amount of DNA bound in the presence of ATP is increased. While the wild-type protein, as well as mutant proteins, Q175A, D212N, V214A, G218A, show a 3 - 10 fold increase in the ssDNA bound in the presence of ATP. Mutant proteins, P210G, G211A, and PG210,211GA have less than 2 fold increase. This result is consistent in that these mutants showed high basal ATPase activity that was not stimulated in the presence of DNA. The ATP seems to induce the conformation of the MCM helicase to have a higher affinity for DNA.

4.3.6 The Cdc6-2 protein does not inhibit the G211A, P210G, PG210,211GA mutant proteins

The eukaryotic homologues of putative helicase loader and the origin binding proteins, Cdc6-1 and Cdc6-2 proteins from *M. thermotrophicus* have been shown to inhibit the helicase activity of MthMCM (104). This Cdc6 inhibition has been shown to be via protein-protein interaction, more specifically, between domain B and C of N-terminal portion of MCM, and Cdc6-1 WH domain and Cdc6-2 full length protein (35,81,117). Since it is possible that Cdc6 proteins specifically interact with these conserved residues found in domain C, the effect of Cdc6 proteins on wild-type and mutant MCM was determined using helicase assay. As shown in Figure 4-6, when forked substrate is used, the wild-type and all mutants show relatively similar level of

inhibition. This result suggests that these are not the residues that interact with Cdc6 proteins.

4.4 Summary and Discussion

The biochemical characterization of mutations of conserved residues found in N-terminal region of MCM helicase revealed variations in biochemical properties compared to the wild-type. Despite the variations observed in DNA binding property, oligomeric states, thermostability, and ATPase activity, all mutant proteins show helicase activity (Figure 4-4). Here possible roles of these conserved residues are discussed.

In mutant proteins, P210G, G211A, and PG210,211GA, ATPase activity and DNA binding are not stimulated in the presence of DNA or ATP, respectively, suggesting that these two functions are partially decoupled (Figure 4-2 and Figure 4-5 C). The P210 and G211 are located on domain C facing the catalytic domain. According to the recent structure of SsoMCM and MkaMCM2, these residues are in close proximity to the C-terminal helix 2, which is disrupted by a motif known as helix 2 insertion (H2-ins). They are in close proximity to second half of helix 2, which consists of hydrophobic residues. Deletion of H2-ins motif results in significant increase in ATPase activity in the presence of dsDNA and increased affinity for both ss and dsDNA (102). However this deletion mutant has no helicase activity (102). It is possible that P210 and G211 may have hydrophobic interaction with the helix 2 and contribute to the regulation of H2-ins function. Since mutations are made to hydrophobic residues, it is possible that the drastic effect could not be observed.

Interestingly, the D212N which is located just next to PG210, 211 shows almost no similarity in biochemical behavior. D212 is further away from the helix 2, and it is

possible that its conservation may not be due to its interaction with the C-terminal region. D212N shows similar activity to Q176A mutant protein which is located on the $\beta 7$ that runs across domains B and C (Figure 4-1). Q176A and D212N mutants showed drastic decrease in ssDNA binding but the binding was recovered in the presence of ATP. These mutants both show ATPase activity stimulated in the presence of DNA, although, D212N show relatively weaker ATPase activity compared to the wild-type. They both showed relatively broad gel filtration elution profile, suggesting mutations cause heterogeneous population. But this heterogeneity does not seem to inhibit any activity, since both mutant proteins have helicase activity.

The mutant proteins, V214A and G218A show almost no obvious difference in all biochemical properties characterized compared to the wild-type protein. According to their location and polar contacts, they are involved in main chain interaction to stabilize the $\beta 9$ hairpin that runs parallel to $\beta 3$ (see Table 4-3). If there are hydrophobic interactions in nearby residues with these conserved residues, mutations to Ala may not cause substantial changes. But it is also possible that the methods available cannot address the reasons for such high conservation.

It is worth noting that the DNA binding properties of these mutant proteins were different not only relative to the wild-type protein, but also between ss and dsDNA. The P210G, G211A, and PG210,211GA proteins showed better ssDNA binding yet slightly reduced binding to dsDNA compared to the wild-type MCM. However, both Q176A and D212N mutants show low ssDNA binding and relatively better dsDNA binding, that is opposite of the trend seen in the wild-type protein. It is well established that the highly charged residues inside the central cavity of the hexameric protein are

responsible for both ss and dsDNA binding (59,127). From this mutational analysis, inconsistent trends were observed in ss and dsDNA binding relative to the wild-type MCM. This suggests that mutations cause different “mode” of binding to ss and dsDNA. Recently, EM study of MthMCM revealed that long dsDNA wraps around the MCM helicase, suggesting the alternative site for MCM and DNA interaction (107). Thus, slight changes in structure of MCM can interfere with either ss or dsDNA binding. Further study will be required to determine the actual mechanism in which MCM discriminates between ss and dsDNA.

Table 4-1 List of primers used to generate mutant proteins

Mutation	Sequence (5'-3')
Q176A F	TTC CTG GAC ACC GCG ACA CTG AAA CTC
Q176A R	GAG TTT CAG TGT CGC GGT GTC CAG GAA
P210G F	CTG GTT GAC ACC CTC ACA GGT GGG GAT ATT GTG AGG GTG ACC
P210G R	GGT CAC CCT CAC AAT ATC CCC ACC TGT GAG GGT GTC AAC CAG
G211A F	CTG GTT GAC ACC CTC ACA CCC GCT GAT ATT GTG AGG GTG ACC
G211A R	GGT CAC CCT CAC AAT ATC AGC GGG TGT GAG GGT GTC AAC CAG
PG210,211GA F	CTG GTT GAC ACC CTC ACA GGT GCT GAT ATT GTG AGG GTG ACC
PG210,211GA R	GGT CAC CCT CAC AAT ATC AGC ACC TGT GAG GGT GTC AAC CAG
D212A F	CTC ACA CCC GGG AAT ATT GTG AGG GTG ACC
D212A R	GGT CAC CCT CAC AAT ATT CCC GGG TGT GAG
V214A F	CCC GGG GAT ATT GCA AGG GTG ACC GGC
V214A R	GCC GGT CAC CCT TGC AAT ATC CCC GGG
G218A F	GTG AGG GTG ACC GCT ACC CTC AGG ACG
G218A R	CGT CCT GAG GGT AGC GGT CAC CCT CAC

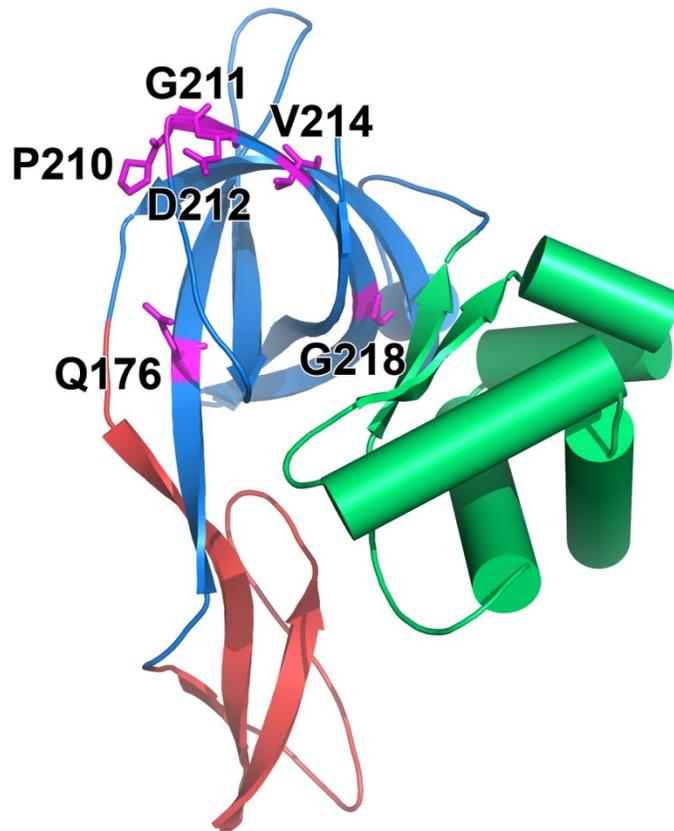


Figure 4-1 Locations of conserved residues in monomeric MCM.

Monomeric structure of N-terminal MCM helicase with domain A colored in green, B colored in red, and C colored in blue. The conserved residues that were mutated are highlighted and stick are shown in magenta. PyMOL was used to construct the figure (PDBID:1LTL).

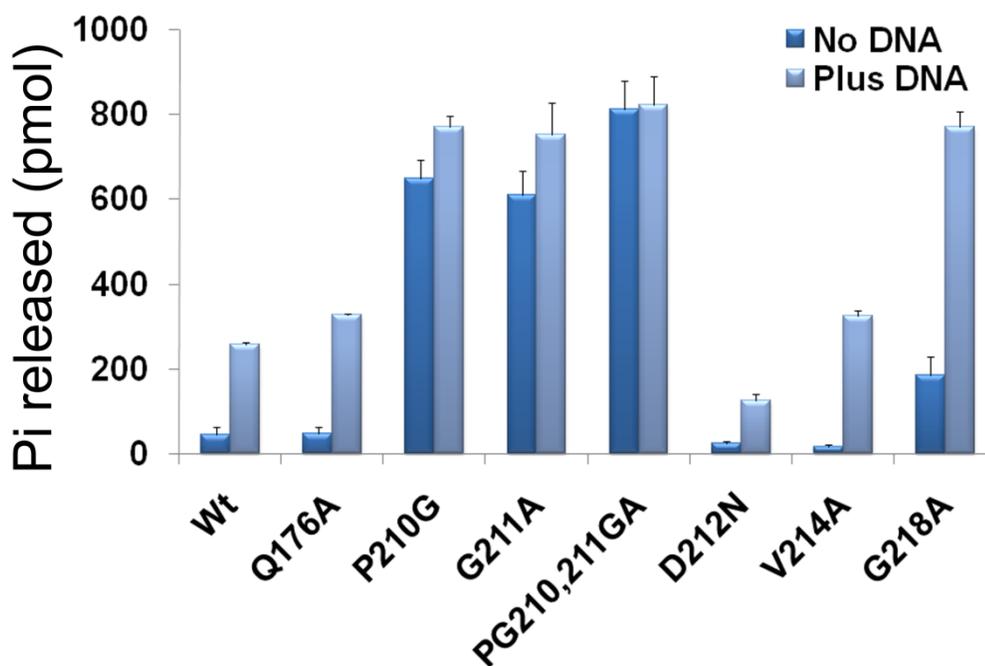


Figure 4-2 Mutations cause various levels of ATPase activity.

ATPase activity of wild-type and mutant MCM proteins in the absence and presence of ssDNA was determined as described in “Materials and Methods” using 30 nM of MCM (as monomer) in the presence or absence of 50 ng of ssDNA (49mer).

Table 4-2 List of mutant proteins and their location and polar contacts.

The PyMOL was used to determine the location and polar contact residues.

Mutation	Location	Polar contacts
WT	N/A	
Q176A	β 7	Sc (side chain) of O from Q176 interacts with bb (backbone) N of Asp206 and sc N interacts with bb O of Asp120. Bb O and bb N of Q176 interact with bb N and bb o of Leu 200.
P210G	1 aa before β 9	No polar contacts.
G211A	β 9	bbN of G211 interacts with bbO of Val116.
PG210,211GA	see above	See above.
D212N	β 9	sc O interacts with both bb N and sc OH of thr209
V214A	β 9	bb O and bb N of V214 interact with bb N and bb O of Gly114 to hold beta strands together
G218A	β 9	bb O and bb N of G218 interact with bbN and bb O of V110, it is in close proximity to F109

Table 4-3 DSC analyses of wild-type and mutant proteins.

	<u>T_m ave (°C)</u>	<u>ΔH_{VH} (kJ/mole)</u>	<u>ΔH_{cal} (kJ/mole)</u>
Wild type	66.2 ± 1.7*	1246.8 ± 159.0	530.6 ± 3.1
Q176A	66.2 ± 0.8	886.0 ± 25.2	248.2 ± 56.1
P210G	66.1 ± 0.2	1003.7 ± 89.5	242.7 ± 135.5
G211A	67.2 ± 0.2	1105.5 ± 21.9	335.5 ± 120.1
PG210,211GA	65.6 ± 1.1*	1137.5 ± 64.3	174.8 ± 60.3
D212N	68.1 ± 0.3	1191.4 ± 370.0	373.7 ± 156.5
V214A	65.7 ± 0.7	1066.0 ± 168.3	393.1 ± 142.2
G218A	65.9 ± 0.9	675.9 ± 136.4	425.8 ± 205.2

*Average of medium and slow scan.

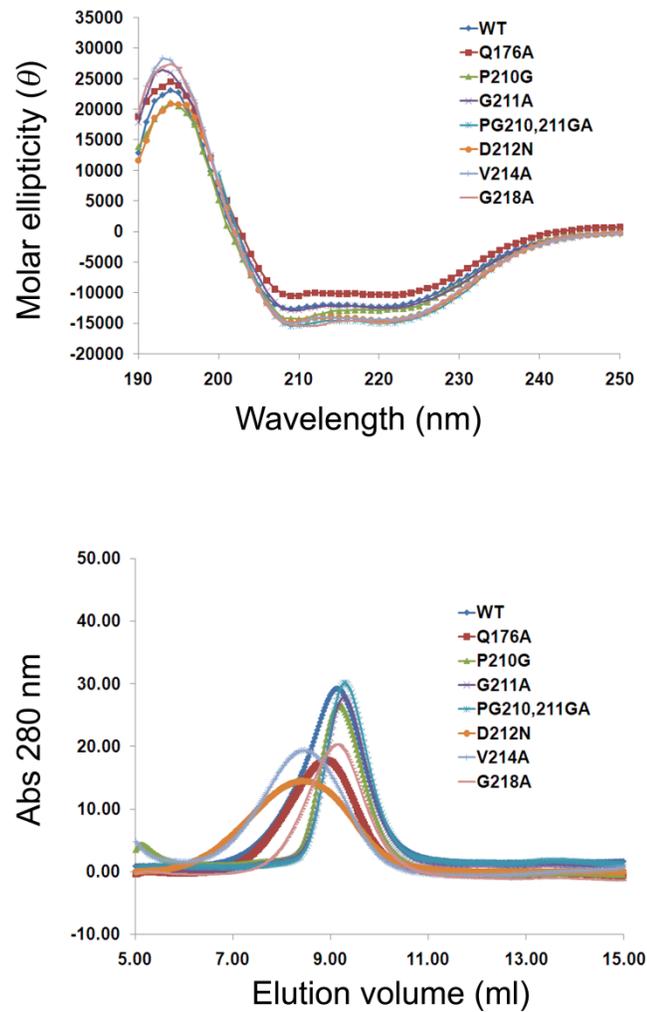
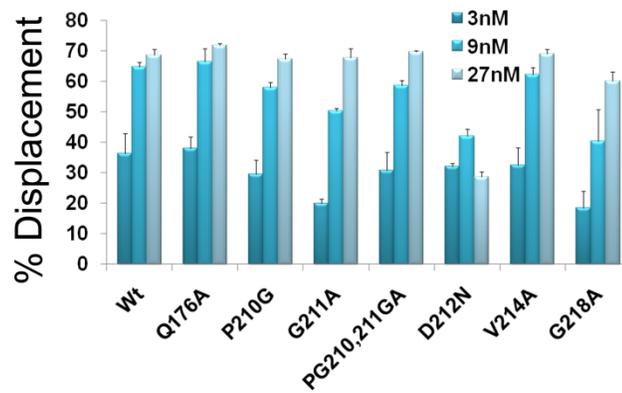


Figure 4-3 Mutations in the MCM protein alter the oligomeric state of MCM.

Circular dichroism (CD) measurements were performed as described in “Materials and Methods” (Top). The average result of three experiments is shown. Gel filtration analysis of wild-type and mutant proteins were performed as described in “Materials and Methods” (Bottom).

A) Forked substrate



B) Flat substrate

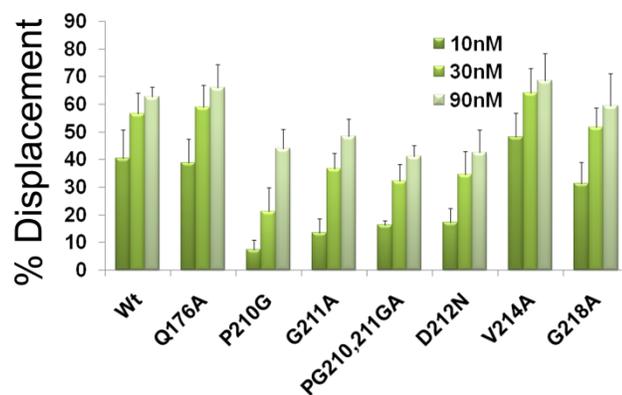


Figure 4-4 Mutations do not substantially affect helicase activity.

Helicase assays of wild-type and mutant MCM proteins were performed as described in “Materials and Methods” with 10 fmol of forked (A) or flat (B) substrates. In panel A the protein concentrations used were 3, 9 and 27 nM (as monomer) while in panel B they were 10, 30, and 90 nM (as monomer). The average results of three experiments are shown.

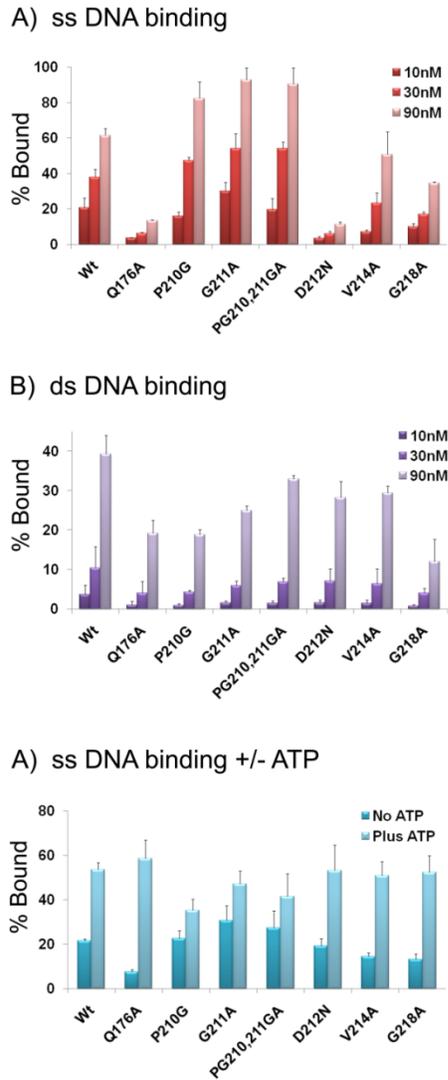


Figure 4-5 Mutations affect ss and dsDNA binding.

Nitrocellulose filter binding assays were performed as described under “Materials and Methods” using 50 fmol of 32 P-labeled single strand (A) or double strand (B) DNA (49-mer) in the presence of 10, 30, and 90 nM of proteins (as monomers). ss DNA binding assay with 25 nM of proteins (as monomers) in the absence or presence of 1 mM ATP (C). The average of three experiments is shown.

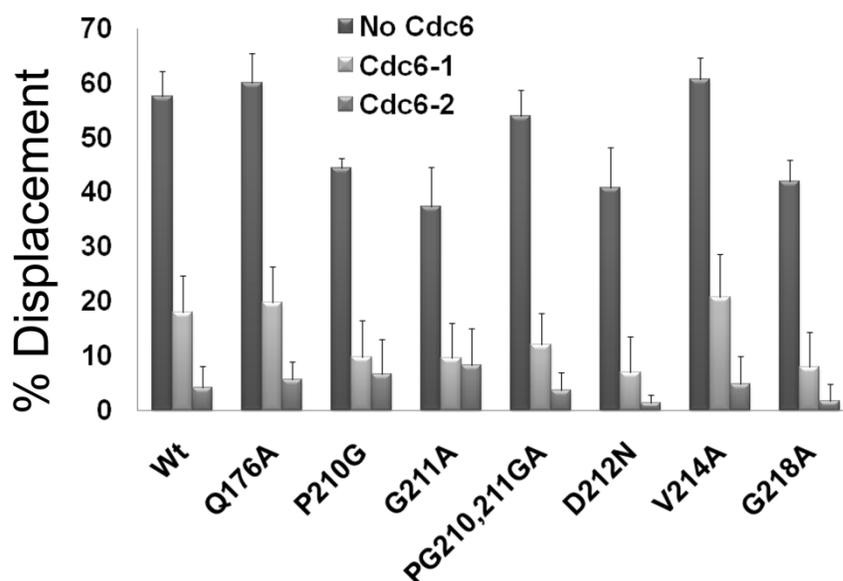


Figure 4-6 Cdc6 proteins do not interact with the conserved residues.

Helicase assays of wild-type and mutant MCM proteins were performed as described in “Materials and Methods” with 10 fmol of forked substrates. For both substrates, 10 nM of MCM was used (as monomer) in the presence or absence of 270 nM of Cdc6-1 and Cdc6-2. The average result of three experiments is shown.

Chapter 5 Effect of ATP hydrolysis and DNA binding on the thermostability of the MCM helicase

5.1 Abstract

Minichromosome maintenance (MCM) helicase is the replicative helicase in archaea. The enzyme utilizes the energy derived from ATP hydrolysis to translocate along one strand of the DNA and unwind the complementary strand. Here, the effect of ATP and DNA on the thermostability of the MCM protein from the archaeon *Methanothermobacter thermautotrophicus* was determined. It is shown that the thermostability of the enzyme is dramatically increase (about 10°C) in the presence of ATP and DNA. When ATP γ S was used, instead of ATP, however, no thermal stabilization could be observed. Thus, the results suggest that ATP hydrolysis is required for MCM stabilization in the presence of DNA. It was also found that MCM stabilization is dependent on the presence of the oligonucleotide.

5.1 Introduction

The minichromosome maintenance (MCM) helicase is thought to function as the replicative helicase in archaea. It uses the energy derived from ATP hydrolysis to translocate along one strand of the DNA and to unwind the complementary strand. In contrast to eukarya, in which MCM is a family of six related polypeptides, most archaeal species contain a single homologue of MCM. However, the single archaeal MCM homologue possesses biochemical properties which are similar to the eukaryotic enzyme (summarized in: (115,128)). One of the extensively studied archaeal MCM proteins is from the archaeon *Methanothermobacter thermautotrophicus* (Mth). The enzyme

possesses an ATP-dependent 3'→5' helicase activity, the ability to bind and translocate along single-strand (ss) and double-stranded (ds) DNA, unwind DNA-RNA hybrids while translocating along the DNA strand and can displace proteins from DNA (93).

The *M. thermautotrophicus* MCM protein is a member of the AAA+ family of ATPases and can be divided into three parts; a N-terminal part, a AAA+ catalytic region, and a C-terminal helix-turn-helix (HTH) domain. The N-terminal part is involved in multimerization and ss and dsDNA binding and regulation of helicase activity (59,81,117). The center part of the molecule consists of the AAA+ conserved motifs needed for ATPase and helicase activity (72,73) and the C-terminal part of the molecule was suggested to fold into a HTH domain involved in regulation (102).

The ATPase activity of the enzyme was shown to be stimulated in the presence of DNA (71-73) and mutations that abolish DNA binding also abolish the stimulation of ATPase activity (61). Conversely, ATP was shown to stimulate MCM binding to DNA (61,71), while ATP γ S had only limited effect on DNA binding. The stimulatory effect of ATP on DNA binding can only be observed when the reaction is incubated at 50-60°C which is the optimal growth temperature for *M. thermautotrophicus* and for *in vitro* helicase and ATPase activities. When the DNA binding reactions were incubated at 22°C, no change in DNA binding could be detected in the presence of ATP (72,102,103). These observations suggest that heat, ATP and DNA binding change the conformation of the enzyme. This is also supported by electron micrograph (EM) structural determination of MCM in the presence of nucleotides and DNA which revealed different conformations of the enzyme (87,88).

5.2 Materials and Methods

5.2.1 Materials

ATP and [γ - 32 P]ATP were obtained from GE Healthcare, AMP-PNP was obtained from Roche, ADP and ATP γ S were obtained from Sigma-Aldrich, ϕ X174 ssDNA was obtained from New England Biolabs and oligonucleotides were synthesized by the CARB DNA synthesis facility or Integrated DNA Technologies. All proteins used in the study were purified as previously described (81).

5.2.2 Differential scanning calorimetry (DSC) measurements

The purified proteins were dialyzed at room temperature against DSC buffer containing 20 mM HEPES-NaOH (pH 7.5), 100 mM NaCl and 10% glycerol. The protein concentrations were determined using an absorbance at 280 nm and a calculated extinction coefficient of ϵ_{280} of $28730 \text{ cm}^{-1} \cdot \text{M}^{-1}$. The solution outside the dialysis bag was retained and used as reference for the experiments. Oligonucleotide substrates were resuspended in DSC buffer. ϕ X174 DNA was dialyzed against H₂O, lyophilized and resuspended in DSC buffer. DNA concentration was calculated from absorption at A₂₆₀.

DSC measurements were performed using a VP-DSC Microcalorimeter from Microcal Inc. (Northampton, MA). The volume of the solution and the reference vessels were 0.511 ml and the scan rate was either at a slow rate of 15 K hr^{-1} or a medium rate 60 K hr^{-1} , with the temperature ranging from 25-85°C. Since the scans of protein samples were irreversible, the second scan for each sample was used as the baseline. After subtraction of the baseline from the protein scan, the resulting raw data of the differential power as a function of time were divided by the scan rate to convert the data into a heat

capacity versus temperature scan. Although the transitions did not re-appear upon a rescan of the solutions, the transition temperatures were found to be scan-rate independent and, thus, could be analyzed by application of a thermodynamic transition model to the results (124). Utilizing the EXAM program (122), a two state, $A \rightleftharpoons B$, transition model was then fitted to the heat capacity as a function of temperature to determine the van't Hoff enthalpy ($\Delta_{\text{trs}}H_{\text{VH}}$) for the scan from the shape of the transition peak; a transition temperature (T_G), and a calorimetric transition enthalpy ($\Delta_{\text{trs}}H_{\text{cal}}$) calculated from the area under the transition peak (mJ) divided by the moles of protein in the sample cell (concentration of protein x 0.511 ml). DSC scans on the samples were repeated twice.

For each experiment 10 μM MCM protein, 5 mM MgCl_2 , 1 mM nucleotides, 108 μg N190, PB4, N191, N192 oligonucleotides (Table 1) or ϕX174 ssDNA were prepared in 700 μl of DSC. The samples were incubated at 60°C for 10 min, placed on ice for 5 min, and centrifuged at 13,000 rpm for 10 min. The protein concentration of the supernatant was determined. The remaining of the sample was loaded to the DSC sample cell.

The binding of ligands, such as ATP and DNA, exclusively to the folded state of MCM results in a temperature shift of the transition temperature T_G to higher temperatures. Estimates of the temperature shifts at the transition temperature are as follows (129),

$$1/T_G - 1/T_G(\text{L}) = (R/\Delta_{\text{trs}}H) \ln\{1.00 + ([\text{L}_o] - [\text{P}_o]/2) K_b(T_G(\text{L}))\} \quad (1)$$

where $K_b(T_G(\text{L}))$ is the binding constant of the ligand at the transition temperature, T_G is the transition temperature of the MCM alone, $T_G(\text{L})$ is the transition temperature of the

complex, $\Delta_{\text{trs}}H$ is the transition enthalpy of the complex, $[L_0]$ is the total initial ligand concentration, and $[P_0]$ is the initial MCM concentration.

For extrapolation of the ligand binding constants determined at 25°C up to the transition temperature of the MCM-ligand complex, $T_G(L)$,

$$\ln[K_b(T_G(L))/K(T_0)] = [(-\Delta_b H(T_0) + \Delta_b C_p T_0)/R](1/T_G(L) - 1/T_0) + (\Delta_b C_p/R)\ln(T_G(L)/T_0) \quad (2)$$

where $\Delta_b H(T_0)$ is the binding enthalpy at $T_0 = 298.15$ K (25°C) and $\Delta_b C_p$ is the heat capacity change for the binding reaction.

5.2.3 ATPase assay

To study of nucleotide analogues ATPase activity was measured in reaction mixtures (15 μ l) containing 25 mM Hepes-NaOH (pH 7.5), 1 mM dithiothreitol, 5 mM $MgCl_2$, 100 μ g/ml bovine serum albumin (BSA), 50 ng ϕ X174 ssDNA, 1500 pmol of [γ - ^{32}P]ATP (3000 Ci/mmol; GE Healthcare) (100 μ M final ATP concentration) and 50 nM MCM protein (as monomer) in the absence or presence of ADP, AMP-PNP, ATP γ S, and ADP- AlF_4^- (10, 50, 100, 200, and 400 μ M). The ATP and the analogues were pre-mixed before their addition to the reactions. After incubation at 60°C for 60 min, samples were placed on ice, then an aliquot (1 μ l) was spotted onto a polyethyleneimine cellulose thin layer chromatography (TLC) plate, and ATP and P_i were separated by chromatography in 1 M formic acid and 0.5 M lithium chloride. The extent of ATP hydrolysis was quantitated by phosphorimager analysis. All ATPase assays were repeated three times, and their averages with standard deviations are shown in the figure.

To study the effect of the different DNA substrates, the reaction conditions described above were used without the nucleotide analogues and in the absence or

presence of 1 ng N190, N191, N192, oligonucleotides (Table 1) or ϕ X174 ssDNA. After incubation at 60°C for 30 and 60 min, samples were placed on ice, then an aliquot (1 μ l) was spotted onto a polyethyleneimine cellulose thin layer chromatography (TLC) plate, and ATP and P_i were separated by chromatography in 1 M formic acid and 0.5 M lithium chloride. The extent of ATP hydrolysis was quantitated by phosphorimager analysis. All ATPase assays were repeated three times, and their averages with standard deviations are shown in the figure.

5.2.4 Helicase assay

Substrates for helicase assays were made as previously described (44) using the DF74 and DF61 oligonucleotides (Table 1). DNA helicase activities were measured in reaction mixtures (15 μ l) containing 20 mM Tris-HCl (pH 8.5), 2 mM DTT, 10 mM MgCl₂, 100 μ g/ml BSA, 1 mM ATP, 10 fmol (0.66 nM) of ³²P-labeled substrate and 25 nM MCM protein (as monomer) in the absence or presence of ADP, AMP-PNP, ATP γ S and ADP-AlF₄⁻ at 0.2, 0.5, 1, 2, and 4 mM. The analogues were pre-mixed with ATP before added to the reaction. Following incubation at 60°C for 1 hr the reactions were stopped by adding 5 μ l of loading buffer containing 1% SDS, 100 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue, and 50% glycerol and chilling on ice. Aliquots (10 μ l) were loaded onto an 8% native polyacrylamide gel in 0.5X TBE and electrophoresed for 40 min at 180V. Gels were visualized and quantitated by phosphorimaging. All helicase experiments were repeated three times, and their averages with standard deviations are shown in the figure.

5.2.5 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) were performed using a MicroCal VP-ITC. The solution vessel (1.43 ml) contained 10 μ M of MCM proteins (as monomer) and 200 μ M DNA in the stirrer syringe at 25°C in buffer containing 25 mM Hepes-NaOH (pH 7.5), 100 mM NaCl, 10% glycerol and 10 mM MgCl₂. DNA was injected 4-8 μ l per injection.

5.2.6 Tryptophan fluorescence study

Intrinsic tryptophan fluorescence of MCM proteins were measured in the presence or absence of DNA using FluoroMax-3 (Jon Yvon). Each sample was prepared in a separate tube with a final volume of 150 μ l, containing 25 mM Hepes-NaOH (pH7.5), 2 mM DTT, 5 mM MgCl₂, 100 nM MCM protein (as monomer), with or without 1 μ g of N190, N191 N192 oligonucleotides (Table 1) or ϕ X174 ssDNA. The samples were incubated at room temperature for 10 min, and spectra were taken at 25°C. Excitation was set at 295 nm and emission spectra of 310 – 400 nm were taken.

5.3 Results

5.3.1 ATP increases the thermal stability of MCM in the presence of DNA

MCM activity is dependent upon ATP binding and hydrolysis, perhaps through alterations in the interactions between the structural domains of MCM upon ATP binding. In order to determine if this indeed occurs, DSC analysis of the wild-type enzyme was performed in the presence and absence of ATP (Table 5-2). A mutant protein (K325A) which cannot bind (125) nor hydrolyze (72) ATP was used as a control.

Another control was a β -hairpin (RK227,229AA) MCM mutant protein that cannot bind DNA (59,117). ATP binding alone increases the transition temperature of MCM (Table 5-2) and of the β -hairpin mutant (Table 5-3) about 5°C while no such increase was observed with the K325A mutant (Table 5-3). If the inverse of the K_M from the enzyme assay of ATP interacting with MCM near the transition temperature is assumed to be the binding constant of ATP to MCM, then the temperature shift can be estimated from Equation (1). The K_M is 0.4 mM at 60°C (data not shown) and this would produce at 1 mM ATP and 10 μ M MCM concentrations and a transition enthalpy of 366 kJ mol⁻¹ a shift of the transition temperature to 69.2°C, close to the observed value of 71 \pm 0.6°C in Table 2.

In addition to ATP binding, the helicase also interacts with DNA. It was shown that DNA binding to MCM changes the structure of the helicase (88), and, thus, may also alter the interactions between the structural domains of MCM, resulting in changes in its conformational stability. To determine the effect of DNA on MCM stability, DSC experiments were performed in the presence and absence of oligonucleotide. As shown in Table 5-2 and 5-3, DSC scans of the wild-type MCM, the K325A and the β -hairpin mutant proteins are unaffected by the presence of DNA in the sample. This lack of a transition temperature shift indicates that DNA does not bind to MCM at the transition temperature in the absence of any co-factor, although at 25°C, MCM binds to ssDNA with a high affinity of 7.5 X 10⁶ M⁻¹ (103). If the binding constant of DNA at the transition temperature shifts the temperature within the uncertainty of 1°C in the T_G values, then the binding constant at the transition temperature would be less than 1.245 X 10² M⁻¹ according to Equation (1). This would imply a binding enthalpy ($\Delta_b H(25^\circ\text{C})$) of -

226 kJ mol⁻¹ in the absence of any heat capacity change, $\Delta_b C_p = 0$ according to Equation (2). Isothermal titration calorimetry measurements performed at 25°C on the binding of DNA to MCM were, however, unsuccessful since the DNA-MCM binding enthalpy at this temperature is close to zero. Since $-\Delta_b H(25^\circ\text{C}) = 0$ in Equation (2), then a reduction of K_b from $7.5 \times 10^6 \text{ M}^{-1}$ at 25°C to below $1.245 \times 10^2 \text{ M}^{-1}$ at 66°C would yield a binding heat capacity change of $\Delta_b C_p = 443 \text{ J K}^{-1} \text{ mol}^{-1}$, a heat capacity change characteristic of the unfolding of a protein in solution.

The effect of the cofactor ATP on the conformational stability of MCM in the presence of DNA was also determined by DSC. When the wild-type MCM was studied in the presence of ATP and DNA, two transition peaks were observed at higher transition temperatures as shown in Figure 5-1 and summarized in Table 5-2. Such enhancement in stabilization was not observed for the K325A mutant of MCM, which does not bind to ATP, and for the β -hairpin mutant protein, which does not bind to DNA (Table 5-3). The presence of two unfolding transitions would indicate the independent unfolding of two thermodynamic domains of MCM that could be identified with two structural domains of MCM upon binding of ATP and DNA to the MCM protein. The lower temperature transition exhibits a transition temperature within the experimental uncertainty of the transition temperatures of MCM in just the presence of ATP. Then the higher temperature transition must be that of the DNA binding to its domain with a binding affinity enhanced by the presence of ATP. The average ratios of $\Delta_{\text{trs}} H_{\text{cal}} / \Delta_{\text{trs}} H_{\text{VH}}$ for MCM with either ATP or DNA have an upper limit of 0.8, close to one so that the two domains unfold as a single entity in the absence and presence of either DNA or ATP. The higher transition temperature reveals a substantial stabilization of the enzyme in comparison to MCM

alone (about 8°C). This temperature increase would correspond to a binding constant of $3.14 \times 10^4 \text{ M}^{-1}$ for DNA binding to its structural domain at the transition temperature according to Equation (1).

Since MCM can both bind and hydrolyze ATP, the bimodal transition stabilization may be due to an allosteric effect between the domain binding ATP and the domain binding DNA. In order to determine the source of this effect between the two domains of MCM, DSC measurements were performed on MCM in the presence of the ATP analogues ADP-AlF₄⁻, AMP-PNP, ATPγS and ADP. ADP-AlF₄⁻, AMP-PNP and ATPγS have been used as non-hydrolysable analogues of ATP while ADP is the product of ATP hydrolysis. In several systems, ADP-AlF₄⁻ was also shown to be a transition state analogue in the hydrolysis of ATP, mimicking the γ-phosphate of ADP-AlF₄⁻ is attacked by a water molecule ((130) and references therein).

The different analogues exhibit different effects on the thermostability of MCM (Fig. 5-1 and Table 5-2). AMP-PNP has very little effect on the stability of the enzyme in comparison to the MCM whether DNA is present or not (Table 5-2) and this may suggest that AMP-PNP either does not bind to MCM or binds it very weakly (discuss later). ADP stabilized the enzyme almost the same as that of ATP and this stabilization did not change in the presence of DNA. When ATPγS was used as a co-factor in the absence of DNA, the stability of the enzyme was the same as in the presence of ATP, indicating that it has the same binding affinity to MCM as ATP at 60°C. The stabilization effects of the different analogues according to the transition temperatures, are in the order ATPγS > ATP > ADP-AlF₄⁻, > ADP > AMP-PNP (Table 5-2). The additional presence of DNA had no effect on the thermal stabilities of MCM complexed

with ATP γ S, ADP, and AMP-PNP. The results of the DSC experiments with ADP, AMP-PNP, and ATP γ S suggest that ATP and ADP-AIF $_4^-$ are alone responsible for the allosteric effect between the two domains as evident in the bimodal stabilization of MCM in the presence of DNA. While the first transition of the ADP-AIF $_4^-$ MCM complex was similar to that observed with ATP and all the other analogues, the second transition of the complex shows a substantial stabilization effect on MCM of about 12°C. This larger shift in the high temperature transition temperature, larger than that observed in the presence of ATP and DNA, would result from a higher binding constant of $2 \times 10^6 \text{ M}^{-1}$ of DNA binding to its structural domain according to Equation (1). Thus, although both ATP and ADP-AIF $_4^-$ stabilized the MCM in the presence of DNA, ADP-AIF $_4^-$ has greater effect on the stability. If ADP-AIF $_4^-$ represents a transition state analogue during the hydrolysis process, as suggested in several systems (130,131), the DSC data with this analogue may suggest that the transition state complex is responsible for the communication between the two domains, whether the transition state complex is produced in the hydrolysis of ATP or by the binding of ADP-AIF $_4^-$ to MCM.

5.3.2 ATP γ S and ADP-AIF $_4^-$ are efficient competitors of ATP Binding

The data presented in Table 5-2 show that all nucleotide analogues affect the stability of MCM, though to different extents with shifts in the transition temperatures ranging from $68.2 \pm 0.2^\circ\text{C}$ for AMP-PNP to 72.5°C for ATP γ S, suggesting that they all interact with the helicase. Their ability to bind to MCM was demonstrated more directly by their ability to compete for binding with ATP in ATPase (Fig. 5-2) and helicase (Fig. 5-3) assays. As shown in Figures 5-2 and 5-3, all analogues could compete with ATP in

the order $\text{ATP}\gamma\text{S} > \text{ADP-AlF}_4^- > \text{ADP} > \text{AMP-PNP}$. This is in the same order as their increases on the stabilization of the MCM. This is an unexpected result as AMP-PNP was used as a non-hydrolysable analogue of ATP in a number of studies (for example: (44,72)). The data presented here suggest that it may not be the right choice for an analogue.

5.3.3 Binding to DNA ends required for MCM stabilization

The data presented in Table 5-2 and Figure 5-1 show the presence of two transitions when MCM is incubated with short oligonucleotides in the presence of ATP or ADP-AlF_4^- . The presence of two transition peaks suggests the unfolding of two structural domains of MCM with ATP or ADP-AlF_4^- bound to one structural domain and DNA bound to the other structural domain. It is possible that when provided with short oligonucleotides the N-terminal domain of MCM can either interact with the end of the DNA or/and the DNA can bind in the central cavity. This will result in two different structures of the DNA-MCM complex and, thus, may result in two different transitions with different stabilities. To check this hypothesis, the effect of short oligonucleotide and closed circular ssDNA on MCM stability in the presence of ATP or ADP-AlF_4^- was evaluated (Table 5-4, Fig. 5-4). In the absence of ATP and ADP-AlF_4^- , a single thermal denaturation transition appears for the unfolding of MCM in the presence of the oligonucleotide or plasmid DNA (Table 5-4 and Fig. 5-4). However, when ATP or ADP-AlF_4^- was used together with the DNA, the type of DNA used has a major effect on the thermostability of MCM as shown in Table 5-4 and Figure 5-4. In the presence of a oligonucleotide or plasmid DNA, two transition peaks are observed. In the presence of a circular single-stranded plasmid DNA (ϕX174) and ATP or ADP-AlF_4^- , only a single

transition peak but at a higher transition temperature was detected. This single transition is at the same temperature as the first transition peak (lower temperature) observed with the oligonucleotide (Table 5-2 and Fig. 5-1), as would be expected for ATP or ADP- AlF_4^- binding to the structural domain undergoing unfolding at this temperature. The results are not due to the inability of MCM in binding to the circular DNA since it has been demonstrated that the helicase can assemble on circular DNA (71) and binding does occur as shown by the reduction of fluorescence in the presence of DNA (Fig. 5-5).

The observed differences between the stability of MCM in the presence of oligonucleotide and circular DNA in the presence of ATP or ADP- AlF_4^- suggest that the protein interacts with an end of the DNA. One possibility is that the enzyme does not interact with any end but has a preference for binding to either 3' or 5' ends. To test this possibility, the experiment was repeated in the presence of oligonucleotides that contain either only 3' ends, 5' ends or both (Table 5-1). No differences in the thermal stabilities could be observed between the three different DNAs and all the DSC scans resulted in two transition peaks separated by similar temperatures in the presence of ATP or ADP- AlF_4^- (Table 5-4 and Fig. 5-4). In order to see whether the length of the oligonucleotide has an effect on the thermostability of MCM, a 136-mer oligonucleotide (PB4) was used in the presence of ATP. The results were similar to those obtained with the 50-mer (N190) oligonucleotide (data not shown).

Qualitatively, however, the three oligonucleotides differ in their effect on MCM stability (Fig. 5-4). When ATP is used in the presence of the oligonucleotide containing both 3' and 5' ends the two transition peaks appear equivalent (Fig. 5-4B). In the presence of the oligonucleotide containing only 3' ends the distribution of the intensities

of the transition peaks shifts to the second (higher temperature) peak (Fig. 5-4B), while the one containing only a 5' ends has the opposite effect and the equilibrium distribution of peak intensities was shifted toward the first transition peak making it more similar to circular DNA (Fig. 5-4B). When ADP-AlF₄⁻ was used instead of ATP all three oligonucleotides resulted in equilibrium shift toward the second peak (Fig. 5-4C). Nevertheless, taken together, the results suggest that although MCM protein binding to DNA ends stabilized the enzyme, it does not have specific preference to 3' or 5' ends.

The presence of DNA end not only affects the stability of MCM protein but also its ATPase activity. The ATPase activity of MCM is stimulated in the presence of DNA (71-73). Therefore, it is possible that DNA ends will have different stimulatory effect than circular DNA. As shown in Figure 5-6 this is indeed the case. While circular single-stranded DNA stimulates the activity about 2 fold, the oligonucleotides stimulate the activity about 7-9 fold. These results suggest that the different binding mode to DNA affect the catalytic ATPase activity of MCM in addition to the effect of stability of the enzyme.

5.4 Discussion

MCM utilizes the energy of ATP hydrolysis to translocate along one strand of the DNA and unwind the complementary strand. The ATP binds to the catalytic domain while the DNA binds to the N-terminal DNA binding domain. Based on the three dimensional structures of other helicases, it is clear that the enzymes change their conformations when they bind to nucleotides and DNA (99,132) and these changes alter the interactions between the domains. These changes upon binding DNA alone are also evident in the large positive heat capacity change implied by Equations (1) and (2). In

this study, the thermostability of MCM was used as an indicator of any conformational change. The data show that the stability of the enzyme is slightly increased (2-4°C) when the protein binds to ATP or its analogue. This is validated in particular for ATP since the temperature increase of 4°C is predictable from Equation (1) by employing the ATP binding constant at 60°C. This temperature shift is maintained for the lower temperature transitions in the presence of DNA and, thus, likely results from the unfolding of the catalytic domain binding the ATP or its analogue ADP-AIF₄⁻. The other nucleotide analogs that exhibit only a single transition in the presence of DNA also exhibit temperature shifts similar to those observed for the lower temperature transition of MCM in the presence of ATP or ADP-AIF₄⁻ and DNA. However, there is a major difference in the thermostability of the high temperature peak between MCM binding to ATP and ADP-AIF₄⁻ in the presence of DNA with the stability of the second transition being further enhanced in the present of ADP-AIF₄⁻ (summarized in Fig. 5-7). This high temperature transition has been tentatively assigned to the unfolding of the N-terminal DNA binding domain.

The inability of the ATP analogues other than ADP-AIF₄⁻ to stabilize the enzyme to the same extent as ATP so as to produce a bimodal transition pattern, may suggest that hydrolysis is required and hydrolysis proceeds through formation of a transition state complex. This idea may be supported by the observation, as several studies have suggested, that the binding of ADP-AIF₄⁻ to MCM may mimic a transition state complex in which the γ -phosphate is forming a phosphoryl transfer transition coordinate with the MCM (130). As summary shown in Figure 5-7, we suggest that the transition state (ADP-Pi state) in the presence of DNA is possibly the most stable state. ATP binding

itself, on the other hand, does not significantly contribute to the change in the presence of DNA, and thus not as stable as the transition state. It is supported by the observation that the further stabilization of the higher temperature transition of MCM in the presence of ADP-AlF₄⁻ in comparison to ATP (about 5°C). This suggest that the transition state complex between MCM and ADP-AlF₄⁻ formed during hydrolysis is more effective in increasing the binding affinity of DNA to MCM. The stabilization of transition state in the presence of DNA may explain the cause for the stimulation of ATPase activity observed in the presence of DNA.

It was also found that the effect of ATP and ADP-AlF₄⁻ on the stability of MCM could be observed only in the presence of oligonucleotides and not with close circular ssDNA. Thus, suggesting that the enzyme binds differently to the DNA end in comparison to circular DNA. The physiological role for MCM interactions with DNA ends is not clear. During chromosomal DNA replication the helicase will encounter DNA ends at sites of DNA breaks. The binding of MCM to the DNA ends may serve as a cellular signal for DNA breaks. MCM was suggested to play a role in S-phase checkpoint which is activated upon DNA damage (40). The specific binding of MCM to the DNA end may be used in these checkpoint processes. MCM was also implicated to have other cellular roles in addition to its function as the replicative helicase. It was suggested that the MCM may play a role in transcription (133) and was implicated in chromosome remodeling (134,135). MCM may need to bind the DNA end during its roles in those processes.

Table 5-3 Thermostabilities of mutant MCM proteins from DSC Results

MCM mutant proteins (10 μ M) in the absence or presence of 1 mM nucleotide and 10 μ M N190 DNA substrate was incubated at 60°C for 10 min then loaded onto the cell.

Mutant	Nucleotide	DNA	T_G °C	$\Delta_{\text{trs}}H_{\text{vH}}$ kJ mol ⁻¹	$\Delta_{\text{trs}}H$
K _{325A}	-	-	68.1 ± 0.8	923 ± 14	396 ± 7
	-	+	67.7 ± 0.1	814 ± 23	412 ± 27
	ATP	-	68.3 ± 0.2	950 ± 44	336 ± 7
	-	+	68.1 ± 0.2	1015 ± 181	298 ± 203
	ADP-ALF ₄ ⁻	-	68.7 ± 0.8	887 ± 30	584 ± 204
	-	+	68.6 ± 0.8	889 ± 63	370 ± 56
	ATPγS	-	68.2 ± 0.2	433 ± 273	599 ± 323
	-	+	68.2 ± 0.2	721 ± 211	426 ± 228
	AMP-PNP	-	68.4 ± 0.2	849 ± 80	484 ± 81
	-	+	67.6 ± 0.2	944 ± 68	384 ± 8
	ADP	-	68.8 ± 0.4	893 ± 1	422 ± 19
	-	+	68.4 ± 0.3	847 ± 75	385 ± 37
β-hairpin	-	-	65.9 ± 0.6	1379 ± 76	305 ± 95
	-	+	65.3 ± 0.1	1034 ± 173	317 ± 56
	ATP	-	70.9 ± 0.3	787 ± 62	494 ± 66
	-	+	70.2 ± 0.3	819 ± 33	386 ± 97
	ADP-ALF ₄ ⁻	-	68.2 ± 1.0	947 ± 134	299 ± 23
	-	+	68.7 ± 0.4	942 ± 10	283 ± 26
	ATPγS	-	71.6 ± 0.3	642 ± 112	345 ± 208
	-	+	71.1 ± 0.1	486 ± 8	560 ± 250
	AMP-PNP	-	66.4 ± 0.2	1129 ± 97	283 ± 26
	-	+	66.2 ± 0.3	1064 ± 37	342 ± 23
	ADP	-	68.7 ± 0.0	748 ± 58	479 ± 63
	-	+	68.3 ± 0.0	903 ± 51	322 ± 28

Table 5-4 Thermostabilities of wild type MCM from DSC Results

MCM mutant proteins (10 μ M) in the absence or presence of 1 mM nucleotide and 10 μ M N190 DNA substrate was incubated at 60°C for 10 min then loaded onto the cell.

Nucleotide	DNA	T_g °C	$\Delta_{\text{trs}}H_{\text{trH}}$	$\Delta_{\text{trs}}H$ kJ mol ⁻¹
-	-	66.7 \pm 1.1	836 \pm 77	376 \pm 149
	3'-5'	65.9 \pm 0.6	693 \pm 228	378 \pm 8
	3'-3'	66.2 \pm 0.2	713 \pm 48	395 \pm 35
	5'-5'	66.5 \pm 0.1	737 \pm 19	360 \pm 32
	ϕ X174 (med)	66.8	791	465
	ϕ X175 (slow)	64.4	1098	378
ATP	-	71.5 \pm 0.6	731 \pm 33	569 \pm 30
	3'-5'	68.9 \pm 0.5, 75.3 \pm 0.3	639 \pm 95, 981 \pm 65	625 \pm 68, 179 \pm 121
	3'-3'	68.7 \pm 0.4, 75.2 \pm 0.2	647 \pm 23, 913 \pm 49	574 \pm 39, 287 \pm 4
	5'-5'	69.0 \pm 0.6, 73.5 \pm 0.3	666 \pm 25, 931 \pm 104	470 \pm 100, 152 \pm 28
	ϕ X174	68.8 \pm 0.0	874 \pm 115	486 \pm 283
ADP-ALF ₄ ⁻	-	70.3 \pm 0.5	778 \pm 4	433 \pm 91
	3'-5'	69.5 \pm 0.4, 80.9 \pm 0.5	762 \pm 77, 1289 \pm 1	106 \pm 42, 283 \pm 35
	3'-3'	68.7 \pm 0.1, 81.4 \pm 0.0	690 \pm 33, 1029 \pm 15	254 \pm 115, 240 \pm 62
	5'-5'	68.5 \pm 0.4, 80.3 \pm 0.1	728 \pm 48, 962 \pm 42	178 \pm 64, 356 \pm 24
	ϕ X174	69.8 \pm 0.2	542 \pm 40	511 \pm 142

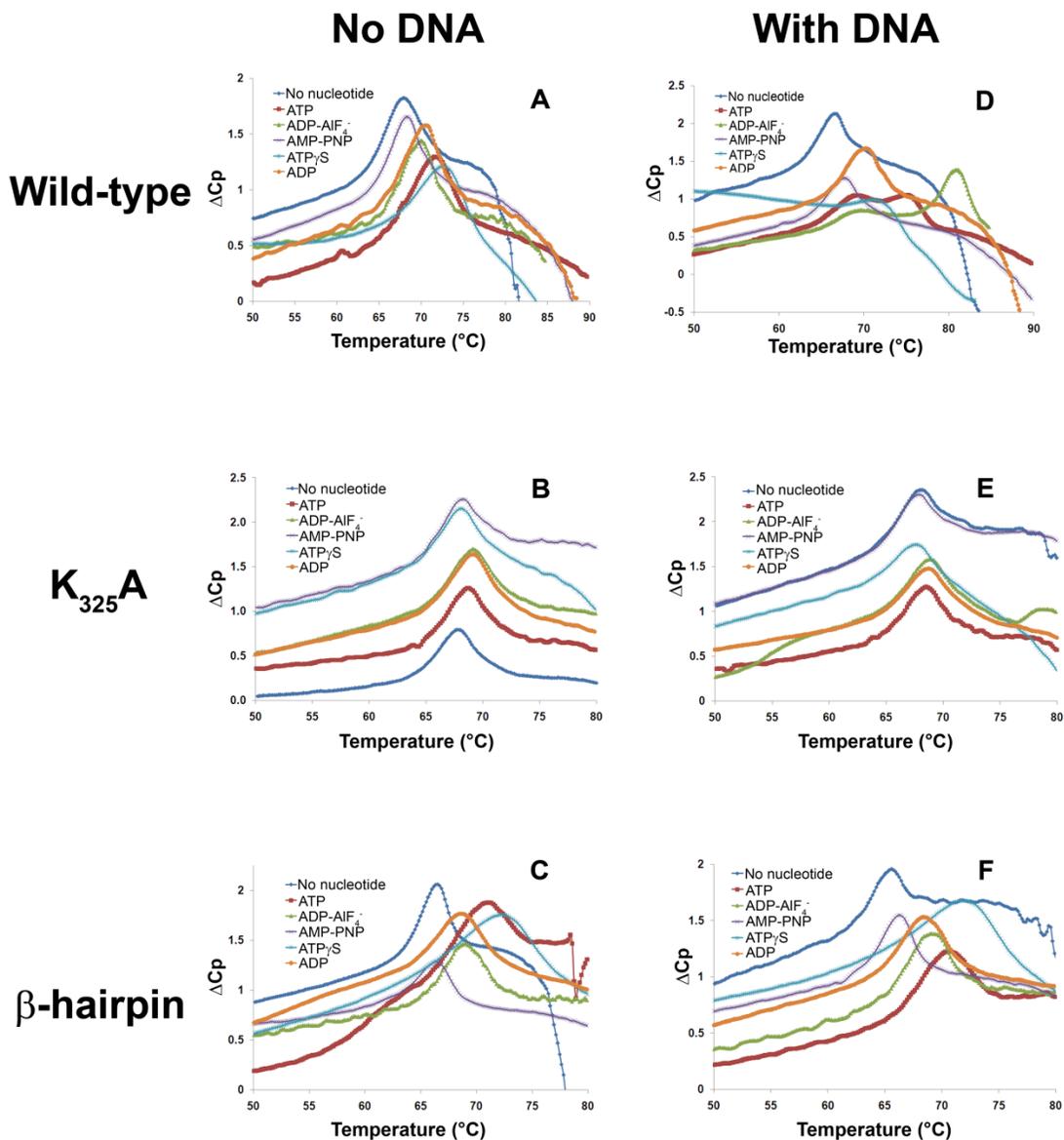


Figure 5-1 Effect of DNA and nucleotides on MCM thermostability.

DSC assays were performed as described in “Materials and Methods” using wild-type (A and D), K325A mutant (B and E) or β -hairpin mutant (C and F) MCM protein in the

absence (A-C) or presence (D-F) of N190 oligonucleotide. DSC results of excess heat capacity versus temperature are shown.

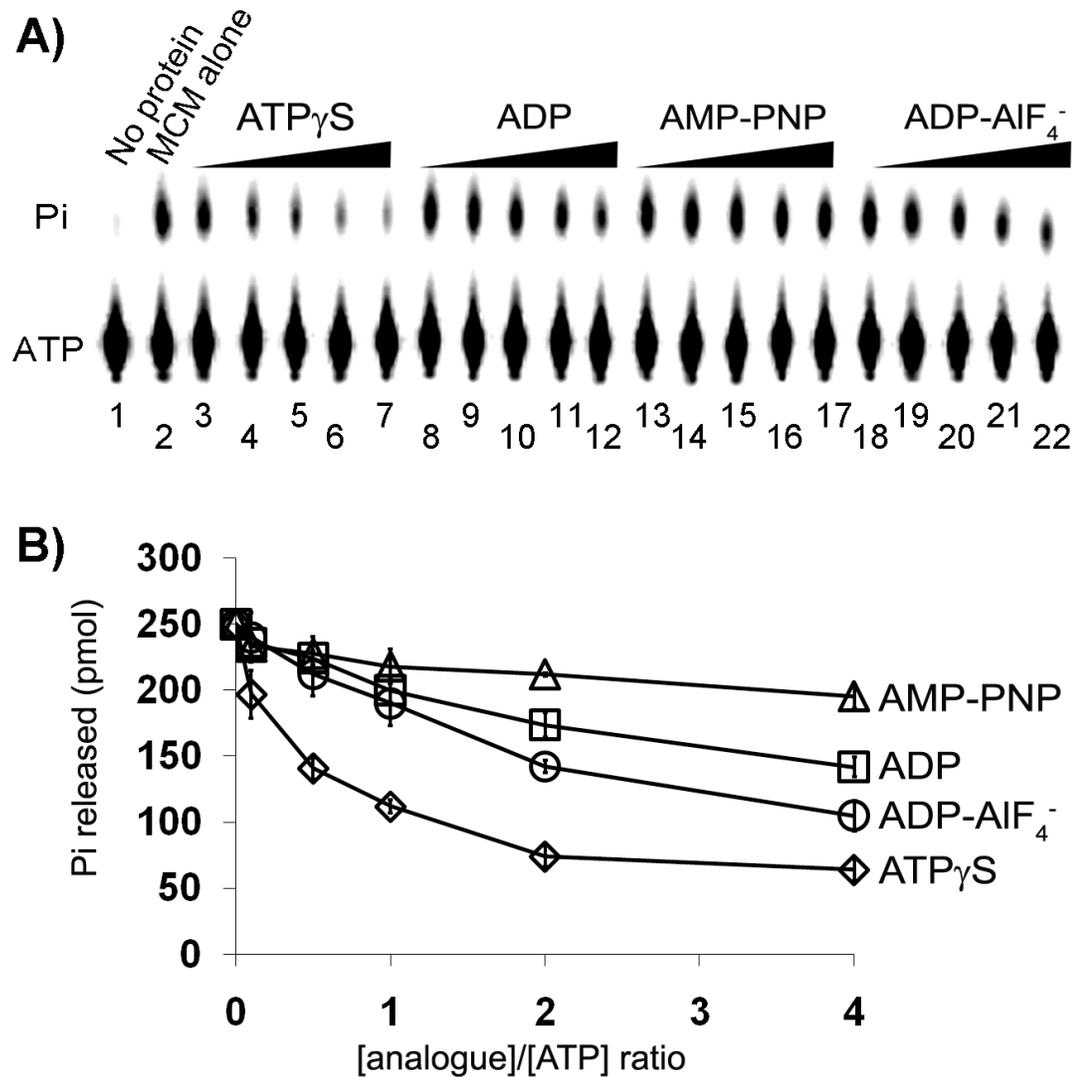


Figure 5-2 Effects of nucleotide analogues on MCM ATPase activity.

ATPase assays were performed as described in “Materials and Methods” in the absence or presence of 0.01, 0.05, 0.1, 0.2, and 0.4 mM of nucleotide analogue in the presence of 0.1 mM ATP. A) A representative image of ATPase inhibition assay. B) Average of three independent experiments with standard deviation.

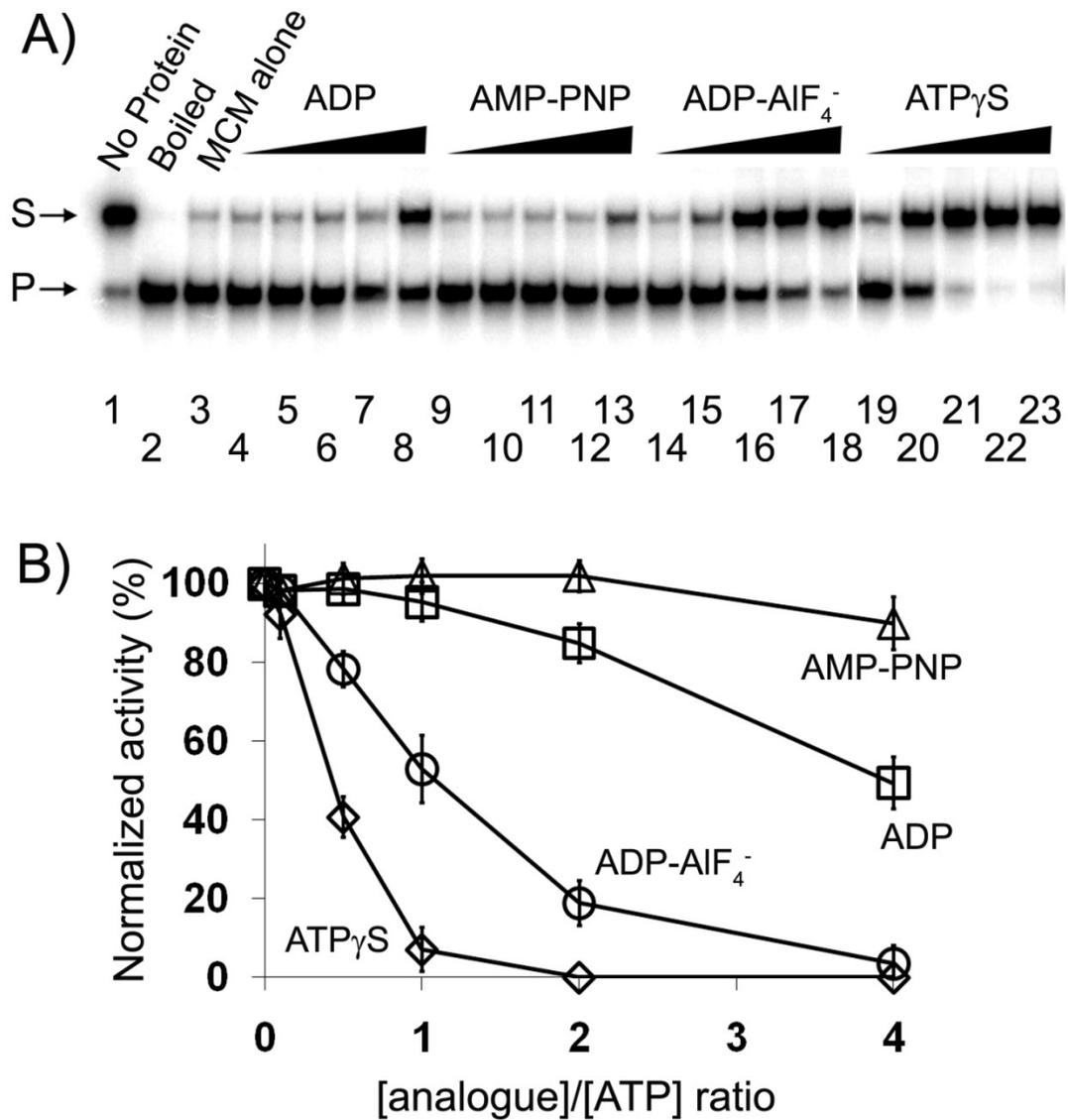


Figure 5-3 The effects of nucleotide analogues on MCM helicase activity.

Helicase assays were performed as described in “Materials and Methods” in the absence or presence of 0.1, 0.5, 1, 2, and 4 mM of nucleotide analogue in the presence of 1 mM ATP. A) A representative gel. B) Average of three independent experiments with standard deviation.

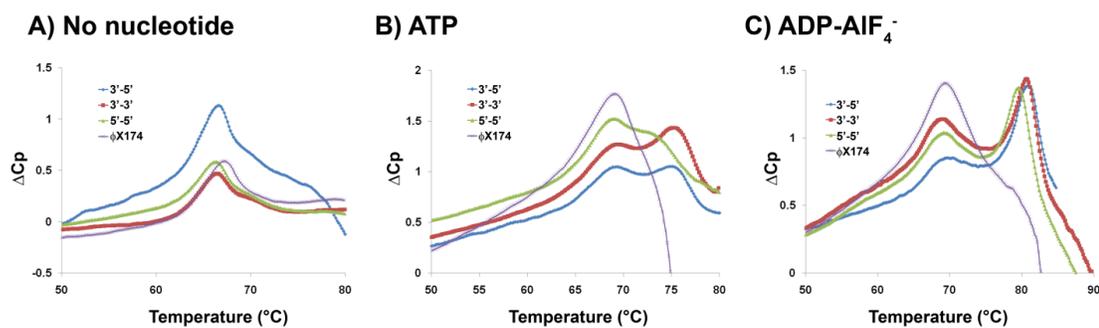


Figure 5-4 The effects of DNA substrates on MCM thermostability.

DSC assays were performed as described in “Materials and Methods” using wild-type MCM protein and N190 (3'-5'), N191 (3'-3'), N192 (5'-5') oligonucleotides or $\phi X174$ ssDNA. DSC results of excess heat capacity versus temperature are shown. A. no nucleotide; B. ATP; C. ADP-AIF₄⁻.

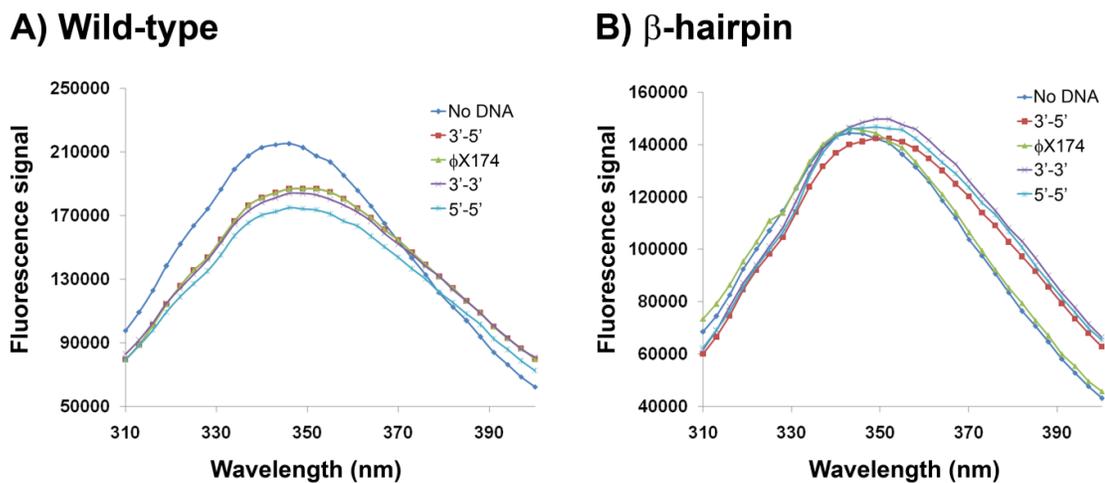


Figure 5-5 MCM interacts with DNA substrates.

Intrinsic tryptophan fluorescence of wild-type (A) or β -hairpin mutant (B) MCM in the presence of N190 (3'-5'), N191 (3'-3'), N192 (5'-5') oligonucleotides or ϕ X174 ssDNA were measured as described in "Materials and Methods". Samples were excited at 295 nm and emission spectra of 310 – 400 nm were taken.

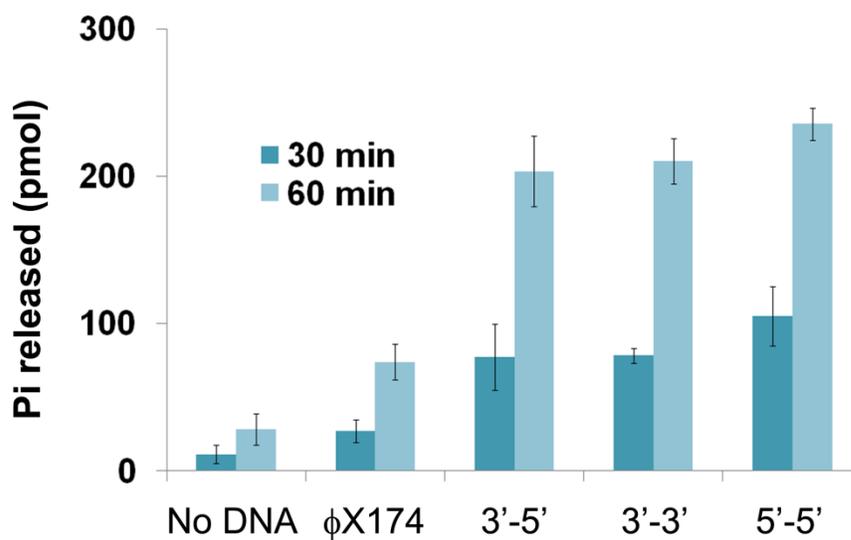


Figure 5-6 Effect of DNA on MCM ATPase activity.

ATPase activities of wild-type MCM protein (50 nM) measured in the absence or presence of 1 ng of N190 (3'-5'), N191 (3'-3'), N192 (5'-5') oligonucleotides or φX174 ssDNA as described in Materials and Methods. Reaction mixtures were incubated at 60°C and quenched after 30 and 60 min.

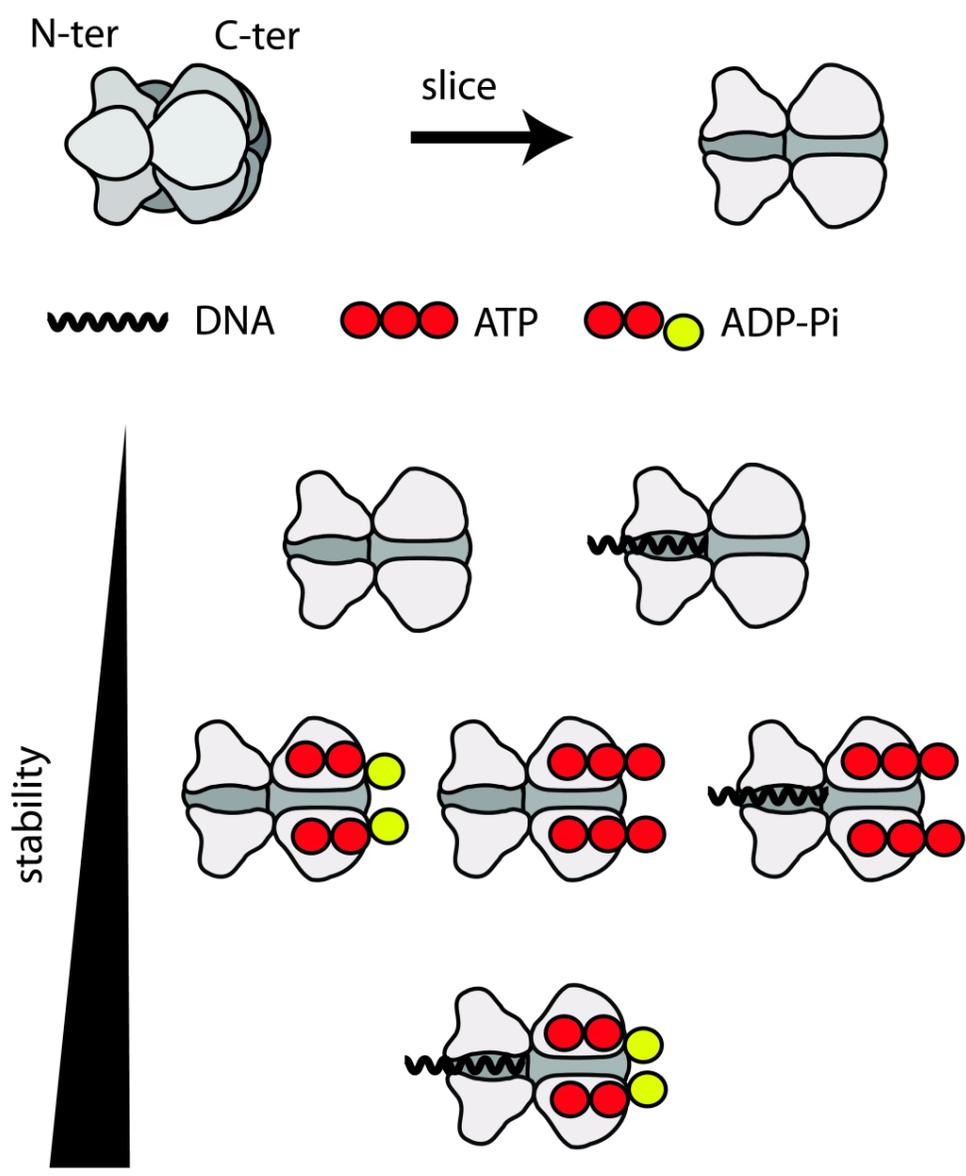


Figure 5-7 Summary of MCM protein thermostability in the presence of different co-factors.

Schematic representation of MCM thermostability in the presence of non-hydrolyzable ATP, hydrolysis transition state analogue, ADP•Pi and DNA. MCM proteins (divided

into two domains) have different thermal stability in the presence of non-hydrolyzable ATP (ATP in three red circles), transition state analogue (ADP-Pi in two red circles with one yellow ball), and DNA. The C-terminal domain binds to and hydrolyzes ATP, and the N-terminal domain may be involved in DNA binding. In the presence of both DNA and hydrolysis transition state analogue, MCM changes the conformation and decouples two domains thus giving rise to two transitions. The second transitions in the presence of transition analogues are significantly higher (10~12°C) than the nucleotide or DNA alone suggesting stronger interactions between these three components.

Chapter 6 Cloning, purification, and partial characterization of the *Halobacterium* sp. NRC-1 minichromosome maintenance (MCM) helicase

6.1 Abstract

The MCM gene from the archaeon *Halobacterium*, with and without its intein, was cloned into an *Escherichia coli* expression vector, overexpressed and the protein was purified and antibodies were generated. The antibodies were used to demonstrate that *in vivo* only the processed enzyme, without the intein, could be detected.

6.2 Introduction

Chromosomal DNA replication is an essential process for all organisms. It ensures the accurate duplication of the genetic information prior to cell division. The process is highly controlled and coordinated, as aberrant DNA replication can lead to malignancies and cancer. The process is conserved in all organisms including bacteria, archaea, eukarya, viruses, and bacteriophages.

Halobacterium sp. NRC-1 is a halophilic archaeon with an optimal growth temperature of 42°C and a 2 hr generation time (15). The genome of *Halobacterium* consists of 2.5 Mb and encodes for about 2,600 proteins. The entire genome consists of one large chromosome (2 Mb) and two extrachromosomal replicons, pNRC100 (191 kb) and pNRC200 (365 kb).

The minichromosome maintenance (MCM) complex is thought to function as the replicative helicase of archaea and eukarya (50,116). One MCM homologue has been identified in the genome of *Halobacterium*. In contrast to other archaeal MCM

proteins studied to date, the *Halobacterium* protein contains an intein. Inteins are encoded as a part of the protein. After the protein is synthesized the intein excises itself from the rest of the protein and ligate the ends resulting in a mature protein (MCM in this case) and the intein molecule (Figure 6-1).

In this study we describe the isolation, purification and *in vivo* identification of the single MCM helicase from *Halobacterium* sp. NRC-1. It was found that *in vivo* only the mature enzyme, from which the intein has been removed, could be identified.

6.3 Materials and Methods

6.3.1 Sequence alignment of MCM proteins

The *Halobacterium* sp. NRC-1 MCM protein sequence was aligned with four other archaeal MCM proteins (*Methanothermobacter thermautotrophicus*, *Thermoplasma acidophilum*, *Methanosaeta thermophila* and *Archaeoglobus fulgidus*) using the MUSCLE algorithm (<http://www.ebi.ac.uk/muscle/>).

6.3.2 Cloning the MCM gene

The *Halobacterium* sp. NRC-1 MCM gene was amplified using PCR from genomic DNA (kindly provided by Brian Berquist). The PCR reaction was performed with 100 ng genomic DNA, 200 μ M dNTPs, 5 units of OptimaseTM polymerase (Transgenomic), 1 x Optimase reaction buffer, 1.5 mM MgSO₄, 10% DMSO and 0.4 μ M of primers, N101 (5'-CCGCTCGAGGCTAGCCATATGGATCCGGACCTGGCCGACGATTACATCAGCC-3') and N102 (5'-CCGCTCGAGCCATGGCTAGATCGAGCGCAAGTGGTCCGTGTTC-3'), where the

XhoI restriction site is underlined, and the NdeI site shown in bold type. Twenty PCR cycles were performed as follows: denaturation at 95°C for 2 min, annealing at 70°C for 45 seconds, and elongation at 72°C for 5 minutes. The PCR product was purified using the QIAquick PCR

purification kit (Qiagen), digested with XhoI and ligated into the XhoI site of pBluescript KS- (Stratagene). The sequence was confirmed by sequencing and the clone was designated pBS-MCM. The gene encoding MCM was excised from the pBS-MCM vector by cleavage at the NdeI and XhoI sites and cloned into pET-16b vector (Novagene). This clone is designated pET-MCM.

To generate the MCM gene construct without the intein, a PCR-based method was used (81) using pBS-MCM as template. The PCR reactions were performed as described above with the following two primer sets to amplify the two exons. One set was N101 together with 5'-GGGGCGCGATGTTCTGCACGTAGGAGATCATCTGGGACTTCCCGGTTCCCGG GTCACC-3' and the other set was N102 together with 5'-GGTGACCCGGGAACCGGGAAGTCCCAGATGATCTCCTACGTGCAGAACATCG CGCCCC -3'. Two exons encoding PCR products were purified using the QIAquick purification kit and used as template (10 ng each) for PCR amplification using N101 and N102 as primers. The resulting PCR product contained the MCM gene minus the intein. This product was cloned into pET-16b, sequenced, and designated pET-MCMint.

6.3.3 Purification of the MCM proteins

Plasmids containing the genes encoding for the full-length MCM and the protein without intein were transformed into Codon plus *Escherichia coli* cells (Stratagene). Cells were grown in LB media at 37°C. When the OD₆₀₀ reached 0.6, protein expression was induced by the addition of IPTG (1 mM final concentration) and the cells were grown for an additional 4 hours. Cells were harvested and stored at -80°C. Protein purification was carried out at 4°C as follows. Cells were thawed on ice in lysis buffer containing 10 mM imidazole, 3 M NaCl, 0.5 M KCl, 20 mM Tris-HCl (pH 7.6) and 20% glycerol and disrupted by sonication. The lysate was centrifuged for 15 min at 15,000 rpm in JA-17 rotor (Beckman). The pellet was kept and the supernatant was incubated with Ni-column resin for 1 hr with gentle shaking. Following incubation, the resin was poured into a column and washed with lysis buffer containing 10 mM imidazole. The MCM protein was step eluted with 50, 100, 200 and 300 mM of imidazole.

Since only a small fraction of the induced MCM protein was found in the soluble fraction, the protein was also purified from the pellet using denaturing conditions in urea. The cell pellet was resuspended in buffer containing 8M urea and 20 mM Tris-HCl (pH 7.6) followed by centrifugation for 15 min at 15,000 rpm in a JA-17 rotor. The supernatant was incubated with Ni-column resin for 1 hr with gentle shaking. Following incubation, the resin was poured into a column and washed with lysis buffer containing 10 mM imidazole. The MCM protein was eluted using step elution in lysis buffer containing 6 M urea and 50, 100, and 300 mM imidazole. The fraction with the highest MCM concentration (300 mM imidazole) was dialyzed in buffer containing 20 mM Tris-

HCl (pH 7.6), 3 M NaCl, 0.5 M KCl, and 20% glycerol. The proteins were flash frozen in liquid nitrogen and kept at -80°C.

6.3.4 Mass spectrometry analysis of the *E. coli* expressed proteins

To confirm that the purified proteins are indeed the recombinant *Halobacterium* MCM proteins a matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometric (MS) analysis was used. MALDI-TOF MS was performed using an AB4700 Proteomics Analyzer (Applied Biosystems, Framingham, MA) using in-gel tryptic digest of Coomassie-stained protein bands. Dry peptide samples were dissolved in 5 mg/ml α -cyano-4-hydroxycinnamic acid in 50% acetonitrile containing 0.1% TFA, and manually spotted onto an ABI 01-192-6-AB target plate. MS-mode acquisitions consisted of 1,000 laser shots averaged over 20 sample positions. For MS/MS-mode acquisitions, 3,000 laser shots were averaged over 30 sample positions for PSD fragments. Automated combined acquisition of MS and MS/MS data was controlled with 4000 Series Explorer software 3.0. Data analysis was performed with GPS Explorer software 3.5 utilizing Mascot 2.0 (MatrixScience, London, UK) as the search engine. During searching, the mass tolerance was 0.08 Da for the precursor ions and 0.2 Da for the fragment ions. A protein was listed as identified protein when the MOWSE score was higher than the MOWSE score at which statistical significance ($p < 0.05$) occurred for that particular search.

6.3.5 Generation of antibodies

Urea purified MCM protein without intein was used to immunize a rabbit to generate polyclonal antibodies by Cocalico Biologicals, Inc.

6.3.6 Preparation of *Halobacterium* cell extract

Halobacterium sp. NRC-1 (ATCC number 700922) was grown in GN101 media (25 g/L NaCl, 20 g/L MgSO₄, 2 g/L KCl, 3 g/L sodium citrate, 10 g/L Oxoid brand bacteriological peptone) with the addition of 1 mL/L trace elements solution (31.5 mg/L FeSO₄·7H₂O, 4.4 mg/L ZnSO₄·7H₂O, 3.3 mg/L MnSO₄·H₂O, 0.1 mg/L CuSO₄·5H₂O) at 42°C with shaking at 220 rpm. Beveled flasks were used to ensure proper oxygenation. Cultures were centrifuged at 8000 x g and a cell pellet from 25 ml of culture was resuspended in 1 ml of buffer containing 50 mM potassium phosphate (pH 7.0), 1 M NaCl, and 10% β-mercaptoethanol. The resuspended cells were sonicated on ice followed by centrifugation at 13,000 rpm for 10 min at 4°C. The supernatant was kept at -20°C.

6.3.7 Western blotting analysis

MCM protein with intein (20 ng) and without intein (1ng) and *Halobacterium* cell extract (0.5 µg and 1 µg) were separated on 10% SDS-PAGE. The gel was then electroblotted onto nitrocellulose BA83 membranes (Whatman) followed by Western analysis using 1:500 dilution of rabbit anti-MCM polyclonal antibodies and goat anti-rabbit antibody coupled to horse radish peroxidase (Epicentrics) as secondary antibody. The blot was developed using ECL (GE Healthcare) followed by exposure to X-ray film.

6.4 Results and Discussion

6.4.1 Sequence analysis of *Halobacterium* MCM protein

Multiple alignment sequence analysis of *Halobacterium* sp. NRC-1 MCM protein revealed that the *Halobacterium* protein contains a large insertion (Figure 6-2). Careful analysis of the insertion sequence and its integration site (136,137) clearly suggest that it is an intein. The intein is inserted in the middle of the Walker-A motif (138). This is a highly conserved motif which is found in all members of the AAA+ family of ATPases (62,66). The sequence of the motif in *Halobacterium* MCM is GDPGTGKS and the intein is located between the highly conserved Lys and Ser residues.

The alignment also revealed that the C-terminal part of the MCM protein is more conserved than the N-terminal portion. The C-terminal part of the molecule contains the AAA+ catalytic domains, known to be highly conserved among different members of this family of enzymes. Nevertheless, the *Halobacterium* MCM clearly shows amino acid sequence similarity to other member of this family of helicases.

The sequence of the *Halobacterium* MCM also revealed that it contains a relatively small number of basic residues in comparison to other MCM proteins, and the ratio between negative and positive residues is much higher than in MCM proteins from non-halophilic archaea (Table 1). Similar observations of charge distribution on proteins have also been observed when the entire proteome of *Halobacterium* was analyzed (139).

6.4.2 Purification and identification of *Halobacterium* MCM helicase

The MCM protein containing the intein was cloned into *E. coli* expression vector and purified as described in “Material and Methods”. Although most of the proteins were

not soluble, a small fraction was soluble and this fraction was purified (Fig. 6-2 lane 4). However, the intein is still present in the protein. It is known that in other intein-containing proteins expressed in *E. coli*, the intein is excised during induction prior to purification. The observation that the intein does not excise itself but stays as a part of the purified MCM suggests that either the *in vivo* conditions in *E. coli* (e.g. salt) are not favorable for the intein protease activity or that the protein is not properly folded.

In order to get a MCM protein without the intein, the intein was recombinantly removed from the gene and the expressed protein was purified as described in “Material and Methods” (Fig. 6-2 lane 7).

When analyzed on SDS-PAGE, both proteins migrate more slowly than predicted from their calculated molecular mass. The protein with intein migrates as a 123 kDa instead of 90 kDa molecule and the recombinant protein without intein migrates as 98 kDa instead of 71 kDa molecule. To confirm that the purified proteins are indeed the MCM proteins MALDI-TOF MS analysis was performed. The proteins containing intein (Fig. 6-3A) and without intein (Fig. 6-3B) were analyzed. Both proteins contain peptides corresponding to the *Halobacterium* MCM proteins. The peaks at m/z position 2158.16 and 2907.48 (Fig. 6-3 A) are present only in the MCM protein containing the intein; they are absent in Figure 6-3 B. This suggests that these peaks originate from the MCM intein. A MS/MS analysis of the 2158.16 peak revealed that it is, indeed, the peptide from the intein sequence (Fig. 6-3C and E). Another peak, 2333.20, is found in both proteins (Fig. 6-3A and B). A MS/MS analysis of the peak showed that it is derived from a peptide located in the MCM protein but outside the intein (Fig. 6-3D and E). The combined results of the MS and MS/MS analysis recovered 43% and 50% of MCM

protein with and without intein, respectively. This suggests, with high confidence, that the two recombinant proteins isolated are the *Halobacterium* MCM protein.

6.4.3 Detection of *Halobacterium* MCM protein *in vivo*

As described above (Fig. 6-1), the *Halobacterium* MCM gene encodes for a protein which contains an intein. As the intein is located in a highly conserved and important domain for helicase activity, it is presumed that, *in vivo*, the intein excises itself from the protein after translation and ligates the two other peptide fragments (exteins). Thus, the mature MCM protein will not contain the intein and will therefore be active (140). It was therefore expected that in *Halobacterium* cells only the mature protein will be present. To determine if this is the case a western analysis was performed using cell extract and anti-MCM antibodies. As shown in Figure 4 lanes 1 and 2, only the mature MCM, without the intein, could be detected.

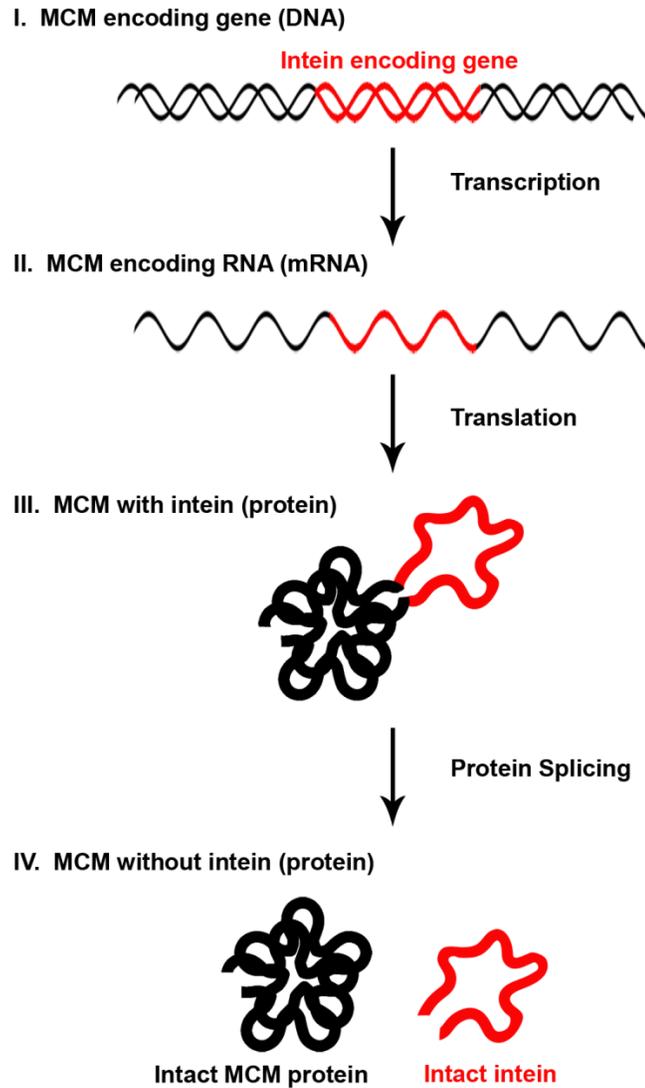


Figure 6-1 General scheme of how intein is expressed and excised from the *Halobacterium* MCM helicase.

Intein is integrated into a MCM gene, transcribed into mRNA, and translated into a protein. Protein splicing is the step that intein cleaves itself from the rest of the protein, ligating the exteins thus leaving the rest of the protein intact.

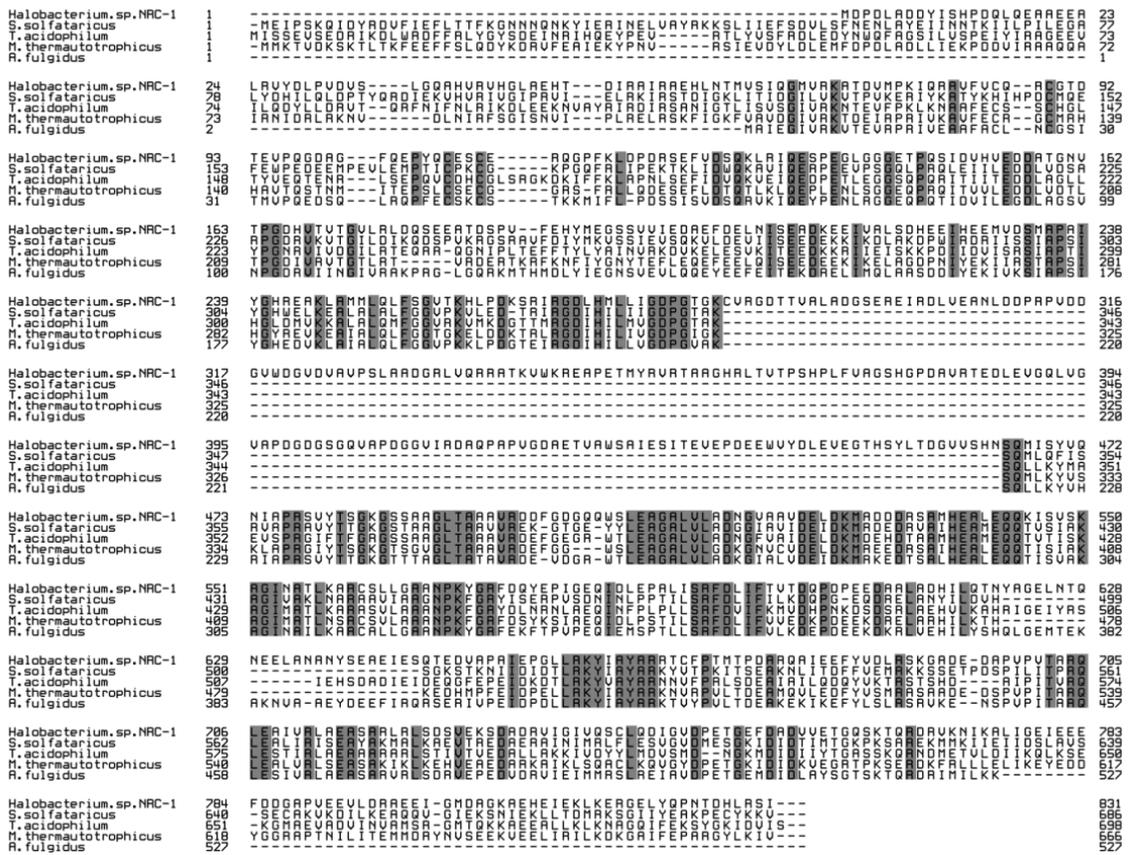


Figure 6-2 Alignment of full length *Halobacterium* and other archaeal MCM proteins. Amino acid sequence alignment of five archaeal MCM helicases: *Halobacterium sp. NRC-1*, *M. thermotrophicus*, *T. acidophilum*, *M. thermophila* and *A. fulgidus*. Highlighted residues are conserved in all five proteins.

Table 6-1 Amino acid analysis of MCM proteins.

	Length	KDa	Negatively charged: E + D (%)	Positively charged: K + R (%)	Ratio: (E+D)/(K+R)	pI
<i>Halobacterium sp. NRC-1 with intein</i>	831	90.5	165 (19.9)	74 (8.9)	2.2	4.4
<i>Halobacterium sp. NRC-1 without intein</i>	649	71.2	130 (20.0)	59 (9.1)	2.2	4.4
<i>M. thermautotrophicus</i>	666	75.6	121 (18.2)	100 (15.0)	1.2	5.2
<i>T. acidophilum</i>	698	78.9	109 (15.6)	94 (13.5)	1.2	5.5
<i>S. solfataricus</i>	686	77.4	107 (15.6)	101 (14.7)	1.3	6.0
<i>A. fulgidus</i>	698	78.8	116 (16.6)	102 (14.6)	1.1	5.7

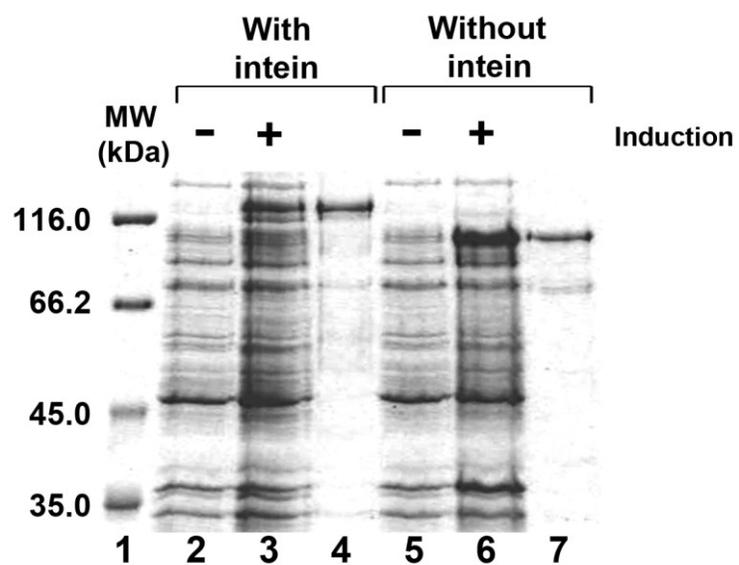


Figure 6-3 The purification of the *Halobacterium* MCM helicase.

Uninduced and induced *E. coli* cells harboring pET-MCM (lanes 2-3) and pET-MCMint (lanes 5 and 6) and the purified MCM protein (0.5 μ g) with intein (lane 4) and without intein (lane 7) were fractionated on 10% SDS-PAGE. Lane 1, molecular weight markers.

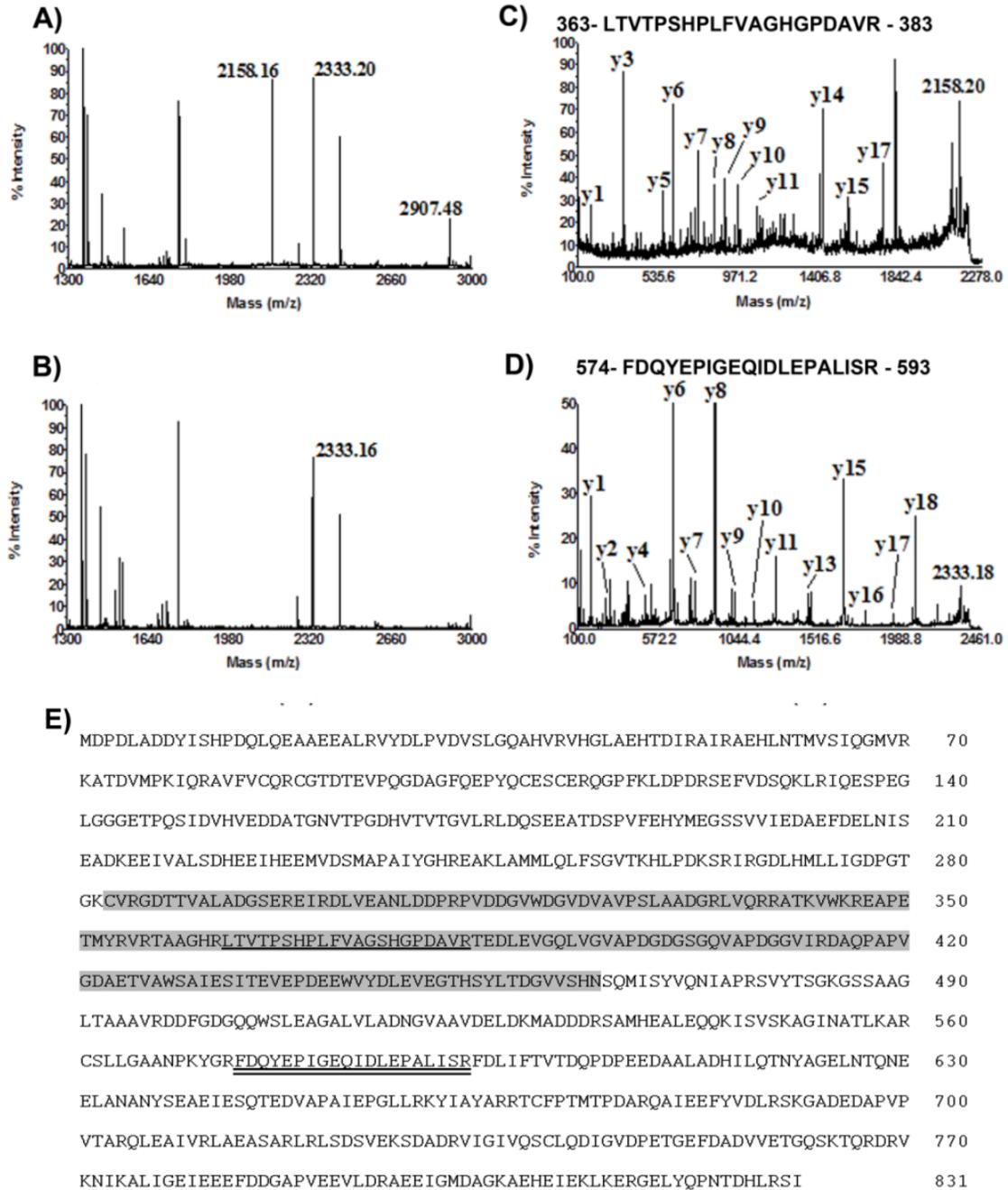


Figure 6-4 The *E. coli* expressed proteins are the *Halobacterium* MCM proteins.

MALDI-TOF spectra of *Halobacterium* MCM with (A) and without (C) intein were performed as described under “Material and Methods”. The 2158.16 and 2333.16 peaks were analyzed using MS/MS (B and C, respectively) which was performed as described

in Material and Methods. Panel C, 2158.16 peak. Panel D, 2333.16 peak. Panel E shows the position of the peptide in the MCM amino acid sequence. Shaded areas show the location of the intein. The peptide corresponding to the 2158.16 peak is underlined and that for the 2333.16 peak is double underlined. Experiment was performed by Dr. Illarion Turko.

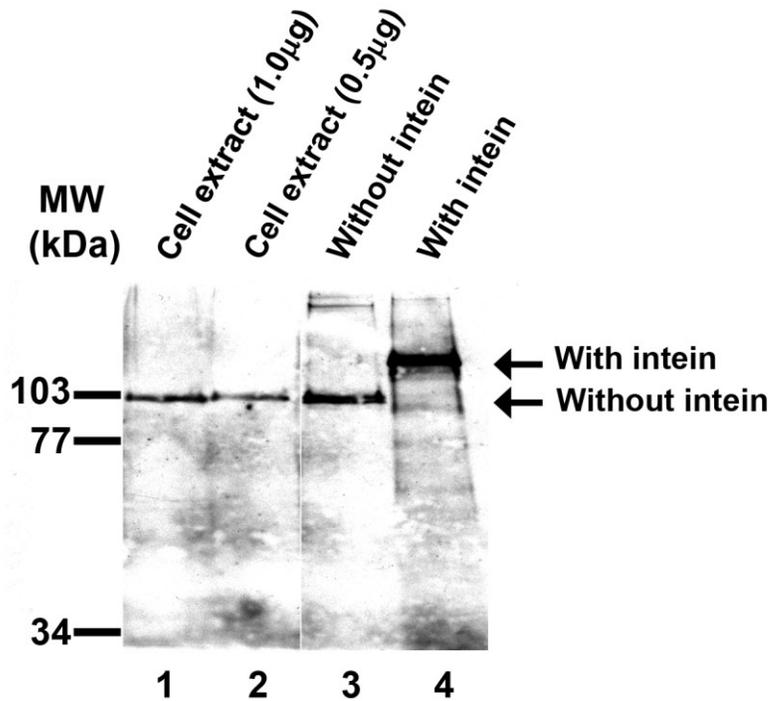


Figure 6-5 Only the MCM protein without intein can be detected in *Halobacterium* cell.

Western analysis of *Halobacterium* cell extract and recombinant MCM protein was performed as described in “Material and Methods”. Lane 1, cell extract (1 μg); lane 2, cell extract (0.5 μg); lane 3, purified MCM protein without intein (1 ng); lane 4, purified full length MCM (20 ng). To the left are the positions of the molecular weight markers. Experiments were performed by Claire Rollor.

Chapter 7 How is the archaeal MCM helicase loaded at the origin? Possible mechanisms.

7.1 Abstract

In order for any organism to replicate its DNA a helicase must unwind the duplex DNA in front of the replication fork. In archaea the replicative helicase is the minichromosome maintenance (MCM) helicase. While much is known about the biochemical properties of the MCM helicase, the mechanism of assembly at the origin of replication is unknown. Here, several possible mechanisms for the loading process will be described.

7.2 Introduction

The minichromosome maintenance (MCM) complex is the replicative helicase in archaea. Several archaeal MCM helicases have been studied, and their biochemical properties described. The helicases from most archaea form hexamers in solution, have an ATP-dependent 3'→5' helicase activity, bind and translocate along single-stranded (ss) and double-stranded (ds) DNA, can unwind DNA-RNA hybrids while moving along the DNA strand, and can displace proteins from DNA (reviewed in (50,115,116,128)).

7.3 MCM structure

High-resolution structural information is not yet available for a full-length MCM helicase, but has been determined for the N-terminal, non-catalytic domains of two MCM proteins, from *Methanothermobacter thermautotrophicus* (59) and *Sulfolobus solfataricus* (60). These have a three-domain structure. Biochemical studies suggest that

domain A plays a role in regulation (81,102,110), and domain B participates in DNA binding via a zinc-finger motif (61). Domain C, which connects the N-terminal portion of the protein to the C-terminal catalytic region, is involved in protein multimerization (60,81), involved in DNA binding via a β -hairpin motif (59,90,117) and plays a role in the communication of signals between the N-terminal and C-terminal portions of the MCM protein ((103), see Chapter 2). Only low-resolution electron micrograph reconstruction studies have been performed on the full-length MCM protein from the archaeon *M. thermautotrophicus*. These studies suggest that the protein can adopt multiple forms; hexamers, heptamers, dodecamers filaments and open circles were identified ((86) and references therein). Conformational changes upon DNA and nucleotide binding have also been observed (87).

7.4 Helicase loaders

Replicative helicases in bacteria and viruses require a loader protein or complex to assemble the helicase onto the origin of replication. In *Escherichia coli* the DnaC protein functions as the helicase loader for the DnaB helicase. It is currently unknown whether MCM helicases have loader proteins. Since the archaeal replication machinery possesses some of the characteristics of bacteria it is possible that an archaeal loader exists. One candidate protein for the archaeal helicase loader is the Cdc6 protein. This is based on primary amino acid sequence similarity to the eukaryotic initiator protein Cdc6. In addition, the enzymes were shown to interact with MCM (35), regulate MCM helicase activity (75,104,105) and affect the oligomeric state of the helicase (114).

7.5 Formation of the replication bubble

Replication begins at origins of replication, which are A/T rich and contain one or more A/T-rich stretches, known as duplex unwinding elements (DUE), which are essential for origin function. Origin binding proteins (OBP) bind origin recognition boxes (ORB) and cause localized unwinding. All archaeal origins identified contain multiple short inverted repeats, and some also contain two long inverted repeats flanking the origin region. The short inverted repeats in different archaeal species share sequence similarity ((31) and references therein) and serve as ORBs to which the OBPs bind. It was shown that when the archaeal Cdc6 protein (OBP) binds to the inverted repeat at the origin it causes a distortion of the duplex DNA (36,38).

There are at least two mechanisms by which the replication bubble can form (Figure 7-1). Multiple OBPs binding to multiple ORBs might result in sufficient deformation to result in bubble formation (Figure 7-1 A). Support for this hypothesis comes from *E. coli*, where the DnaA protein (the OBP) binds to the DnaA-boxes (ORB) and results in replication bubble formation. Alternatively, under supercoiling conditions, such as those found *in vivo*, the inverted repeat sequences may form cruciform structures with stems formed from the inverted repeats and single-stranded loops composed of intervening sequences (50). The binding of the OBP could stabilize the stems (Figure 7-1 B).

Recent data demonstrated that DNA could interact with the outer surface of the MCM helicase, bending the DNA (107). Thus it was suggested that the helicase itself could aid in bubble formation (Figure 7-1 C) (108).

7.6 Possible mechanisms of archaeal MCM helicase loading at the origin

Helicase assembly in archaea may or may not use a loader protein; investigation into this is ongoing. Several possibilities for helicase loading are presented here.

7.6.1 Helicase loader-dependent assembly

7.6.1.1 Ring breaker

If archaea employ a loader protein, as do bacteria, helicase assembly at the origin could occur as it does in bacteria. Much is known about the mechanism of helicase loading at the *E. coli* origin of replication (*oriC*). The DnaC protein functions as a helicase loader, assembling the DnaB helicase at the origin in an ATP-dependent reaction. The protein forms a complex with the hexameric ring of DnaB resulting in a DnaB₆:DnaC₆ complex. The complex then interacts with origin-bound DnaA. When the DnaA protein binds to the origin it causes a distortion in the DNA and forms the replication bubble. After binding the DnaA-DNA complex the DnaC protein opens the DnaB hexameric ring and assembles it onto the replication bubble. Thus, the DnaC protein functions as a “ring breaker” (Figure 7-2 A) (141).

Biochemical studies with a number of archaeal Cdc6 proteins show that they have biochemical properties that are similar to those of the *E. coli* DnaC loader. Both proteins interact with the helicase and inhibit helicase activity *in vitro* and both bind and hydrolyze ATP. These similarities led to the hypothesis that the archaeal Cdc6 may be the functional homologue of the *E. coli* DnaC protein and function as a “ring breaker” (Figure 7-2) (115).

7.6.1.2 Ring dissociater

The archaeal replication machinery may be a modified version of the *E. coli* ring breaker model that involves dissociation of the helicase (Figure 7-2 B). Studies with the Cdc6 and MCM proteins from the archaeon *M. thermautotrophicus* showed that interaction between the Cdc6 and MCM proteins dissociates the helicase, and no large complexes, which may include hexamers of MCM that associate with Cdc6 protein, could be observed (114). These observations are different from those made with the *E. coli* proteins in which a complex between hexamers of the DnaC and DnaB proteins can easily be detected. Thus, it was suggested that the Cdc6 protein may still have a MCM loading function but instead of opening the ring at one interface, within the hexameric structure, it dissociates the MCM ring prior to assembly. This would make Cdc6 function as a “ring dissociater” (114) (Figure 7-2 B).

7.6.1.3 Ring maker

A potential mechanism could be to assemble monomeric subunits onto the origin to form the active helicase, referred to as a “ring maker” (141). However, as all archaeal MCM proteins are stable hexamers or other multimers, this mechanism is unlikely.

7.6.2 Helicase loader-independent assembly

If the initial replication bubble (Figure 7-1) is sufficiently large, the helicase may be able to assemble itself on it without the need for a loader (Figure 7-3). There are several observations that may support this hypothesis. The archaeal MCM complex was shown to assemble around circular single-stranded closed plasmid DNA ((50) and references therein). The *E. coli* DnaB helicase, on the other hand, cannot assemble

around such a substrate (and therefore requires a helicase loader). Also, the size of the single-stranded loop required for MCM assembly is about 50 bases. This size loop may be available between the long inverted repeats in several archaeal origins (for example, in *Pyrococcus abyssi* the distance is 240 bases (142) and in *S. solfataricus* the distance is about 65 bases (26)). Therefore bubble formation of sufficient size could occur via the mechanisms shown in Figure 7-1.

7.6.2.1 The MCM helicase may open to load onto DNA

The hexameric ring can assemble itself through complex “breathing” resulting in temporary opening of the ring allowing assembly onto the DNA (Figure 7-3 A). The interaction with DNA may stabilize the ring and prevent dissociation.

7.6.2.2 The MCM helicase may thread itself onto the initial replication bubble

However, even if the initial bubble is too small for hexamer assembly the MCM might still thread itself onto it. The electron micrograph reconstruction studies with the *M. thermautotrophicus* MCM complex revealed that the protein can form open circles and filaments (85,86). Both of these helix-like structures have an opening in the ring. Therefore, it was suggested this could facilitate self-loading of the helicase (86). This loading could take place even on a small replication bubble created by the OBPs if the helical MCM protein can wrap itself around the DNA strand. The open ring could act as a “corkscrew” which would melt the DNA and form the bubble as it wound onto the DNA. Structural studies show that when Cdc6 proteins bind the origin they form a distortion in the duplex (36,38). Although there are multiple ORBs in most of the archaeal origins the distortion formed by them may not result in a bubble sufficiently large for helicase assembly; only a small ssDNA region may be exposed. If the MCM

complex could melt the DNA using a corkscrew motion while it threads on, the small bubble formed may be sufficient for helicase self assembly (Figure 7-3 B).

7.6.2.3 The oligomeric state of the MCM helicase may change during loading

Electron micrograph reconstruction studies with the *M. thermautotrophicus* MCM have shown that the enzyme can form hexamers and heptamers (83,86,88). It was suggested that DNA binding by the helicase might result in a switch from heptamer to hexamer and facilitate the assembly of the active helicase (88). It is also possible that during assembly at the origin, the helicase loader removes one of the protomers from the heptamer and closes the hexamer around the DNA (Figure 7-3 C) (83).

7.7 What is the role of duplex DNA translocation by the MCM helicase?

All MCM proteins studied bind and translocate along duplex DNA. It is not yet clear if during replication the helicase moves on ss or dsDNA and models for both possibilities have been proposed (143). It has also not been established whether the helicase assembles at the origin prior to origin melting by the OBPs, and thus may be loaded on dsDNA, or after origin melting, and thus on either ss or dsDNA.

It is possible, however, that duplex translocation plays an indirect role in helicase loading by speeding the process by which the helicase locates the origin. Most of the genes encoding archaeal OBPs are located in close proximity to the origin of replication. It was suggested that this proximity enables the proteins to bind to the origin as soon as they are synthesized (50). However, the genes encoding the MCM proteins are located in different locations along the chromosome, and not always near the origin. The helicase

therefore has to travel to find the origin, either by itself or in combination with the helicase loader. As diffusion in one dimension is more efficient than diffusion in three dimensions, one possibility is that the helicase binds to the duplex DNA of the chromosome and translocates along the duplex until it reaches the origin. At the origin the helicase may either encounter the replication bubble and self load, as described above, or the helicase may encounter the helicase loader associated with the origin.

Several biochemical properties of MCM support this hypothesis. The MCM proteins from several archaea can associate with circular dsDNA and translocate along the duplex (44,72). It was also shown that the MCM protein can displace histone and other proteins from DNA, suggesting translocation can occur *in vivo* (93).

7.8 Do all archaea utilize the same loading mechanism?

As described above, the mechanism of helicase assembly at the origin of replication is not yet known, although several mechanisms have been proposed. It is also not clear whether all archaeal species use the same loading mechanism or whether each organism, or group, developed their own assembly process. Several lines of evidence suggest that more than one loading mechanism may exist. 1) Not all archaeal species contain a clear homologue of Cdc6, and some species contain only one homologue. In species lacking a Cdc6 homologue the loader may not exist (loader-independent assembly) or may not have been identified. In some species the Cdc6 protein may function both in origin recognition and as the helicase loader. 2) Not all archaeal MCM complexes behave the same. Different forms in solution have been observed (for examples see (72,89)) and different activities have been noted in different species. For example, some enzymes are active on their own but some require additional factors such

as Cdc6 (75) or the GINS complex (78). 3) Some archaea contain multiple MCM homologues. If these form hetero-oligomers the loading mechanism may be different from that of a homomultimer. 4) The relative abundance of MCM helicase within the cell appears to differ between different organisms. Studies have shown that *P. abyssi* contains about 400 molecules of MCM in rapidly dividing cells (29), *M. thermautotrophicus* has about 300 molecules, but *Thermoplasma acidophilum* contains about 1000 molecules per cell. The abundance of the MCM and other initiation proteins (e.g. Cdc6) may influence helicase assembly. 5) The Cdc6 proteins from different archaea have different effects on helicase activity *in vitro* (for examples see (75,104)). These differences could possibly be a reflection of different loading mechanisms.

Support for different loading processes in different archaea comes from studies with bacteria. It was shown that while the *E. coli* helicase is loaded at the origin via a “ring breaking” process the helicase from *Bacillus subtilis* is loaded via a “ring making” mechanism (141). It is therefore possible that different archaea may also use different loading mechanisms.

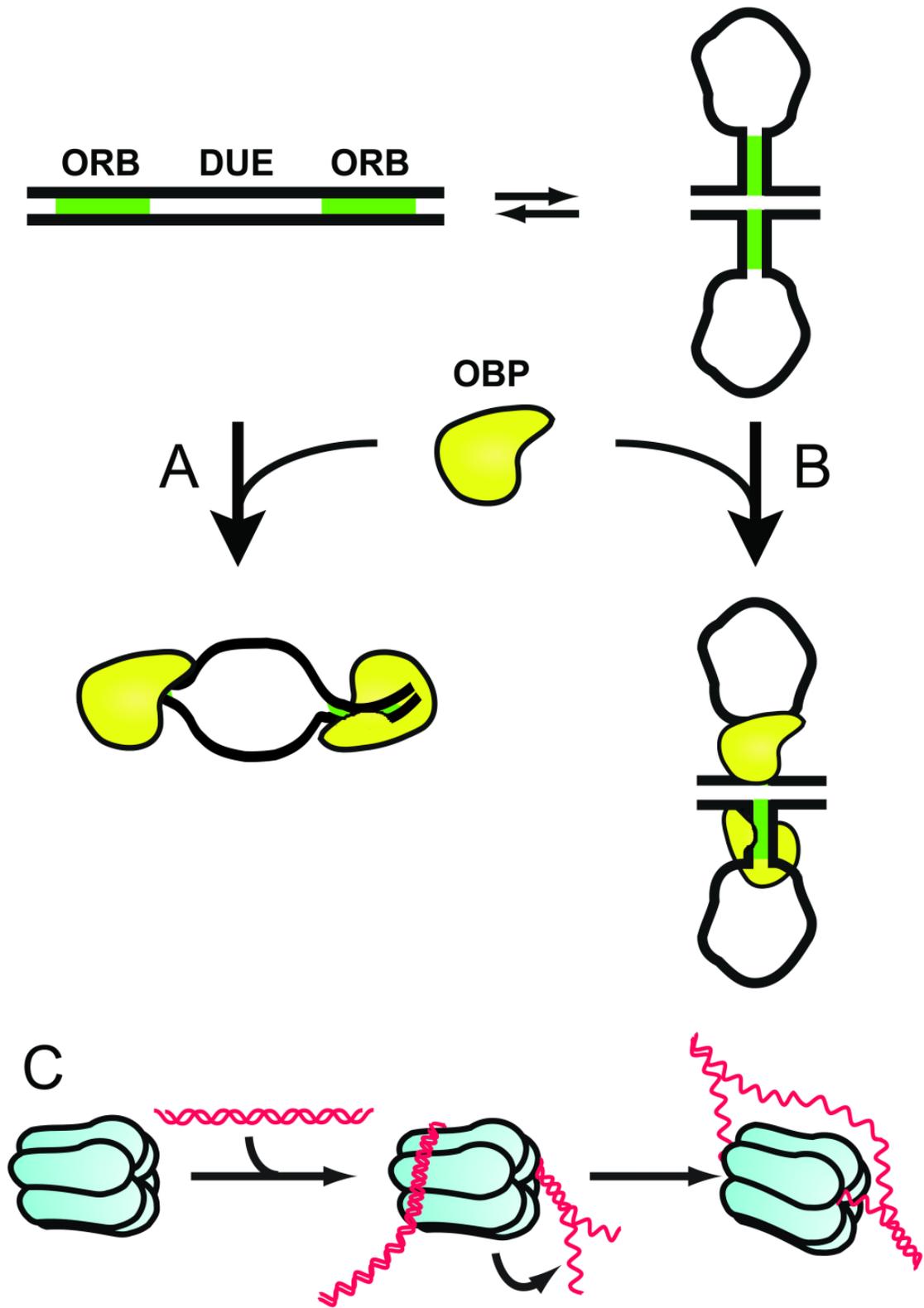


Figure 7-1 Possible mechanisms of replication bubble formation.

The duplex unwinding element (DUE) is flanked by origin recognition box (ORB, green). Although most archaeal origins contain multiple ORBs only two are shown for simplicity. The origin binding protein (OBP) is shown in yellow. A. Bubble formation may be the result of cumulative binding of OBPs to multiple ORBs at the origin, resulting in deformation of the DUE. B. Bubble formation may result from cruciform formation of the origin. The ORBs could form a stem, which could be stabilized by OBP binding. See text for details. C. MCM interacts with duplex DNA with its surface and bends the duplex DNA distorting DUE. Once ss DNA is available, it is directly loaded.

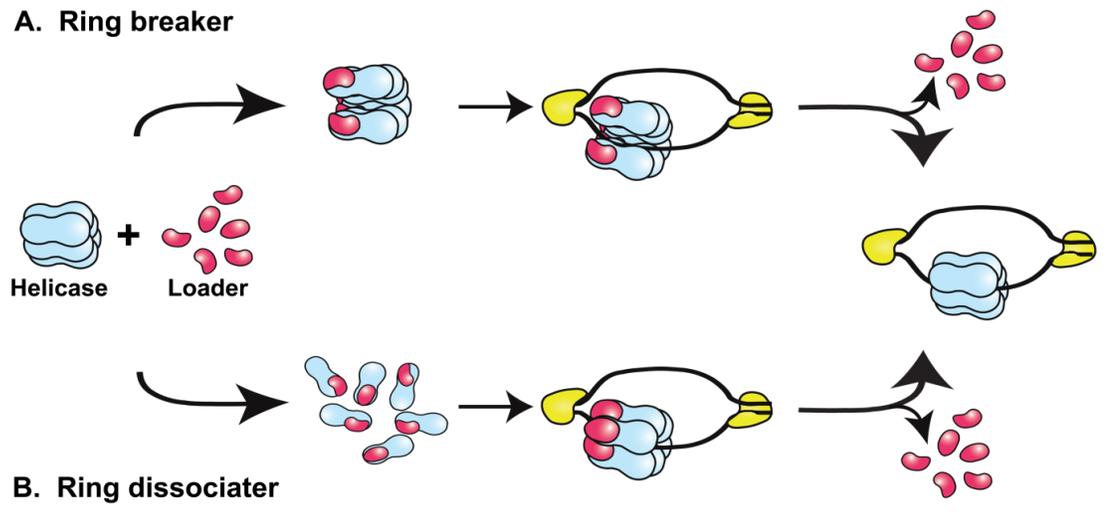


Figure 7-2 Possible mechanisms of loader-dependent helicase assembly.

The loader (pink) associates with the helicase (blue) and facilitates loading onto the replication bubble. The OBP is shown in yellow. A. “Ring breaker” model. B. “Ring dissociater” model. See text for details.

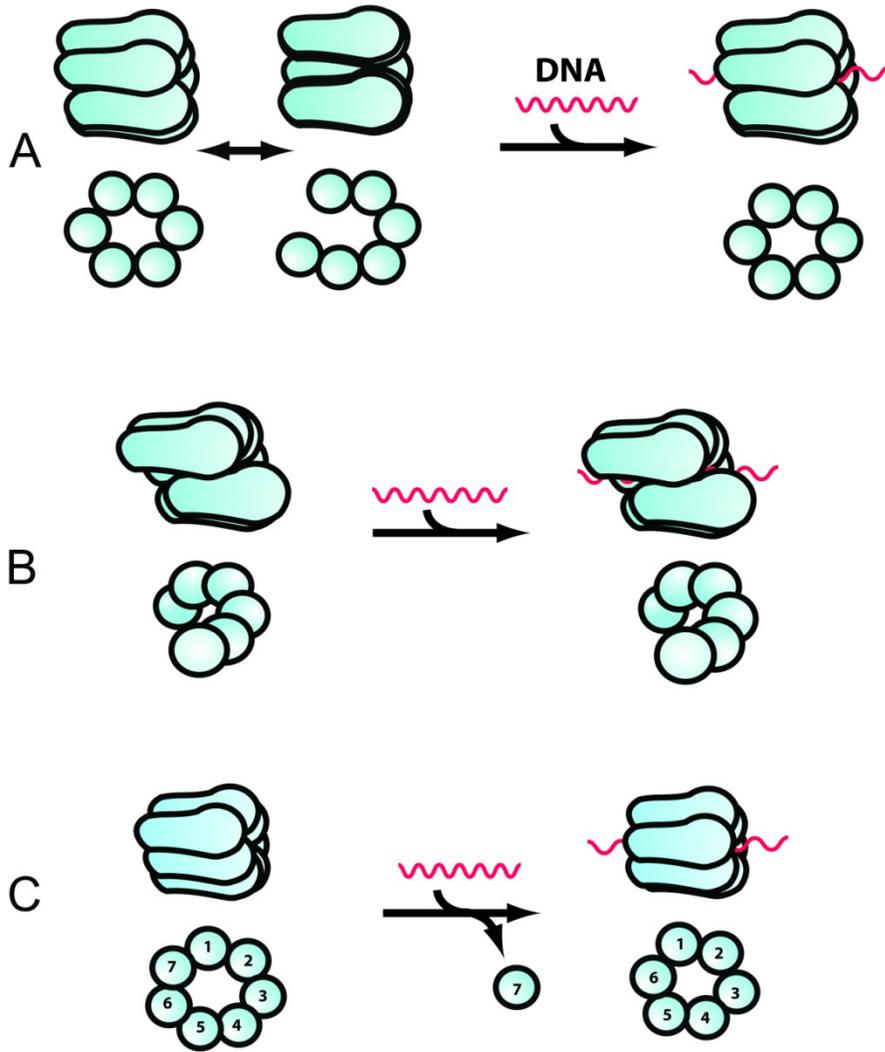


Figure 7-3 Helicase loader-independent assembly.

The helicase (blue) is shown in top- and side-views. DNA is shown in red. A. The MCM complex may open to load onto DNA. B. The MCM helicase may thread itself onto the initial replication bubble. C. The oligomeric state of the MCM helicase may change during loading. See text for details.

Chapter 8 Concluding Remarks

MCM helicase is an essential enzyme in both the initiation and elongation stages of DNA replication. Its function is to translocate on one strand of the duplex DNA, and displace the complementary strand providing the ssDNA template for the DNA polymerase. A part of this dissertation elucidates the possible roles of the conserved residues found in the N-terminal region of MthMCM using various biochemical assays.

The conserved loop between $\beta 7$ and $\beta 8$ was shown to be essential for the communication between the catalytic domain and the N-terminal domain (see Chapter 2). According to the recent crystal structure of SsoMCM full length protein, the loop extends to the AAA+ domain of the neighboring subunit. Although the hexameric form of SsoMCM is a modeled structure, the loop is clearly in the close proximity to the β -hairpin (PS1BH) of the AAA+ domain. This PS1BH is conserved among all MCM helicases, and is one of the key motifs required for helicase activity. Thus having the proximity to such an essential motif, it is possible that the loop may be directly involved in the helicase activity. It is also possible that the loop and PS1BH are interacting during unwinding. It would be interesting to see whether the compensating mutations at PS1BH can rescue the mutation at the loop to determine whether the interactions between the loop and the PS1BH are required for the helicase activity. It is also possible to perform a complementation experiment with the PS1BH counter mutant and loop mutant to see if they can rescue the activity *in vitro*. If these mutants complement the activity, the loop is likely to be involved in helicase activity *in trans*. It is possible that the helicase cooperative activity is not only contributed from AAA+ domain, but also from the interactions between the neighboring N-terminal domain and the catalytic region.

Although genetic tools have not yet been developed for *M. thermautotrophicus*, such tools have been developed for several archaeons. It will be interesting to take advantage of the genetic system to see whether the loop is also essential *in vivo*. And if it is essential for cell survival, is it possible to rescue the cell with counter mutation at PS1BH at the AAA+ domain?

The EM studies have demonstrated that the MCM helicase conformation is altered in the presence of DNA. A part of this dissertation focuses on the effect of nucleotide and DNA on the thermostability of the MCM helicase (Chapter 5). The study suggests that ATP hydrolysis and DNA change the conformation of MCM. It is possible that the ATP hydrolysis transition state is stabilized in the presence of DNA and thus observed increase in ATPase activity in the presence of DNA. However, one needs to explore the kinetics of each step of the hydrolysis step in the absence and presence of DNA to determine whether it is the case. Dissecting the steps of hydrolysis would allow us to understand the mechanism of the MthMCM helicase. Many questions need to be addressed. For example, it is not clear how many nucleotides the MCM helicase unwinds or translocates per hydrolysis step. The mechanism of hydrolysis, whether it is sequential or concerted remains unclear. The stoichiometry of MCM-ATP or DNA remains undetermined.

Some archaeal MCM helicases have ATPase activity that is not stimulated in the presence of DNA. Some archaeal MCM helicases require GINS or Cdc6 proteins to enhance the ATPase and helicase activity. If to make an analogy between the observations from MthMCM, it is possible that, instead of DNA, GINS or Cdc6 proteins

may stabilize the hydrolysis transition state of MCM helicases, and thus increasing the turnover activity of the enzyme.

It would be also interesting to see whether the ATP and DNA have similar effect on mesophilic MCM helicases. The ATP stimulates the DNA binding of MthMCM as it was shown in the Chapters 3 and 4. One interesting observation is that one of the mutants which lost DNA binding completely was able to regain its ability to bind to DNA in the presence of ATP (Chapter 3). This mutant protein showed lower thermostability, about 3°C lower than biochemical assay conditions. It is possible that the DNA and ATP form favorable complex that is thermostable and thus active in the presence of ATP. To date, no *in vitro* activity of mesophilic archaeal MCM helicase has been reported. Understanding the activity of MCM helicase from diverse groups would provide more insights into the MCM helicase functions.

While much progress has been made to understand the function of archaeal MCM helicase, much remain to be explored. The crystal structure of SsoMCM monomeric form was recently resolved and would help significantly in the future studies, yet the determined structure reveal only the N-terminal and AAA+ domains (without C-terminal domain) with a resolution of 4.1 Å. It will be important to improve the resolution of the structures and to determine the full length structure of a hexameric complex. Moreover, it would be of great interest in our field to determine the structure of MCM with substrate and nucleotide. It is also important to address whether MCM helicases from several archaea have similar fold overall. The structures of N-terminal domains of MthMCM and SsoMCM show almost identical fold. However, the role of N-terminal in both proteins seems to be slightly different. For instance, ATPase is coupled to DNA binding

in MthMCM while in SsoMCM, ATPase activity is DNA independent and AAA+ domain alone has helicase and ATPase activity. Would the structures of both MCM helicases be expected to be the same? Structural differences may provide insights into the differences in the ATPase and helicase activity between different archaea.

In most of archaeal MCM helicases studied to date, Y-shaped, forked substrate is preferred compared to 3' overhang substrate. As one may expect, the MCM helicase being a replicative helicase may favor substrate that resembles the replication fork. It is likely that as MCM helicase moves along the axis of the one strand of DNA in 3' to 5' direction, the other strand is excluded to the exterior surface of the molecule (also known as the steric exclusion mechanism). The structure of SsoMCM revealed that there is an external β -hairpin (EXT- β hp) in AAA+ domain that is essential for the helicase activity. Interestingly, SsoMCM AAA+ domain alone can unwind various substrate including blunt end duplex, and 5' overhang and 3' overhang, suggesting that 5' tail is not required for the helicase activity of AAA+ domain. Only in the presence of N-terminal domain, the MCM helicase discriminates and shows preferences to specific substrates. The C-terminal end of the molecule was shown to enter the 3' end of DNA first. If it is indeed the case, how does the N-terminal domain discriminate one substrate from the other? What determines the directionality of the enzyme? Do all archaea enter the 3' end from C-terminal end? Do MCM proteins directly assemble at the replication fork, or they translocate on the ss DNA until it finds a complementary strand which stands on its way? What factors regulate MCM from translocating on ss or dsDNA? It has been shown that the MCM helicase is expressed during exponential and stationary phase in *P. abyssi* and *S. acidocaldarius*. While it is bound to the origin in the early S-phase of cell cycle, it

moves away from the origin as the MCM is likely to move with the replication fork. However what regulates association and dissociation of MCM from chromatin in the cell remains unclear. One of the main questions that remains in this field is how, by what mechanism the helicase is assembled at the origin. Although some studies provide clues to the process (Chapter 7) the development of an *in vitro* loading assay will be of great importance.

To date, all archaeal MCM studies are from organism with a single homologue. As it has become apparent that many organisms contain multiple MCM homologues, it would be important to find whether these paralogues form hetero multimeric complex. If they do form a heteromeric complex, it would be interesting to see if all of the MCM in the cell are essential for the cell viability.

With the recently developed genetic tools in several archaeal species some important questions can be addressed. Especially, it would be essential to determine the interacting partners of the MCM helicase and how they may regulate the MCM function during the cell cycle.

Chapter 9 Appendices

9.1 Protein expression and purification

9.1.1 *M. thermautotrophicus* MCM protein

Plasmid containing the MCM gene (pET21a – MCM with C-terminus His₁₀ tag) was transformed into BL21-Codon Plus (DE3)-IRL cells (Stratagene). Single colony was inoculated into 50 ml of minimal medium containing 1 mM MgSO₄, 0.5 % (or 5 g/L) glucose and 1 x NPS [13.5 g/L (NH₄)₂SO₄, 28 g/L KH₂PO₄, 29 g/L Na₂HPO₄] and ampicillin and chloramphenicol. The medium was shaken overnight at 37°C at 250 rpm.

The overnight culture was transferred to a 1 L autoinduction medium, containing, ZY medium (10 g/L N-Z amine AS and 5 g/L yeast extract), 1 x NPS and medium 5052 (5.0 g/L Glycerol, 5.0 g/L glucose, and 2 g/L α-lactose), where ZY medium, NPS, and 5052 were separately autoclaved, and only added before inoculation. The medium was shaken at 37°C for 24 hours at 250 rpm, followed by cell harvesting. Cells were kept at -80°C.

The cell pellet (~13 g/L) was thawed and resuspended in 50 ml lysis buffer containing 50 mM Tris-HCl (pH 8.0), 0.5M NaCl, and 20% glycerol at 4°C. The cells were then lysed by sonication, after which the cell debris was removed by centrifugation at 15,000 rpm in JA-17 rotor (Beckman) at 4°C. The lysate was bound to 5 ml of Ni NTA resin (Qiagen) with gentle shaking for 1 hour at 4°C. Following binding, the mixture containing the Ni NTA resin was poured into a column and washed with 50 ml lysis buffer containing 10 mM imidazole. The column was moved to 22°C, and all subsequent elution steps were carried out at that temperature. The column was further washed with 3

column volumes (CV) of 50 mM imidazole in elution buffer containing 40 mM Tris-OAc (pH 8.0), 0.4 M potassium acetate, and 20 % glycerol. The proteins were then step eluted in 3 CV fractions with increasing concentrations of imidazole (100, 150, 200, 250 and 300 mM) in the same elution buffer. The 1 µg of protein samples from each step was loaded onto 10 % SDS-PAGE and visualized. The peak fractions containing the purified protein samples were pooled for further purification using ion exchange chromatography using Q-Sepharose column.

The pooled proteins were diluted or dialyzed such that the final concentration of buffer (loading buffer) is 20 mM Tris-HCl (pH 8.0), 100 mM NaCl and 20 % glycerol. The proteins were then loaded onto Q Sepharose column pre-equilibrated with the same loading buffer. After protein binding, the column was washed with 3 CV of 20 mM Tris-HCl (pH 8.0), 300 mM NaCl and 20% glycerol. The protein was then eluted using 3 CV of 20 mM Tris-HCl (pH 8.0), 600 mM NaCl and 20 % glycerol. The eluted proteins were dialyzed 2 times against 20 mM Tris-HCl (pH 8.0), 100 mM NaCl and 20% glycerol to bring down the salt concentration of the buffer. Finally, protein concentration was measured by Bradford (Bio-Rad) using BSA as the standard and the samples were flash-frozen in liquid nitrogen and stored at -80°C.

9.1.2 *Halobacterium sp.* NRC-1 MCM protein

Plasmids containing the genes encoding for the full-length MCM and the protein without intein were transformed into BL21-Codon Plus (DE3)-IRL cells (Stratagene). Cells were grown in LB media at 37°C. When the OD₆₀₀ reach 0.6, proteins expression was induced by the addition of IPTG (1 mM final concentration) and the cells were grown for an additional 4 hours. Cells were harvested and the pellet was store at -80°C.

Protein purification was carried out at 4°C, unless otherwise stated. Cells were thawed on ice followed by addition of lysis buffer containing 10 mM imidazole, 3 M NaCl, 0.5 M KCl, 20 mM Tris (pH 7.6) and 20 % glycerol. Cells were incubated for 1 hour followed by sonication and centrifugation for 15 min at 15,000 rpm in JA-17 rotor (Beckman). The supernatant was incubated with Ni charged resin that was pre-equilibrated with lysis buffer. The cell pellet was saved for urea based protocol. The column was then washed with lysis buffer containing 10 mM imidazole. The MCM proteins were eluted with increasing concentration of imidazole (50 to 300 mM), similar to *M. thermotrophicus* MCM protein preparation. Since only a fraction of protein remained in soluble fraction, MCM protein was also purified in the presence of urea. For the urea based protein preparation, cell pellet which was collected after sonication stage was re-suspended in 8M urea containing buffer [8M urea, 20 mM Tris (pH 7.6)] with glass homogenizer at room temperature. The cells were centrifuged for 15 min at 15,000 rpm in JA-17 rotor, and supernatant was incubated with Ni-NTA resin. The column was washed several times with 10 mM Imidazole containing 8 M urea, and then eluted with 50, 100, and 300 mM imidazole containing 6 M urea. Fraction with highest homogeneity (300 mM imidazole eluent) was dialyzed step wise, from buffer containing 5 M urea up to 0M urea with 1M increments decreasing concentration of urea, and NaCl and KCl concentrations were increased step wise from 0 to 3 M NaCl, with initially increasing by 1 M increments and continued at 3 M NaCl, and 0 to 1.5 M KCl with initially increasing the concentration by 0.5 M increments and continued at 1.5 M KCl with constant concentration of 20 mM Tris (pH 7.6) and 20 % glycerol and product was flash frozen in liquid nitrogen.

9.2 Fluorescence polarization anisotropy (FPA) measurements

9.1.3 General principle

Fluorescence polarization anisotropy can be used to measure DNA-protein interaction in solution, in equilibrium condition. Fluorophores, when excited in polarized light, only molecules with particular orientation which is aligned with the polarized incident light can be excited, and emit the photon. This photon released by the fluorophore is also polarized, since dipole of emitted photons depend on the dipole of excited photons. Smaller the molecule that fluorophore is attached, faster it rotates, and it is less probable to align with the position of polarized light, which result in low anisotropy. When the small molecule is bound by larger molecule, the molecule tumbles slower, and has longer time of being aligned with the polarized light, which results in higher anisotropy value. More detailed background on the method can be found in (121).

In order to determine the equilibrium constant (K_d) for MCM binding to DNA, the 50mer oligonucleotide was labeled with Cy3 or ROX. DNA was titrated with MCM protein at room temperature after the reaction was equilibrated for 10 min. Detail is described below.

9.1.3.1 Calculation

9.1.3.1.1 Quadratic equation

The following assumptions have to be clarified before MCM-DNA binding is derived in terms of equilibrium equation; 1) The interaction is reversible, the association is bimolecular and dissociation is unimolecular. 2) All the DNA molecules are equivalent

and independent. 3) The signal response is proportional to the number of occupied DNA binding sites. 4) The interaction and response are measured after the reaction has reached the equilibrium. With these assumptions in mind, we can derive equations to determine the equilibrium constant for MCM – DNA binding;

Where MCM and DNA are interacting in solution;

Equation 9-1



At equilibrium, where the concentration of both product and the reactants are relatively constant;

Equation 9-2

$$\frac{[\text{MCM}_{\text{Free}}][\text{DNA}_{\text{Free}}]}{[\text{MCM : DNA}]} = \frac{[\text{M}_F][\text{D}_F]}{[\text{B}]} = \frac{k_{-1}}{k_1} = K_d$$

This equation can be re-written by substitutions; $\text{DNA}_{\text{Free}} = \text{DNA}_{\text{Total}} - \text{DNA}_{\text{Bound}}$; thus,

Equation 9-3

$$\frac{[\text{M}_F][\text{D}_T - \text{B}]}{[\text{B}]} = K_d$$

Moreover, we can write the equation in terms of fractional binding, in other words, ratio of bound versus free DNA, which gives us the information of where 50% of substrate is bound;

Equation 9-4

$$\frac{B}{D_T} = \frac{M_F}{(K_d + M_F)} = \text{fractional occupancy}$$

This can be further simplified by making $M_{\text{Free}} = M_{\text{Total}} - M_{\text{Bound}}$, and since $M_{\text{Bound}} = D_{\text{Bound}} = B$, simply.

Equation 9-5

$$B = \frac{D_T (M_T - B)}{K_d + (M_T - B)}; \text{ while}$$

Equation 9-6

$$\{K_d + (M_T - B)\}B = D_T (M_T - B);$$

Equation 9-7

$$K_d B + M_T B - B^2 = D_T M_T - D_T B;$$

Equation 9-8

$$0 = B^2 - (K_d + M_T + D_T)B + D_T M_T;$$

Solving for B would result in quadratic equation;

Equation 9-9

$$B = \frac{-(M_T + D_T + K_d) \pm \sqrt{(M_T + D_T + K_d)^2 - 4D_T M_T}}{2}$$

In order to convert this equation as signal in the anisotropy experiments, it is required to understand the relationship between the fraction bound to the anisotropy

values. Because the DNA molecule which is labeled with Cy3 can exist in two states, bound by MCM or free, anisotropy value of either high (for the bound state) and low (for the free state) can be treated as additives in the equation. Therefore, the observed anisotropy value, A, depends on the sum of fractional saturation of anisotropy of either bound or free;

Equation 9-10

$$A = F_f A_f + F_b A_b ;$$

Where $F_f + F_b = 1$, and A is the observed anisotropy value, F_f is the fraction of DNA that is free, F_b is the fraction of DNA that is bound, A_f is the anisotropy value of DNA that is free, and A_b is the anisotropy value of DNA that is bound by MCM. Since the fraction of DNA bound (F_b) is $DNA_{Bound} / DNA_{Total}$, and F_f is $1 - F_b$, the equation 7 can be written as follows;

Equation 9-11

$$[DNA_{Bound}] = D_T \frac{A - A_f}{A_b - A_f} = B$$

Combining the equations 6 and 8, since they represent DNA bound, following equation can be written for DNA binding for anisotropy experiments;

Equation 9-12

$$A = A_f + (A_b - A_f) \times \frac{(M_T + D_T + K_d) - \sqrt{(M_T + D_T + K_d)^2 - 4D_T M_T}}{2D_T} ;$$

or could be rewritten as

Equation 9-13

$$\frac{A - A_f}{A_b - A_f} = \frac{\Delta A}{\Delta A_{\text{Final-Initial}}} = \frac{(M_T + D_T + K_d) - \sqrt{(M_T + D_T + K_d)^2 - 4D_T M_T}}{2D_T}$$

Since we have the information on A, the observed anisotropy value, how much protein is added (M_T), and the amount of DNA (D_T), the K_d was determined by fitting the data to the **Equation 9-13**.

It is possible to simplify the quadratic equation. Let us assume that the depletion of free protein is very low (due to total DNA being much smaller than the K_d value), such that it is assumed to be the same as the total protein concentration, where $M_{\text{free}} = M_{\text{total}}$. The quadratic equation, derived in Equation 9-13 can be simplified to:

Equation 9-14

$$\Delta A = \left\{ \frac{\Delta A_T (M/K_d)}{1 + M/K_d} \right\}$$

In fact, when both **Equation 9-13** and **Equation 9-14** were used to compare the fit of a data set using GraFit program, the K_d values turned out to be the similar (see section on Data Analysis). The advantage of having this equation is that Hill coefficient can be easily included to find out whether there is a cooperative or multi-site binding (detail described below, and example is shown in Figure 9-3).

Equation 9-15

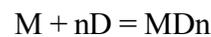
$$\Delta A = \left\{ \frac{\Delta A_T (M^n/K_d^n)}{1 + M^n/K_d^n} \right\}$$

9.1.3.1.2 Hill equation

In some proteins, especially in multimeric proteins, it is observed that binding of ligand to the particular site affect the affinity for the ligand in another unbound site, showing cooperativity of the molecule toward the ligand (144).

Consider the reaction:

Equation 9-16



Where M is MCM protein, and D is the DNA, and n is the Hill coefficient. K_a (association constant) can be expressed as:

Equation 9-17

$$K_a = \frac{[MD_n]}{[M][D]^n}$$

And consider that the fractional occupancy, which is the bound species over the total species, (Bound DNA/Total DNA = B/D_T from Equation 9-4) is expressed in terms of multiple binding sites.

Equation 9-18

$$\theta = \frac{[MD_n]}{[MD_n] + [M]} = \text{fractional occupancy}$$

Substituting **Equation 9-17** for **Equation 9-18** gives rise to **Equation 9-19**,

Equation 9-19

$$\theta = \frac{K_a[M][D]^n}{(K_a[M][D]^n) + [M]} = \text{fractional occupancy}$$

And mathematical rearrangement results in following;

Equation 9-20

$$\theta = \frac{[D]^n}{[D]^n + (1/Ka)} = \frac{[D]^n}{[D]^n + Kd}$$

The ratio between bound and unbound species will be expressed as;

Equation 9-21

$$\frac{\theta}{1-\theta} = \frac{[D]^n}{Kd}$$

Taking the log of the **Equation 9-21** results in **Hill equation**,

Equation 9-22

$$\log\left(\frac{\theta}{1-\theta}\right) = n \log[D] - \log Kd$$

Where θ is the fractional saturation of DNA, which is the bound species, and $1-\theta$ is the free DNA.

9.1.4 Experimental procedure

9.1.4.1 DNA probe preparation

DNA was labeled with Cy3 at the 5' end, considering the directionality of MCM is 3' to 5' where Cy3 (Cy3 phosphoramidite) was ordered from Glen Research (3'-CCACACTCCCACAGTTGGTCCAGCAAGAGGTCCTGTTGACAATAGACGCy3-5'). Cy3 has maximum excitation at 545 nm and maximum emission at 570 nm. Labeled DNA was purified in BioCAD HPLC, using C18 column. Column C18 was equilibrated in solution containing 5% Acetonitrile and 100 mM TEAA (Triethyl ammonium acetate)

with flow rate of 1 ml/min. The sample (100 μ l suspended in H₂O) was loaded, and percentile of acetonitrile was increased from 5 % to 45 % in 30 min time. The DNA was observed with absorbance at 260 nm, and peaks were collected. Since the fluorescently labeled DNA is colored pink, oligonucleotide with pink color was visually detected and precipitated using following protocol. The eluted DNA was aliquoted into a microcentrifuge tube, and 360 μ l H₂O, and 40 μ l of 10 M NH₄Ac, and 1 ml of cold ethanol were added and placed in -80°C freezer for 15 min. The tube was spun for 20 min at 13,000 rpm, and supernatant was discarded. One milliliter of 70% ethanol was added, and spun again for another 10 min, and supernatant was taken out carefully not to disturb the hot pink colored pellet. It was left to dry for 15-20 min, then H₂O was added to resuspend the pellet. Concentration of DNA was calculated using A₂₆₀ with extinction coefficient of 477300 M⁻¹·cm⁻¹. ROX labeled oligonucleotide (5' ROX NHS Ester) was ordered from integrated DNA technologies (IDT) with HPLC purification. DNA was re-suspended in H₂O.

9.1.4.2 Fluoromax-3 autopolarizer machine setting

First, the Fluoromax-3 machine has to be turned on and set the water bath temperature to 25°C and prepare the sample by following protocol while it is warming up for about 20 – 30 minutes. The cuvette is stored in nitric acid (stock) at room temperature, so remove the cuvette carefully and wash with deionized water. Wash with ethanol, and wash one more time with deionized water. Dry the cuvette, and acquire the fluorescence buffer containing 5 mM MgCl₂, 25 mM Hepes-NaOH (pH 7.5), 2 mM DTT, and dilute proteins accordingly. Thaw Cy3 stock (1.5 μ M), Keep protein and DNA samples on ice at all times.

Log into the computer and click on the FluroMax-3 icon on the desktop of the computer. A box will prompt you to make a choice: choose FluoroMax-3 with Autopolarizers. Another box will appear, and on the tool bar, choose applications, and scroll down to instrument set-up. Click on the icon that represents the sample (the bottom box on the part that looks like an eye). Double click on polarizers, and switch both of the options to in (this step takes few minutes). Close the box.

Return to the applications option, click on constant wavelength option. In the top of the window, excitation and emission values must be typed in (in the case of Cy3, it is 545 and 570 and in the case of ROX, it is 588 and 608 for excitation and emission respectively). Below it, check off use polarization modes in the box, and then add all of the acquisition parameters by selecting them one by one and clicking add. At the bottom of this window, change the number of maximum trials to the desired value, which we usually set to 3. And integration time should be set between 5 – 10 seconds. Then click on “proceed to acquisitions”.

Click the “Start the acquisitions”, place cuvette containing 149 μl of the fluorescence buffer. Click the blank option, and the “Dark correction” option and click “Acquire now”, this will automatically collect the dark values, when the values appear on each row, then click “OK”. Then finally click on “Run the sample”. Once the values are recorded in the spreadsheet, add 1 μL DNA stock (1.5 μM). Instead of “Blank”, choose “Unknown”, and give it a name. Click “Run sample”. Record the values of anisotropy. In the case of Cy3, the anisotropy starts around 0.28, whereas ROX starts around 0.17. If the values are off, clean the cuvette, and start from the blanking step. Add small quantity of protein (0.5- 1.0 μl), and let it sit for at least 10 min. The time based experiments are

recommended for different proteins, since different proteins have different time it takes to reach its equilibrium. After the incubation, change name and click on “Run sample”. Keep increasing the concentration of protein as been added so that anisotropy value is saturated at some point.

9.1.4.3 Data Analysis

Once the anisotropy for each concentration of protein, as well as DNA alone is recorded, it can be analyzed using GraFit program (Erithacus Software Limited) (145), which is a software that allows for plotting binding curves, as well as computing Kd values. Usually, the anisotropy change is normalized before analysis. For example, the data collected from anisotropy should look like the following in excel spreadsheet as shown in Table 9-1.

Note that the final concentration of protein is calculated as follows:

Equation 9-23

$$\text{Final Conc}(x) = \frac{\text{Conc of prot}(x) \times \text{vol added}(x)}{150 + \text{Tot vol added}(x)} + \frac{\text{Conc of prot}(x-1) \times \text{vol added}(x-1)}{150 + \text{Tot vol added}(x)}$$

Where final concentration at some titration point x, or in the table above each row represent the titration point, is a sum of concentrations added at point x and the concentration of protein prior to addition of protein at point x, therefore x-1. Since the total volume is different from point x-1, the prior concentration must be readjusted by a factor of total volume at point x (see **Equation 9-23**). The original volume that the cuvette contains with DNA is 150 µl, as described in 9.1.4.2.

If you plot the protein concentration vs. anisotropy from Table 9-1, you get a graph as shown in Figure 9-1.

This plot can be directly analyzed by applying the **Equation 9-13**, but we can normalize the anisotropy value by:

Equation 9-24

$$\text{Normalized Anisotropy } (A_N) = \frac{A - A_{\min}}{A_{\max} - A_{\min}};$$

where A is the anisotropy value, A_{\max} and A_{\min} are maximum and minimum values for the anisotropy during the course of titration. Once the anisotropy is converted to the normalized anisotropy, the $A_{\text{final-initial}}$ value becomes 1, and the equation becomes a little simpler.

Once this is ready on the spreadsheet, open the GraFit program. Since there was no anisotropy equation available, the new equation called “ANISOTROPY” was created. Under the “definition” tab, write the quadratic equation as derived in **Equation 9-13** as follows:

Equation 9-25

$$dA = (dAt / (2*Dt)) * ((P + Dt + Kd) - \sqrt{(P + Dt + Kd)^2 - (4*P*Dt)}) \text{ for simple binding}$$

For cooperative binding, **Equation 9-18** is written as:

Equation 9-26

$$dA = dAt * ((P/Kd)^{nH}) / (1 + (P/Kd)^{nH})$$

Name the Y-name as: dA, indicating change in anisotropy as a function of concentration. Under the “X-variable” tab, write P, which stands for protein concentration in μM , and under the “Parameter” tab, write Kd, as dissociation constant. Under the “Constants” tab, write Dt and dAt, whose definition is the total concentration

of DNA in μM (D_t), and total change in anisotropy (dA_t), if normalized, dA_t value is 1. Once the equation is properly set, the equation has to be compiled by going to the “Equation” under the tool bar, and hold down to and click “Compile”, give this equation a name, such as “Anisotropy”. Then upload the data by going under “Add”, and add “Data”. Click on the Data icon which shows up on the left side. It is similar to excel, so copy the data on the excel file to the data, which can be named, such as “experiment 1” for convenience. Next, add a “Non-linear Fit”, a type of calculator for this analysis. Open the “Non-linear fit” icon, click the “settings” tab. For equation, choose “Anisotropy” equation that was created earlier. Weighing should be set to “Simple”, and Data table should be set to “experiment 1” that was created prior to this step. Data layout depends on which “Weighing” is chosen, but usually, choose, “x1, y1, y2, y3...”. X column should be the concentration of protein, most probably in the “experiment 1”, the first column, so choose that column. Last Y data column simply means up to where you want to analyze the data. If there are five protein samples, pick the last column in the list so all of them will be analyzed simultaneously. Next, go to “Constant” tab next to “Settings”. For DNA concentration (in μM), write 0.01, because we usually use 10 nM DNA concentration. At 10 nM concentration, the signal from the fluorescent at given condition is at least 10^5 , which is required for consistent results. For total anisotropy change, write 1.0 (for normalized data). Next, go to “Estimatrix” tab next to “Constant”, for Start, write 0.0005 or 0.0001, and for End, write 1. The Steps will be automatically filled out. Now the data can be fit. Go to the tool bar, “Calculate”, scroll down and click “Fit the data”. In the “Results” tab, the estimated K_d is listed. It is important to make the graph, so that we can visualize the fit of the data. So, go to the tool bar, and click on

“Add” and scroll down to “Graph”. Click on the graph icon that appears on the left, window will appear and choose “X/Y scatter” and click on “create” box. The wizard will appear, and for source, choose, “Non-linear Fit”, and “all fits” under the source choice. Click on “next” box, and choose any style or property preferred for the graph. And finish. The graph will appear, but the name of the samples will all be the same. Change from first row till last, in the order of the column (left to right) in the data set. If you want the Kd values to show up, place it by going to the tool bar, and click on “Draw”, scroll down to and click “Results listing”, the window will appear, and will ask for the source. Click “Non-linear fit”, and choose “Fit number”, 1 indicates the first column, 2 indicates the second column, and so on. It is not possible to pull out all the fit values at a time, it has to be drawn one at a time.

When the data is fit to the **Equation 9-25**, the GraFit program shows the graph as in Figure 9-3.

It is worth noting that the fitting of quadratic equation derived in **Equation 9-13** and simplified **Equation 9-14** are almost the same. Thus, it is reasonable to use the simplified cooperative binding equation, Equation 9-26, and determine whether the data fits better to the cooperative binding (see Figure 9-4). In the case of the example used here, the fitting of data to simple binding curve (Figure 9-3) or cooperative curve (Figure 9-4) do not make significant differences. However, it appears that some data fit better to the cooperative binding equation, and it would be interesting to fit all the data and compare whether the wild type or mutant protein has either lost or gained cooperative binding due to mutation.

9.3 Determination of kinetic parameter for ATPase activity

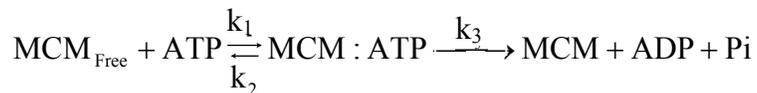
9.3.1 General principle

When we are comparing the same enzyme with mutations, it is useful to have kinetic parameters as a quantitative evaluation tool to understand the differences and the impact of mutation. For MCM and its mutants, we were able to determine steady state kinetic parameters, K_M and k_{cat} of ATPase activity in the absence or presence of ssDNA or dsDNA. The steady state kinetic parameters, K_M and k_{cat} . In steady state kinetics, the rate of activity is determined at state where the formation of product as a function of time is constant. Therefore, it is important that whatever is measured in the linear phase of reaction. In the case of ATPase activity by MCM helicase, the concentration of substrate (ATP) versus the rate of hydrolysis at each concentration is plotted, and it was fit to a Michaelis-Menten equation as described below.

9.3.1.1 Calculation

Derivation of Michaelis-Menten equation using MCM catalyzing ATP hydrolysis:

Equation 9-27



MCMFree + ATP

The Pi released can be quantitated with ^{32}P - γ -ATP.

In steady state kinetics, the concentration of intermediate, MCM:ATP is assumed to be constant, thus, the rate of catalysis is equal to the product of the concentration of MCM:ATP complex and k_3 , and expressed as:

Equation 9-28

$$V_0 = k_3[\text{MCM : ATP}]$$

And the [MCM:ATP] can be expressed in terms of rate of formation and breakdown of MCM:ATP by following:

Equation 9-29

$$\text{Rate of formation of MCM : ATP} = k_1[\text{MCM}][\text{ATP}]$$

$$\text{Rate of breakdown of MCM : ATP} = (k_2 + k_3)[\text{MCM : ATP}]$$

Again, in the steady state kinetics assumption is that MCM:ATP is constant, resulting in following equation:

Equation 9-30

$$\text{Rate of formation} = \text{Rate of breakdown}$$

$$k_1[\text{MCM}][\text{ATP}] = (k_2 + k_3)[\text{MCM : ATP}]$$

Rearranging Equation 9-30 can be written as:

Equation 9-31

$$\frac{[\text{MCM}][\text{ATP}]}{[\text{MCM : ATP}]} = \frac{(k_2 + k_3)}{k_1}$$

; substituting *Michaelis constant*, K_M for $\frac{[\text{MCM}][\text{ATP}]}{[\text{MCM : ATP}]}$

Equation 9-32

$$K_M = \frac{(k_2 + k_3)}{k_1}$$

Inserting the equation 9-32 into equation 9-30, and solving for [MCM:ATP] results in following equation:

Equation 9-33

$$[\text{MCM} : \text{ATP}] = \frac{[\text{MCM}][\text{ATP}]}{K_M}$$

Since the [MCM], the free MCM in reaction can be written as:

Equation 9-34

$$[\text{MCM}] = [\text{MCM}]_T - [\text{MCM} : \text{ATP}]$$

Substituting the expression in equation 9-34 into equation 9-33 gives:

Equation 9-35

$$[\text{MCM} : \text{ATP}] = \frac{([\text{MCM}]_T - [\text{MCM} : \text{ATP}])[\text{ATP}]}{K_M}$$

Solving equation 9-35 for [MCM:ATP] complex gives:

Equation 9-36

$$[\text{MCM} : \text{ATP}] = \frac{[\text{MCM}]_T / K_M}{1 + [\text{ATP}] / K_M}$$

Or

Equation 9-37

$$[\text{MCM} : \text{ATP}] = [\text{MCM}]_T \frac{[\text{ATP}]}{[\text{ATP}] + K_M}$$

Substituting this expression for [MCM:ATP] into equation 9-28 results in:

Equation 9-38

$$V_0 = k_3[\text{MCM}]_T \frac{[\text{ATP}]}{[\text{ATP}] + K_M}$$

The maximal rate, V_{\max} is achieved when all free MCM is saturated with substrate, where

[MCM]=[MCM:ATP], therefore,

Equation 9-39

$$V_{\max} = k_3[\text{MCM}]_T$$

Substituting the equation 9-39 into equation 9-38 yields in Michaelis-Menten equation:

Equation 9-40

$$V_0 = V_{\max} \frac{[\text{ATP}]}{[\text{ATP}] + K_M}$$

9.3.2 Experimental procedure

9.3.2.1 Conditions

Since it was determined that 50 nM of MCM protein for 30 min from 0.01 mM to 1 mM ATP, reaction mixture (15 μ l) containing 25 mM HEPES-NaOH (pH 7.5), 1 mM dithiothreitol, 5 mM MgCl_2 , 100 $\mu\text{g/ml}$ BSA, and 0.2 μM [γ - ^{32}P]ATP (3,000 Ci/mmol; GE Healthcare) with various amount of unlabeled ATP, and 50 nM of MCM (as monomer) in the presence or absence of 50 ng ssDNA (5'-

GGCAGATAACAGTTGTCCTGGAGAACGACCTGGTTGACACCCTCACACCC -

3'). Reaction was quenched by placing the reaction tubes on ice. An aliquot (1 μ l) was spotted onto a polyethyleneimine cellulose thin layer plate, and ATP and P_i were separated by chromatography in 1 M formic acid and 0.5 M lithium chloride. The extent of ATP hydrolysis was quantitated by phosphorimager analysis.

9.3.2.2 Data Analysis

The quantitation of P_i hydrolyzed by MCM helicase was quantitated as follows. First, the QUANTL phosphorimager analysis tool was used to determine the % hydrolysis on each column, at this point, the background was subtracted from all % hydrolysis to obtain the actual % hydrolysis. The % hydrolysis was converted to the actual number of P_i released by multiplying the number of ATP present in the reaction, resulting in μ M ATP hydrolyzed, in a total reaction. Then that value was divided by the time (min) of reaction, resulting in μ M ATP hydrolyzed per min. This then was divided by the amount of protein used to determine the μ M ATP hydrolyzed per min per moles of MCM (as a monomer). Once all the values for each ATP concentration are calculated, one can plot the graph of ATP concentration versus the rate of hydrolysis per moles of protein. The plot was fit to Michaelis-Menten equation in order to determine the K_M , V_{max} , and k_{cat} value using GraFit program.

The detail of how GraFit was used to determine the K_d values was explained in 9.1.4.3.

The Michealis-Menten equation derived in the previous section was used and written as:

Equation 9-41

$$v = V_{\max} * S / (K_m + S);$$

Where v is rate of hydrolysis, and therefore it is the data determined from experiments, and V_{\max} and K_m is the parameters to be determined from the fit, and S , is the substrate concentration, and thus x variable. When the data is available, set the Michaelis-Menten equation and run the non-linear fit. The fitting of the data can be displayed as graph, in linear or logarithmic scale.

9.4 Differential scanning calorimetry (DSC)

9.4.1 General principle

Differential scanning calorimetry (DSC) detects the thermally induced transitions by measuring the differential power of thermally coupled sample and reference cells. Both the sample and reference cells are in adiabatic environment (closed system) which has no exchange of heat with the surroundings. As the protein unfolds, contributions from exothermic (i.e. bond breaking) and endothermic (i.e. disturbances of solvent molecule) processes result in a differential power of cells that gives rise to a peak with its maximum representing a melting temperature, T_m (review from (146)). Integrating the curve of excess heat capacity (C_p) as a function of temperature results in transition enthalpy ΔH°_m (see calculations in later section), and the difference between the baseline before and after transition is ΔC_p .

MCM helicase appears to have irreversible thermal transition (or unfolding), since there is no reappearance of a peak after the first scan, which is followed by irreversible process (Figure 9-6). However, the thermal transition of MCM is independent of scan

rate since T_m calculated for both slow and medium scan rate was within $\pm 3^\circ\text{C}$, suggesting that the thermal transition precedes the irreversible step. This was simply determined by measuring the T_m with different scan rates, slow (15 deg/hr) and medium (60 deg/hr). Since unfolding of the MCM was determined to be independent of rate, the transition can be considered an equilibrium between folded and unfolded states (two state system) (124,147). To be consistent, proper baseline was obtained by subtracting the second scan from the first scan from the first scan (Figure 9-7). At the end of the run, the DSC profile contains three parameters, temperature, power, and scan rate. This profile is transformed by EXAM program to yield excess heat capacity and temperature.

9.4.1.1 Thermodynamics of denaturation

The stability of proteins is quantitated by the standard free energy ΔG° [difference in Gibbs energy between two different states, folded (F) and unfolded (U)]. At equilibrium in a reversible system (where $\Delta G=0$), ΔG° is related to the equilibrium constant (K) between the two states at a certain temperature (T);

Equation 9-42

$$\Delta G^\circ(T) = -RT \ln K(T); \text{ where } K = \frac{[F]}{[U]}$$

where R is the gas constant ($8.314 \text{ J} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$) and T is the temperature in Kelvin.

Also ΔG° can be written in terms of two contributions;

Equation 9-43

$$\Delta G^\circ(T) = \Delta H^\circ(T) - T\Delta S^\circ(T)$$

where ΔH° and ΔS° are the enthalpy and entropy changes at the temperature at which ΔG° is measured. Taking the derivative of this equation (Equation 9-43) as a function of temperature (T) can break it down to following equations:

Equation 9-44

$$\int_F^U dH = \int_F^U Cp dT \text{ Or } \Delta H^{FU} = \Delta Cp dT = \Delta Cp^{FU}(T_U - T_F) + C$$

Where Cp is the heat capacity, and FU denotes the difference in value between folded and unfolded states. And $\Delta S(T)$ is defined by;

Equation 9-45

$$\int_F^U dS = \int_F^U \frac{Cp}{T} dT \text{ Or } \Delta S^{FU} = Cp^{FU} \ln\left(\frac{T_U}{T_F}\right) + C \text{ where C is a constant or reference}$$

Thus, taken together, free energy change between two states as a function of temperature is;

Equation 9-46

$$\Delta G^\circ(T) = \Delta H^\circ(T) - T\Delta S^\circ(T) = \Delta H^\circ(T) + \Delta Cp(T - T_o) - T\left[\Delta S^\circ + \Delta Cp \ln\left(\frac{T}{T_o}\right)\right];$$

Which can be simplified to:

Equation 9-47

$$\Delta G^\circ(T) = \Delta H^\circ(T) - T\Delta S^\circ(T) + \Delta Cp\left[(T - T_o) - T \ln\left(\frac{T}{T_o}\right)\right]$$

Combining the equations 9-42 and 46 yield;

Equation 9-48

$$-RT \ln K(T) = \Delta H^\circ(T) + \Delta C_p(T - T_o) - T \left[\Delta S^\circ + \Delta C_p \ln \left(\frac{T}{T_o} \right) \right];$$

When values of K can be determined experimentally as a function of temperature, the data can be fitted to yield values for ΔH° (van't Hoff enthalpy, not calorimetrically determined enthalpy). If ΔH° is independent of the temperature, ($\Delta C_p=0$), a plot of $\ln K$ as a function of $1/T$ will give a straight line with slope of $-\Delta H^\circ/R$. However, ΔH° is usually temperature-dependent.

When a protein unfolds, heat capacity change (ΔC_p) is observed. This change is said to be caused primarily due to the disorganization of the solvent molecules around the hydrophobic residues being exposed upon denaturation (148).

The unfolding transition occurs at a transition midpoint or melting temperature, T_m . Assuming a two-state transition, where $\Delta G^\circ = 0$, the T_0 is equal to T_m . The term $\Delta S^\circ(T_m)$ will be equal to $\Delta H^\circ(T_m)/T_m$, that would result in an equation with only terms of ΔH and ΔC_p ;

Equation 9-49

$$\Delta G^\circ(T) = \Delta H^\circ_m \left(1 - \frac{T}{T_m}\right) + \Delta C_p \left[(T - T_m) - T \ln \left(\frac{T}{T_m} \right) \right]$$

where ΔH°_m is the value of ΔH° at T_m ($\Delta H^\circ(T_m)$).

DSC measures the ΔC_p over temperature, thus the $\Delta H(T)$. The enthalpy of the denaturation has been associated with the content of secondary structure of a protein. The peak transition seen is an outcome of contributions from endothermic reactions, such

as disruption of hydrogen bonds, and exothermic ones, such as the breaking of hydrophobic interactions.

9.4.2 Experimental procedure

9.4.2.1 Sample preparation

The purified proteins were dialyzed at 4°C against buffer containing 20 mM Hepes-NaOH (pH 7.5), 100 mM NaCl and 10% glycerol. The protein concentrations were determined using an absorbance at 280 nm and a calculated extinction coefficient of ϵ_{280} of $28730 \text{ cm}^{-1}\cdot\text{M}^{-1}$, using the dialysate (the solution outside the dialysis bag) as blank. The dialysate was kept for a reference cell and for protein dilution. Proteins were diluted with dialysate to final concentration of 10 μM in the presence of 10 mM MgCl_2 . DNA and nucleotide, depending on the experiments was dialyzed or suspended in dialysate buffer, and appropriate amount of these substrates were added to final volume of 700 μl (The volume of cell is about 0.55 ml, so it is important have at least 700 μl). The final sample was incubated at 60°C for 10 minutes to activate the protein. The sample was spun for 10 minutes at 13,000 rpm at room temperature. After spinning, two aliquots of 5 μl of sample were added to 1 ml of Bradford reagent each (to determine consistent absorbance), and vortexed and left for 5 minutes. The rest of the sample was loaded onto the cell carefully without any air bubble.

The absorbance at 595 nm from Bradford assay was converted to protein concentration in mg/ml by using the following standard made from BSA (Figure 9-8). To further confirm that Bradford method and absorbance at 280 nm correlates, the same wild type MCM sample used to measure the absorbance at 280 nm was also measured at

595 nm by addition of 5 μl to Bradford reagent. The concentration in mg/ml determined from BSA standard vs. absorbance at 280 nm is plotted (Figure 9-9). The protein concentration determined using Bradford assay after the sample was prepared was in good agreement with protein concentration measured with absorbance at 280 nm (ϵ reported = 28730 cm M^{-1} ; ϵ calculated = 29200 cm M^{-1} ; Error = 1.6 %).

9.4.2.2 DSC machine setting

Enter the scan rate, whether it is a slow or medium scan, enter 60 or 15 deg/hour respectively. The scan can be started from any temperature, but for this particular study described in this dissertation mostly were started from 25°C to 85 or 90°C. Pre- and post-equilibrium times were set to 5 min and 2 min respectively. Enter the name of the sample consisting of 6 letters with information of date, scan rate, and sample name (by alphabetical order).

Before any samples were run, buffer or dialysate was loaded to both sample and reference cells, to make sure that the dialysate does not contain impurities and that the machine is behaving normally. The buffer in the reference cell was kept untouched as long as a series of studies were run.

After the run, the samples cells were washed with H₂O about 10 times with the syringe until there is no precipitant floating inside the syringe. Then buffer was used to rinse the cell 3 times for loading the next sample.

9.4.2.3 Data Analysis using EXAM program

EXAM program is a two state thermodynamic analysis program that was created by William H. Kirchhoff (122). Briefly, this program uses the least square fit of Equation

9-52 (taken from (122)) to DSC data to determine the ΔH and T_m in two state system.

The detail of deriving the equation can be found in (122).

Equation 9-50

$$\frac{dQ}{dT} = (1 - \alpha)[Ba + B'a(T - T_o)] + N \left(\frac{\Delta H(T)^2}{RT^2} \right) F(\alpha) + \alpha[Bb + B'b(T - T_o)]$$

where N is the number of moles in a sample, ΔH is the molar transition enthalpy (van't Hoff enthalpy), terms $[Ba+B'a(T-T_o)]$ and $[Bb+B'b(T-T_o)]$ applies to pre- and post-transition asymptotic baselines. The α , when $n=1$ ($A \rightarrow B$ transition), is;

Equation 9-51

$$\alpha = \frac{1}{1 + e^{-z}} ; \text{ where } z = \frac{\Delta H_o}{RT_o^2} (T - T_o);$$

Where subscript "o" applies to midpoint of the transition

And

Equation 9-52

$$F(\alpha) = \alpha(1 - \alpha)$$

When α is α and $1-\alpha$ have opposite behavior (i.e. α approaches 0 when $T \rightarrow -\infty$, but it approaches 1 when $T \rightarrow +\infty$; while the term " $1-\alpha$ " has the opposite behavior to α).

For analyzing the data, first, you have to convert the DSC (.dsc) files into a text (.txt) document. Once it is done, you can open the EXAM program, and subtract the second peak (baseline) from the first peak by typing the following commands:

```
>subtract
```

```
: a:mcm118m3.txt (a=location of file, name of data (mcm118m3.txt), 1st scan)
```

:a:mcm118m2.txt (2nd or reference scan)

:a:mcm11832 (give a new name for the file), then press [Enter]

Temp: 25 90 (range of temperature to be analyzed), then press [Enter]

:2 (selection for the number of points to use for fit), then [Enter]

:1(selection for enter degree of polynomial), [Enter]

Once this subtraction is done, the data is ready to be analyzed by following commands.

>exam, then press [Enter]

>[Enter] (to continue)

:a:mcm11832 (type in the entry – the name that you saved in the file), then [Enter]

Do you wish ... values? Y/N [N] Type [N], then [Enter]

Enter number....[1] Type [4] (for Transform Data), then [Enter]

Enter selection....: Type [8], then [Enter]

$Y \rightarrow s(Y+\Delta Y)$. Enter value for ΔY [-Ymin]: 0

Enter value for s [1.0]: 0.06975 (converts $\mu\text{Cal}/\text{min} \rightarrow \text{Joules}/\text{sec}$)

Enter selection.....: Type [2] (for $y \rightarrow y/(dX/dt)$ ($t=N\Delta t$));

Transformation.....: [Enter]

Caution! : Continue with transformation? Y/N N] Type [Y]

Replace.....? Y/N [N] Type [N]

Then hit [Enter] until you see the graph.

Place the cursors in regions that contain the peak using arrow key. Enlarge the peak by holding the [Shift] key and hit [F5]. Once the peak is enlarged, it is easier to see where the peak starts to show and also return to the base line position. Lock the cursors on both ends of the peak. Press hit [F7] “Start Fit” for data analysis. On top of the graph,

it asks what type of data is being analyzed. Select [3] for DSC, this automatically makes the graph to fit the Equation 9-51. Then again, hit [F7] to fit the curve. The graph will appear as shown in Figure 9-8. Keep the image by pressing [Ctrl]+[Alt]+[T] at the same time. Once the snapshot is taken, press [Shift]+[F6] to obtain the area under the peak (ΔH_{cal}) (Figure 9-9). This analysis procedure can be repeated by pressing [Esc], and continue pressing [Enter] until the original graph appears. Using [Shift] and [F5], place cursor in different locations to see better fit (lower σ value) can be obtained.

Once the values for $\Delta H_{van't Hoff}$ and area under the curve is obtained, simple calculation is needed to convert the area (mJoule) to ΔH_{cal} .

Equation 9-53

$$\Delta H_{cal} = \frac{\text{area under the peak (mJoule)}}{\text{number of mole present in cell (concentration of protein} \bullet 0.51 \text{ lml)}}$$

9.4.3 Rough estimation of bound species at higher temperature

Since the DSC experiments with MCM and DNA, as well as nucleotide assumed that all proteins were bound by the DNA or nucleotide, it was important to estimate roughly how much of protein was bound to DNA. In order to roughly estimate the fraction of protein bound, the total concentration of proteins and DNA, and known Kd at room temperature can be used in the context of quadratic equation. As described above, bound species, B, can be determined using quadratic equation, Equation 9-56.

Equation 9-54

$$B = \frac{-(M_T + D_T + K_d) \pm \sqrt{(M_T + D_T + K_d)^2 - 4D_T M_T}}{2}$$

In DSC experiments, amount of protein is set to around 10 μM so that enough heat of denaturation can be observed and analyzed. Anisotropy experiments have shown that the MCM binds to short oligo (50mer) with K_d of 130 nM. Using this K_d value as a reference, one can determine how much fraction of MCM-DNA complex is formed, at 10 μM of MCM and 10 μM of DNA in the DSC sample, but simply putting the number into equation 6, as following in excel shown in table 1. Where [Protein] $_i$ is the protein concentration, [DNA] $_i$ is the DNA concentration, K_d is the assumed dissociation constant (therefore, it is variable), a is the coefficient of B^2 , (see equation 5.4 for clarity), b is $-(K_d + M_T + D_T)$ the coefficient of B , and c is the constant term, $M_T D_T$. Since this equation contains “ \pm ” sign, it has both positive and negative answers.

Unfortunately, we do not have the dissociation constant for ATP, or ATP non-hydrolyzable analogue or ADP, although we tried to measure it using isothermal titration calorimetry (ITC). Estimation of complex formed is very important, at least before starting the experiments, so that we can expect whether there should be a complex formed or not.

Also this estimation is at room temperature, so at higher melting temperature, the estimation may be same, greater or less, depending on the protein, and that has to be analyzed separately.

9.5 Circular Dichroism (CD)

9.5.1 General principle

Circular dichroism refers to a light absorption of circularly polarized light. This far-UV CD absorption is sensitive to the protein secondary structures, which are optically active or non symmetrical molecule. There are signature pattern of spectra that predicts whether the protein contains mostly α helix or β sheet, or random coil. One signature pattern is a drop in absorption at 222 nm, this pattern allows for detection of α helix in protein structure. However, CD cannot predict where the alpha helices that are detected are located within the molecule. CD can be used to study how the secondary structure of a molecule changes as a function of temperature (like DSC) or of the concentration of denaturing agents, such as urea or guanidium hydrochloride (Figure 9-12). CD, like gel filtration analysis, is useful to verify the mutants that were generated have the same protein fold as the wild type, before doing any other experiments.

9.5.1.1 Calculation

The circularly polarized light obey Beer's law, and circular dichroism (CD) is defined as the difference in extinction coefficients of right handed or left handed polarized light:

Equation 9-55

$$\Delta A(\lambda) = A_L(\lambda) - A_R(\lambda) = [\varepsilon_L(\lambda) - \varepsilon_R(\lambda)]\ell c = \Delta\varepsilon\ell c$$

Where L and R symbolizes the direction of polarized light (left or right handedness), ε is the molar absorption with units of $\text{L mol}^{-1} \text{cm}^{-1}$, ℓ is the path length in cm, and c is the concentration in mol L^{-1} . In general, the $\Delta A(\lambda)$ is presented as ellipticity, θ , which relates the measurement to optical rotator dispersion. Ellipticity in degrees is related to the difference in absorbance by $\Delta A(\lambda) = \theta/32.98$. The corresponding expression for $\Delta\varepsilon$ is molar ellipticity in $\text{deg}\cdot\text{dl}/\text{mol}\cdot\text{dm}$, $[\theta] = 3298\Delta\varepsilon$. The circularly polarized light, therefore, are produced by asymmetric molecules, while symmetric molecules produce linearly polarized light, and no circularly polarized light.

9.5.2 Experimental procedure

9.5.2.1 Sample preparation

The proteins (1~3 mg/ml) were dialyzed in buffer containing 20 mM HEPES-NaOH (pH7.5), 100 mM KCl, and 10% glycerol. About 5 μl of protein was spotted on one of the circular 0.005 cm path length quartz cell plate and covered with another cell plate. While putting the plate, it has to be carefully done so that there is no bubble in between the cells. The cells were placed in the Model J-720 Jasco Spectropolarimeter at room temperature. Far-UV wavelength scans were recorded from 250 to 190 nm. Averages for three CD spectra were presented.

9.5.2.2 Machine setting

Turn on the N_2 gas (10 scf/h) to prevent the lamp from producing a lot of ozone (O_3) which corrodes the lamp. Turn on the water valve which cools the lamp during the experiment leave both gas and water on for about 10 min. Turn on the power 1, which should be set to 95 Volts. Press “start” briefly, then again press “start” for the second

time briefly. The current should be set to 24-25 Amps. Turn on the power 2 followed by turning on the computer. On the desk top, open the program called Jhardct1. Set the wavelength that you want to run your samples (250 to 190 nm) with sensitivity of 20 mdeg. Also type in the path length of the cell (it is 0.005 cm, for proteins with concentration of about 2 mg/ml).

9.5.2.3 Data Analysis

The raw CD spectra are converted to mean residue ellipticity (MRE), by multiplying the concentration of protein by number of amino acid present in the protein. For example, the MRE of MCM protein at 3.4 mg/ml is converted to MRE as follows:

Equation 9-56

$$3.4\text{mg/ml} \cdot \frac{1\text{mole}}{77000\text{g}} \cdot 666\text{amino acids} = 0.029\text{M}$$

Thus, the concentration of amino acid in a 3.4 mg/ml sample is 0.029 M, and this value is used to correct the raw CD data to MRE.

Ellipticity results were expressed as mean residue ellipticity $[\theta] = \text{degrees} \times \text{centimeter}^2 \times \text{decimole}^{-1}$ over wavelength (nm).

9.6 Tryptophan fluorescence

9.6.1 General principle

MCM protein contains one tryptophan residue at position 363, which has been predicted to be located in the helix 2 insert (H2-ins), where it is a unique insert found only in MCM and few other AAA+ proteins among superfamily 3. This H2-ins has been

shown to play an important role in MthMCM (102), and it may serve as a great probe for conformational change in MCM protein.

Tryptophan has intrinsic fluorescence, a property that when it is excited, it emits fluorescence. The intensity and wavelength of maximum fluorescence emission of tryptophan depend on the surroundings of tryptophan, enabling us to assess the local conformational change. The spectrum of Trp fluorescence shifts to shorter wavelength and the intensity of the fluorescence increases as the polarity of the solvent surrounding the tryptophan residue decreases (149). Tryptophan residues which are buried in the hydrophobic core of proteins can have spectra which are shifted up to 20 nm compared to tryptophan located on the surface of the protein. Tryptophan fluorescence intensity can be quenched by neighboring protonated acidic groups such as Asp or Glu. Therefore it is often used as a reporter for the conformational change or binding of substrate, though it is suggested not to over-interpret the data. The principles and details of tryptophan fluorescence are described in (149).

9.6.1.1 Calculations

The calculations for tryptophan fluorescence can be necessary depending on type of experiments and purpose it is serving. For example, if MCM protein were to be titrated by DNA, the fluorescence signal without DNA, and quenched signals as DNA is titrated are normalized, and fit to a quadratic binding equation, as described in Equation 9-13.

However, according to the experiments, MCM has tryptophan fluorescence signal which continuously quenches (Figure 9-13). This was shown to be likely due to photobleaching effect, where as tryptophan (protein) is being excited, the tryptophan's

intrinsic property to be able to excite the molecule is destroyed by photochemical reaction, resulting in decreased fluorescent signal. When the slit width was decreased to 4 nm from 5 nm, used to repeat the experiment, the effect of photobleaching was decreased but fluorescence intensity also decreased (Figure 9-14). Titration of DNA was performed with slit width set at 4 nm, but single experiment was error prone. Thus, we pursued anisotropy experiments.

Instead, we decided to make several mixtures containing different amount of DNA, and measured the intensity of tryptophan.

9.6.2 Experimental procedure

9.6.2.1 Sample preparation

MCM wild type and β hairpin mutant proteins as negative control (~200 nM to give about 10^5 signal) were measured in the presence or absence of either increasing amount or set amount of DNA in separate tubes. Each sample contained final volume of 150 μ l, containing 25 mM HEPES-NaOH (pH 7.5), 2 mM DTT, 5 mM $MgCl_2$, 100 nM MCM protein, with or without increasing amount of DNA. The samples were incubated at room temperature for 10 min, and spectra were taken at 25°C using FluoroMax-3 (Jon Yvon). Excitation was set at 295 nm and emission spectrum was set between 310 and 400 nm.

9.6.2.2 Data Analysis

Since it was not possible to titrate the tryptophan fluorescence signal with DNA, large premix was aliquoted into separate tubes with different DNA or different amount of DNA. In this case, the data can be qualitatively analyzed. The data can be plotted, signal

vs wavelength from 310 to 400 nm. Since the signal is quenched in the wild type protein, but not in the b hairpin mutant, which has been shown to have no interaction with DNA, the quenched signal is due to DNA.

Table 9-1 Example of how the anisotropy value is recorded as protein is titrated into the fluorescently labeled DNA.

Conc of protein (μM)	Vol added (μl)	Total vol added (μl)	Final conc(μM)	Anisotropy
0	0	0	0	0.1716
5	0.5	0.5	0.016611296	0.2132
5	1	1.5	0.04950495	0.2486
15	0.5	2	0.098684211	0.2756
15	1	3	0.196078431	0.317
15	2	5	0.387096774	0.3442
30	2	7	0.76433121	0.3511
30	2	9	1.132075472	0.3558

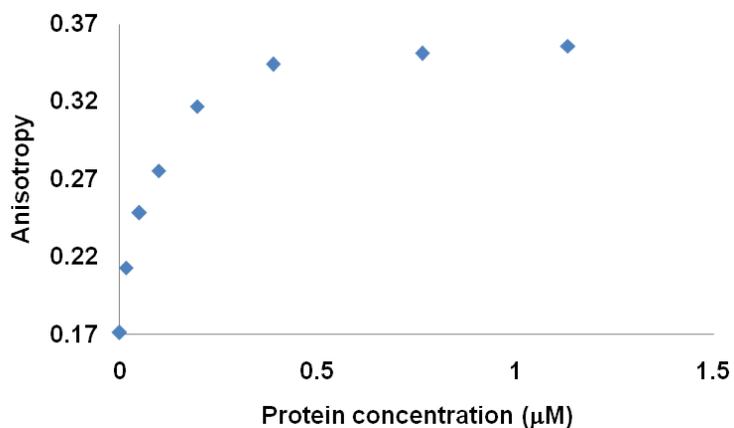


Figure 9-1 Example of anisotropy vs. protein concentrations plot.

The anisotropy value and final protein concentration from Table 9-1 is plotted using excel spread sheet..

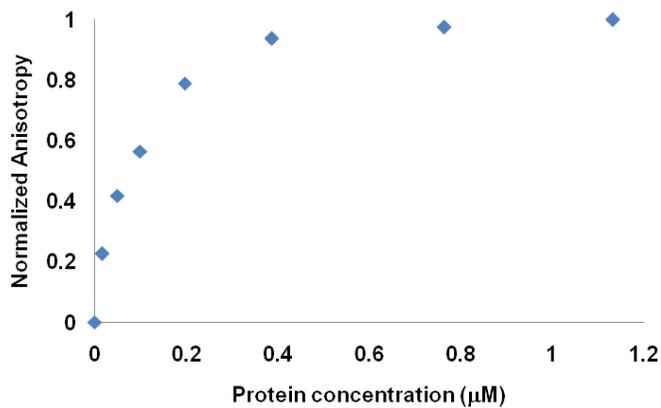
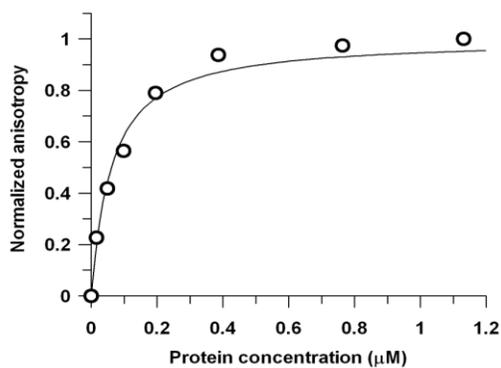


Figure 9-2 Normalized anisotropy.

Example of how the normalized anisotropy graph looks. The y-axis value is changed from 0-1, instead of actual values of anisotropy. The anisotropy values of Table 9-1 was converted to normalized anisotropy using Equation 9-24.

A) Linear scale



B) Logarithmic scale

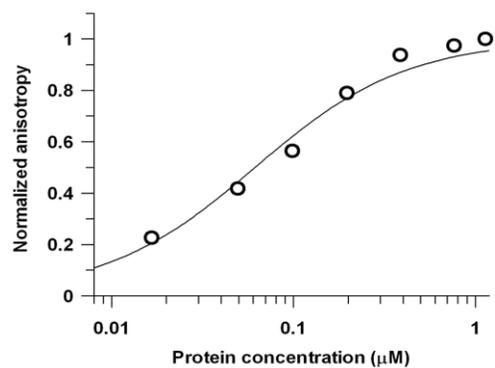
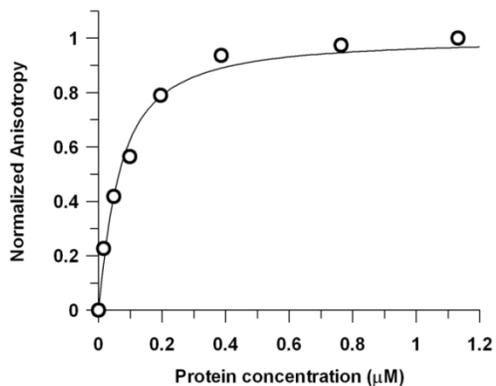


Figure 9-3 Example of MCM mutant (F172I) binding to DNA was fit using GraFit 5.0.

The data from Table 9-1 was normalized, and fit to Equation 9-25, a quadratic equation derived in Equation 9-13. The graph was plotted in linear scale (A) and logarithmic scale (B). The binding constant (K_d) calculated was 56.4 nM, with standard error of 5.8 nM.

A) Linear scale



B) Logarithmic scale

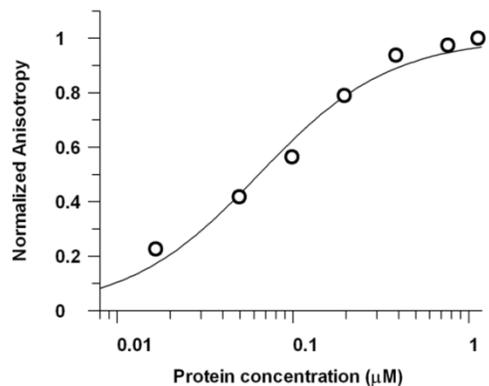
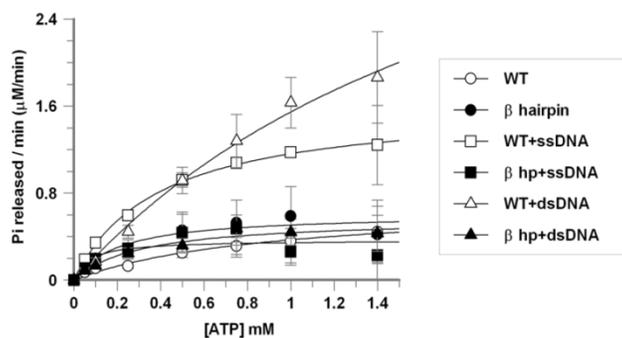


Figure 9-4 Example of MCM mutant (F172I) binding to DNA data fitting to cooperative binding equation.

The data from Table 9-1 was normalized, and fit to Equation 9-25, a cooperative binding equation derived in Equation 9-13. The graph was plotted in linear scale (A) and logarithmic scale (B). The binding constant (K_d) calculated was 64.1 nM, with standard error of 5.7 nM, with Hill coefficient of 1.16 ± 0.11 .

A) Linear scale



B) Logarithmic scale

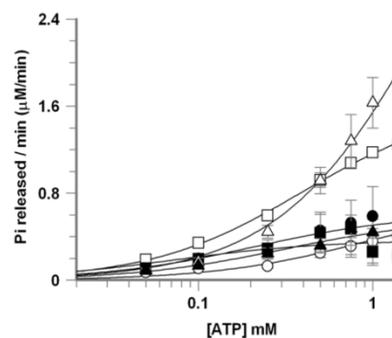


Figure 9-5 Example of ATPase data fitting to Michaelis-Menten equation using GraFit program.

ATPase activity of the wild type (WT) protein and β hairpin mutant (β hp) (50 nM each) was determined in the presence of ss or ds DNA (circular ϕ X174 and RF respectively) with increasing amount of substrate, ATP. The x-axis was plotted in linear (A) and logarithmic (B) scale for visualizing the “wellness” of the fit.

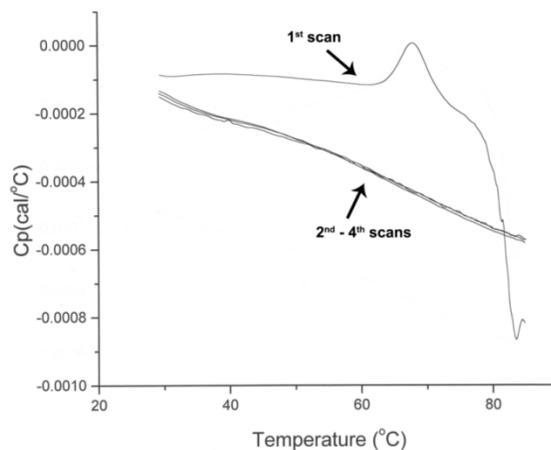


Figure 9-6 Example of DSC scan.

First scan contains endothermic protein denaturation peak, followed by aggregation of protein. It is an irreversible reaction, since after the first scan, there is no reappearance of transition. The excess heat capacity is expressed as Cp (cal/degree) in Y-axis, and temperature is in Celsius (°C).

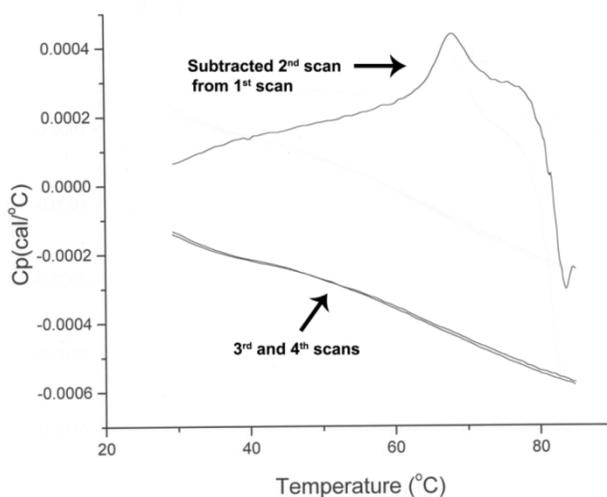


Figure 9-7 Example of subtraction of baseline (second scan) from the first scan.

The second scan from Figure 9-6 is subtracted from the first scan.

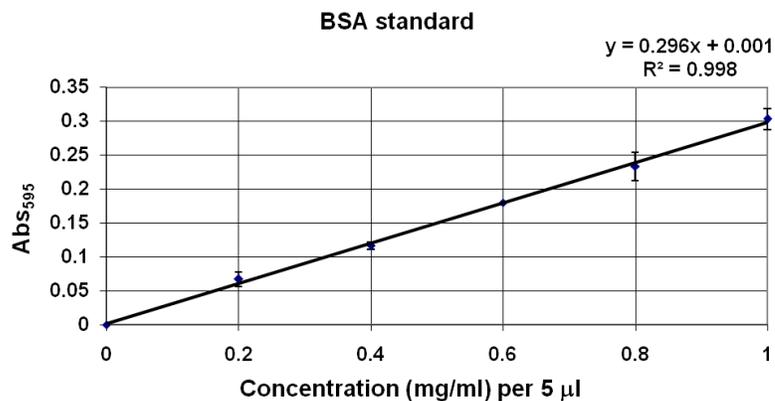


Figure 9-8 BSA standard plot.

Five microliters of BSA at corresponding concentrations were added to 1 ml of Bradford reagent. Absorbance at 595 nm was measured after 5 min incubation at room temperature.

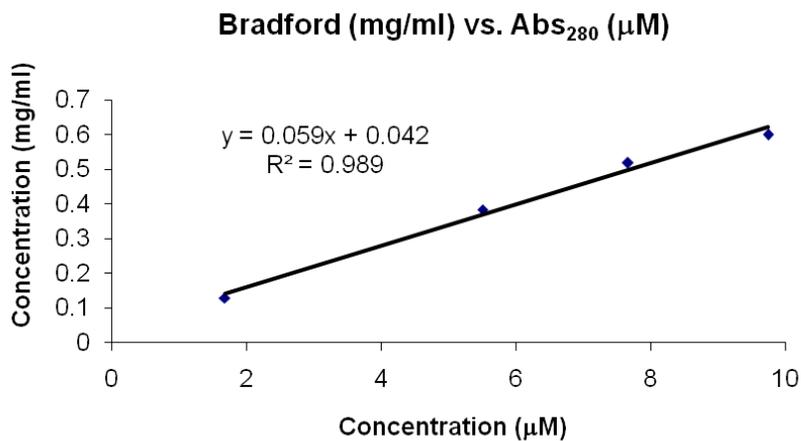


Figure 9-9 Standard curve of Bradford assay and spectroscopic (Abs₂₈₀ nm) determination.

The concentration of MCM protein was initially measured at A₂₈₀. The same sample was diluted to determine the concentration using Bradford assay.

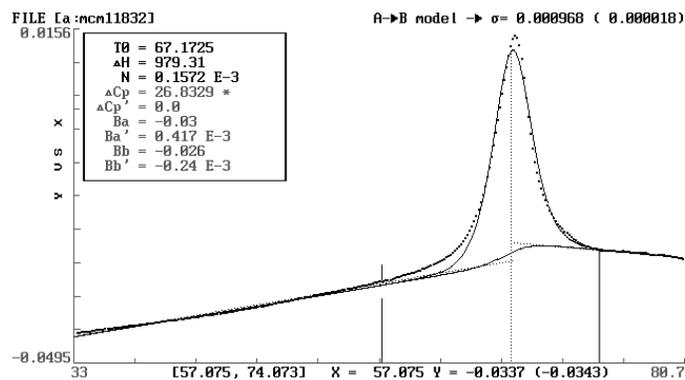


Figure 9-10 Example of DSC analysis using EXAM program.

Using a $A \rightarrow B$ transition model. T_0 is the T_m , ΔH is the transition enthalpy in kJ/mole. σ shown in the upper corner is how well the equation is fit to the actual data. Lower the σ better the fit.

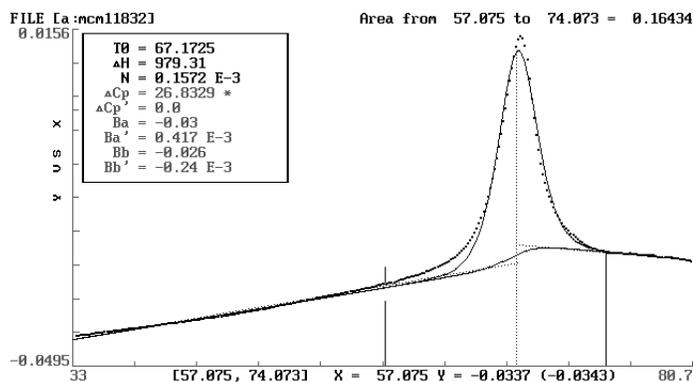


Figure 9-11 Example of DSC analysis to obtain ΔH_{cal} .

The value shown in the right top corner is the area under the curve, in mJoules.

Table 9-2 Rough estimation of MCM bound by DNA based on K_d obtained from room temperature, using quadratic equation.

[Protein]	[DNA]	Kd	a	b	c	minus b	b ²	4ac	2a	sqrt(b ² -4ac)	solve for B		Bound species (%)
											positive	negative	
1.00E-05	1.00E-05	1.00E-07	1	-2.01E-05	1.00E-10	2.01E-05	4.04E-10	4.00E-10	2	2.00E-06	1.10512E-05	9.04875E-06	90.5
1.00E-05	1.00E-05	1.00E-06	1	-2.10E-05	1.00E-10	2.10E-05	4.41E-10	4.00E-10	2	6.40E-06	1.37016E-05	7.29844E-06	73.0
1.00E-05	1.00E-05	1.00E-05	1	-3.00E-05	1.00E-10	3.00E-05	9.00E-10	4.00E-10	2	2.24E-05	2.61803E-05	3.81966E-06	38.2
1.00E-05	1.00E-05	1.00E-04	1	-1.20E-04	1.00E-10	1.20E-04	1.44E-08	4.00E-10	2	1.18E-04	0.000119161	8.39202E-07	8.4
1.00E-05	1.00E-05	1.00E-03	1	-1.02E-03	1.00E-10	1.02E-03	1.04E-06	4.00E-10	2	1.02E-03	0.001019902	9.80486E-08	1.0
1.00E-05	4.00E-05	5.00E-05	1	-1.00E-04	4.00E-10	1.00E-04	1.00E-08	1.60E-09	2	9.17E-05	9.58258E-05	4.17424E-06	41.7
1.00E-05	8.00E-05	5.00E-05	1	-1.40E-04	8.00E-10	1.40E-04	1.96E-08	3.20E-09	2	1.28E-04	0.000134031	5.96876E-06	59.7
1.00E-05	4.00E-05	1.00E-06	1	-5.10E-05	4.00E-10	5.10E-05	2.60E-09	1.60E-09	2	3.16E-05	4.13193E-05	9.68071E-06	96.8
1.00E-05	8.00E-05	1.00E-06	1	-9.10E-05	8.00E-10	9.10E-05	8.28E-09	3.20E-09	2	7.13E-05	8.11406E-05	9.85943E-06	98.6
1.00E-05	1.00E-04	1.00E-06	1	-1.11E-04	1.00E-09	1.11E-04	1.23E-08	4.00E-09	2	9.12E-05	0.00010111	9.89024E-06	98.9
4.00E-05	1.00E-05	1.00E-06	1	-5.10E-05	4.00E-10	5.10E-05	2.60E-09	1.60E-09	2	3.16E-05	4.13193E-05	9.68071E-06	24.2
8.00E-05	1.00E-05	1.00E-06	1	-9.10E-05	8.00E-10	9.10E-05	8.28E-09	3.20E-09	2	7.13E-05	8.11406E-05	9.85943E-06	12.3
4.00E-05	1.00E-05	1.00E-06	1	-5.10E-05	4.00E-10	5.10E-05	2.60E-09	1.60E-09	2	3.16E-05	4.13193E-05	9.68071E-06	24.2
8.00E-05	1.00E-05	1.00E-06	1	-9.10E-05	8.00E-10	9.10E-05	8.28E-09	3.20E-09	2	7.13E-05	8.11406E-05	9.85943E-06	12.3
1.00E-04	1.00E-05	1.00E-06	1	-1.11E-04	1.00E-09	1.11E-04	1.23E-08	4.00E-09	2	9.12E-05	0.00010111	9.89024E-06	9.9

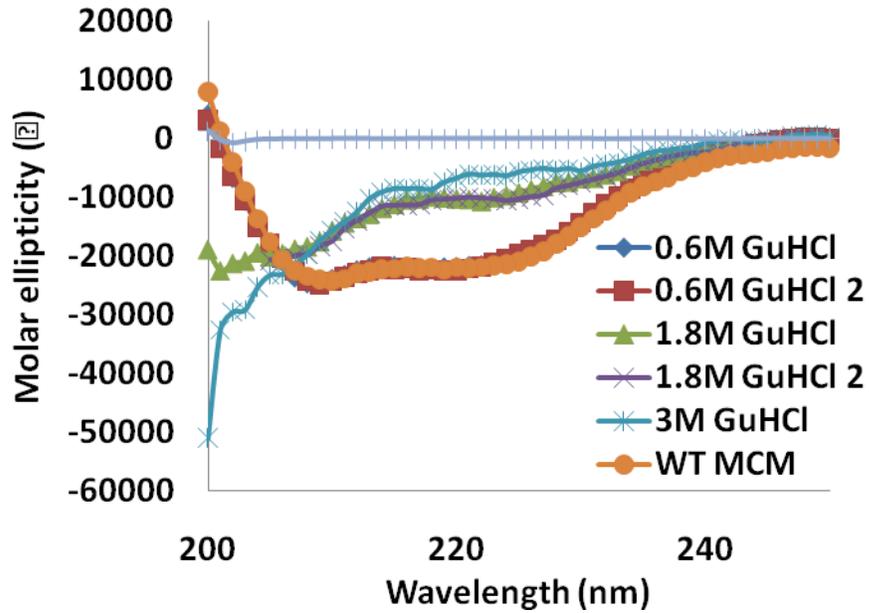


Figure 9-12 CD spectra of wild type MCM and increasing amount of guanidinium hydrochloride (GuHCl).

MCM protein seems to be folded properly up to 0.6 M Guanidinium hydrochloride, however after 1.8 M GuHCl is added, the structure seems to unfold.

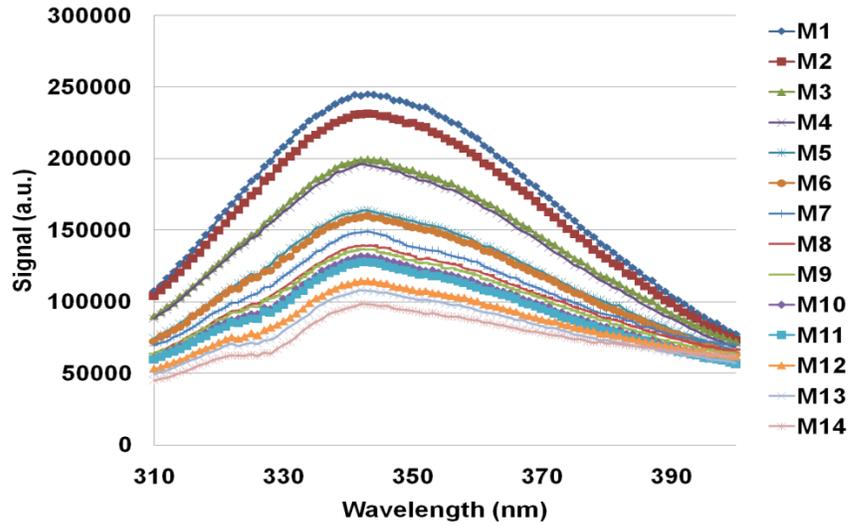


Figure 9-13 Tryptophan fluorescence signal of MCM protein over time using 5 nm slit width.

The measurement of tryptophan fluorescence signal of MCM protein (100 nM) was taken every 10 minutes at 25°C with 5 nm slit width. M1 is measurement immediately after the insertion of cuvette in to the machine. The cuvette was left at 25°C in the cuvette holder for 10 minutes, and measurement was taken, and called M2, and same was repeated up to M14.

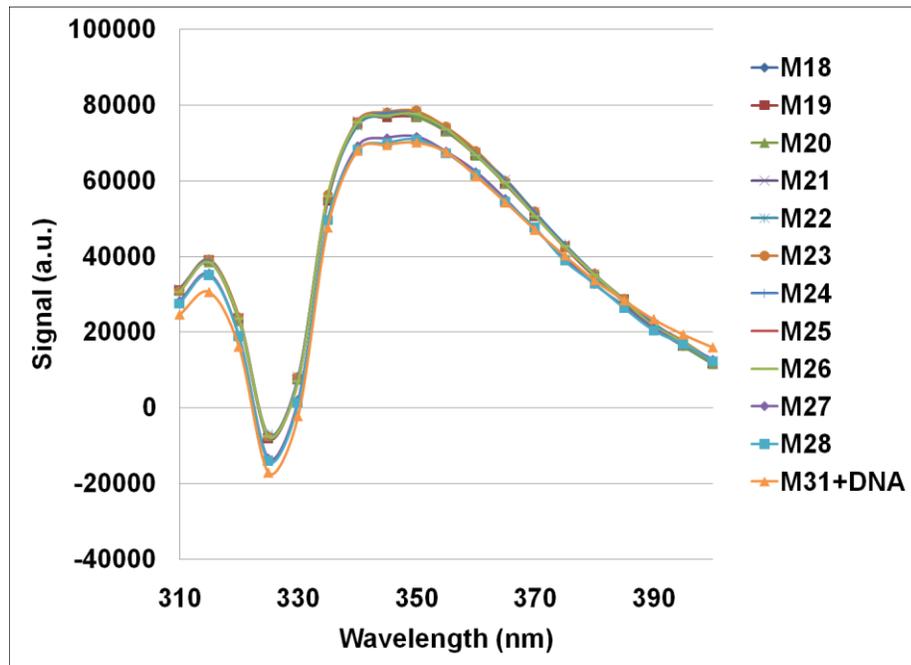


Figure 9-14 Tryptophan fluorescence signal of MCM protein over time using 4 nm slit width.

The measurements of MCM protein (100 nM) were taken every 10 minutes at 25°C with 4 nm slit width. M18 is the measurement of sample immediately after cuvette was placed in the machine. The cuvette was left untouched at 25°C in the cuvette holder for 10 minutes and the fluorescence was measured. M31 +DNA is the sample containing 1 µg of ssDNA (N190).

BIBLIOGRAPHY

1. Grabowski, B. and Kelman, Z. (2003) Archaeal DNA replication: Eukaryal Proteins in a Bacterial Context. *Annu Rev Microbiol*, **57**, 487-516.
2. Robinson, N.P. and Bell, S.D. (2005) Origins of DNA replication in the three domains of life. *Febs J*, **272**, 3757-3766.
3. Jacob, F., Brenner, S. and Cuzin, F. (1963) On the regulation of DNA replication in bacteria. *Cold Spring Harbor Symposium on Quantitative Biology*, **28**, 329-348.
4. Stryer, L. (1995) *Biochemistry* 4th ed. W.H. Freeman and Company, New York.
5. Weaver, R. (2002) *Molecular Biology*. 2nd ed. McGraw-Hill.
6. Woese, C.R. and Fox, G.E. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, **74**, 5088-5090.
7. Woese, C.R., Kandler, O. and Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 4576-4579.
8. Olsen, G.J. and Woese, C.R. (1997) Archaeal genomics: an overview. *Cell*, **89**, 991-994.
9. Kates, M., Kushner, D. and Matheson, A.T. (1993) *The Biochemistry of archaea (archaeobacteria)*. Elsevier, Amsterdam ; New York.

10. Barns, S.M., Delwiche, C.F., Palmer, J.D. and Pace, N.R. (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences [see comments]. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 9188-9193.
11. Huber, H., Hohn, M.J., Rachel, R., Fuchs, T., Wimmer, V.C. and Stetter, K.O. (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature*, **417**, 63-67.
12. Zeikus, J.G. and Wolfe, R.S. (1972) *Methanobacterium thermoautotrophicus* sp. n., an anaerobic, autotrophic, extreme thermophile. *J Bacteriol*, **109**, 707-715.
13. Smith, D.R., Doucette-Stamm, L.A., Deloughery, C., Lee, H., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K. *et al.* (1997) Complete genome sequence of *Methanobacterium thermoautotrophicum* Δ H: functional analysis and comparative genomics. *J Bacteriol*, **179**, 7135-7155.
14. Majernik, A.I., Lundgren, M., McDermott, P., Bernander, R. and Chong, J.P. (2005) DNA content and nucleoid distribution in *Methanothermobacter thermautotrophicus*. *J Bacteriol*, **187**, 1856-1858.
15. Grant, W.D. (2001) *Halobacterium*. 2nd ed. Bergey's Manual Trust, Baltimore.
16. Ng, W.V., Kennedy, S.P., Mahairas, G.G., Berquist, B., Pan, M., Shukla, H.D., Lasky, S.R., Baliga, N.S., Thorsson, V., Sbrogna, J. *et al.* (2000) Genome sequence of *Halobacterium* species NRC-1. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 12176-121781.
17. Zhang, R. and Zhang, C.T. (2003) Multiple replication origins of the archaeon *Halobacterium* species NRC-1. *Biochem Biophys Res Commun*, **302**, 728-734.

18. Berquist, B.R. and DasSarma, S. (2003) An archaeal chromosomal autonomously replicating sequence element from an extreme halophile, *Halobacterium* sp. strain NRC-1. *J Bacteriol*, **185**, 5959-5966.
19. Berquist, B.R., DasSarma, P. and DasSarma, S. (2007) Essential and non-essential DNA replication genes in the model halophilic Archaeon, *Halobacterium* sp. NRC-1. *BMC Genet*, **8**, 31.
20. Speck, C. and Messer, W. (2001) Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. *The EMBO journal*, **20**, 1469-1476.
21. Boulikas, T. (1996) Common structural features of replication origins in all life forms. *J Cell Biochem*, **60**, 297-316.
22. Nasheuer, H.P., Smith, R., Bauerschmidt, C., Grosse, F. and Weisshart, K. (2002) Initiation of eukaryotic DNA replication: regulation and mechanisms. *Prog Nucleic Acid Res Mol Biol*, **72**, 41-94.
23. Blow, J.J. and Dutta, A. (2005) Preventing re-replication of chromosomal DNA. *Nature reviews*, **6**, 476-486.
24. Zhang, R. and Zhang, C.T. (2005) Identification of replication origins in archaeal genomes based on the Z-curve method. *Archaea*, **1**, 335-346.
25. Myllykallio, H., Lopez, P., Lopez-Garcia, P., Heilig, R., Saurin, W., Zivanovic, Y., Philippe, H. and Forterre, P. (2000) Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science*, **288**, 2212-2215.

26. Robinson, N.P., Dionne, I., Lundgren, M., Marsh, V.L., Bernander, R. and Bell, S.D. (2004) Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. *Cell*, **116**, 25-38.
27. Lundgren, M., Andersson, A., Chen, L., Nilsson, P. and Bernander, R. (2004) Three replication origins in *Sulfolobus* species: Synchronous initiation of chromosome replication and asynchronous termination. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 7046-7051.
28. Maisnier-Patin, S., Malandrin, L., Birkeland, N.K. and Bernander, R. (2002) Chromosome replication patterns in the hyperthermophilic euryarchaea *Archaeoglobus fulgidus* and *Methanocaldococcus (Methanococcus) jannaschii*. *Mol Microbiol*, **45**, 1443-1450.
29. Matsunaga, F., Forterre, P., Ishino, Y. and Myllykallio, H. (2001) In vivo interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 11152-11157.
30. Majernik, A.I. and Chong, J.P. (2008) A conserved mechanism for replication origin recognition and binding in archaea. *Biochem J*, **409**, 511-518.
31. Norais, C., Hawkins, M., Hartman, A.L., Eisen, J.A., Myllykallio, H. and Allers, T. (2007) Genetic and physical mapping of DNA replication origins in *Haloferax volcanii*. *PLoS Genet*, **3**, e77.
32. Bell, S.P., Mitchell, J., Leber, J., Kobayashi, R. and Stillman, B. (1995) The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing. *Cell*, **83**, 563-568.

33. Singleton, M.R., Morales, R., Grainge, I., Cook, N., Isupov, M.N. and Wigley, D.B. (2004) Conformational changes induced by nucleotide binding in Cdc6/ORC from *Aeropyrum pernix*. *J Mol Biol*, **343**, 547-557.
34. Capaldi, S.A. and Berger, J.M. (2004) Biochemical characterization of Cdc6/Orc1 binding to the replication origin of the euryarchaeon *Methanothermobacter thermoautotrophicus*. *Nucleic acids research*, **32**, 4821-4832.
35. Kasiviswanathan, R., Shin, J.H. and Kelman, Z. (2005) Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties. *Nucleic acids research*, **33**, 4940-4950.
36. Dueber, E.L., Corn, J.E., Bell, S.D. and Berger, J.M. (2007) Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. *Science*, **317**, 1210-1213.
37. Liu, J., Smith, C.L., DeRyckere, D., DeAngelis, K., Martin, G.S. and Berger, J.M. (2000) Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Molecular cell*, **6**, 637-648.
38. Gaudier, M., Schuwirth, B.S., Westcott, S.L. and Wigley, D.B. (2007) Structural basis of DNA replication origin recognition by an ORC protein. *Science*, **317**, 1213-1126.
39. Maine, G.T., Sinha, P. and Tye, B.K. (1984) Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics*, **106**, 365-385.
40. Forsburg, S.L. (2004) Eukaryotic MCM Proteins: Beyond Replication Initiation. *Microbiol Mol Biol Rev*, **68**, 109-131.

41. Tye, B.K. and Sawyer, S.L. (2000) The Hexameric Eukaryotic MCM Helicase: Building Symmetry from Nonidentical Parts. *The Journal of biological chemistry*, **275**, 34833-34836.
42. Forsburg, S.L. (2008) The MCM helicase: linking checkpoints to the replication fork. *Biochemical Society transactions*, **36**, 114-119.
43. Tye, B.K. (1999) MCM proteins in DNA replication. *Annual review of biochemistry*, **68**, 649-686.
44. Shin, J.H., Jiang, Y., Grabowski, B., Hurwitz, J. and Kelman, Z. (2003) Substrate requirements for duplex DNA translocation by the eukaryal and archaeal minichromosome maintenance helicases. *The Journal of biological chemistry*, **278**, 49053-49062.
45. Bochman, M.L. and Schwacha, A. (2008) The Mcm2-7 complex has *in vitro* helicase activity. *Molecular cell*, **31**, 287-293.
46. Ishimi, Y. (1997) A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex [published erratum appears in J Biol Chem 1998 Sep 4;273(36):23616]. *The Journal of biological chemistry*, **272**, 24508-24513.
47. Maiorano, D., Lutzmann, M. and Mechali, M. (2006) MCM proteins and DNA replication. *Curr Opin Cell Biol*, **18**, 130-136.
48. Ibarra, A., Schwob, E. and Mendez, J. (2008) Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 8956-8961.

49. Kearsey, S.E. and Labib, K. (1998) MCM proteins: evolution, properties, and role in DNA replication. *Biochim Biophys Acta*, **1398**, 113-136.
50. Kelman, L.M. and Kelman, Z. (2003) Archaea: an archetype for replication initiation studies? *Mol Microbiol*, **48**, 605-615.
51. Enemark, E.J. and Joshua-Tor, L. (2008) On helicases and other motor proteins. *Curr Opin Struct Biol*, **18**, 243-257.
52. Matsunaga, F., Glatigny, A., Mucchielli-Giorgi, M.H., Agier, N., Delacroix, H., Marisa, L., Durosay, P., Ishino, Y., Aggerbeck, L. and Forterre, P. (2007) Genomewide and biochemical analyses of DNA-binding activity of Cdc6/Orc1 and Mcm proteins in *Pyrococcus* sp. *Nucleic acids research*, **35**, 3214-3222.
53. Duggin, I.G., McCallum, S.A. and Bell, S.D. (2008) Chromosome replication dynamics in the archaeon *Sulfolobus acidocaldarius*. *Proceedings of the National Academy of Sciences of the United States of America*.
54. Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D. *et al.* (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii* [see comments]. *Science*, **273**, 1058-1073.
55. Marinsek, N., Barry, E.R., Makarova, K.S., Dionne, I., Koonin, E.V. and Bell, S.D. (2006) GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep*, **7**, 539-545.
56. Haugland, G.T., Rollor, C.R., Birkeland, N.K. and Kelman, Z. (2009) Biochemical characterization of the minichromosome maintenance protein from the archaeon *Thermoplasma acidophilum*. *Extremophiles*, **13**, 81-88.

57. Aravind, L. and Koonin, E.V. (1999) DNA-binding proteins and evolution of transcription regulation in the archaea. *Nucleic acids research*, **27**, 4658-4670.
58. Singleton, M.R., Dillingham, M.S. and Wigley, D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annual review of biochemistry*, **76**, 23-50.
59. Fletcher, R.J., Bishop, B.E., Leon, R.P., Sclafani, R.A., Ogata, C.M. and Chen, X.S. (2003) The structure and Function of MCM from archaeal *M. thermoautotrophicum*. *Nature Struct Biol*, **10**, 160-167.
60. Liu, W., Pucci, B., Rossi, M., Pisani, F.M. and Ladenstein, R. (2008) Structural analysis of the *Sulfolobus solfataricus* MCM protein N-terminal domain. *Nucleic acids research*, **36**, 3235-3243.
61. Poplawski, A., Grabowski, B., Long, S.E. and Kelman, Z. (2001) The zinc finger domain of the archaeal minichromosome maintenance protein is required for helicase activity. *The Journal of biological chemistry*, **276**, 49371-49377.
62. Neuwald, A.F., Aravind, L., Spouge, J.L. and Koonin, E.V. (1999) AAA⁺: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res*, **9**, 27-43.
63. Iyer, L.M., Leipe, D.D., Koonin, E.V. and Aravind, L. (2004) Evolutionary history and higher order classification of AAA⁺ ATPases. *J Struct Biol*, **146**, 11-31.
64. Davey, M.J., Jeruzalmi, D., Kuriyan, J. and O'Donnell, M. (2002) Motors and switches: AAA⁺ machines within the replisome. *Nature reviews*, **3**, 826-835.

65. Lee, D.G. and Bell, S.P. (2000) ATPase switches controlling DNA replication initiation. *Curr Opin Cell Biol*, **12**, 280-285.
66. Ogura, T. and Wilkinson, A.J. (2001) AAA+ superfamily ATPases: common structure--diverse function. *Genes Cells*, **6**, 575-597.
67. Erzberger, J.P. and Berger, J.M. (2006) Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu Rev Biophys Biomol Struct*, **35**, 93-114.
68. Hanson, P.I. and Whiteheart, S.W. (2005) AAA+ proteins: have engine, will work. *Nature reviews*, **6**, 519-529.
69. Belfort, M. and Roberts, R.J. (1997) Homing endonucleases: keeping the house in order. *Nucleic acids research*, **25**, 3379-3388.
70. Sakakibara, N., Han, M., Rollor, C.R., Gilson, R.C., Busch, C., Heo, G. and Kelman, Z. (2008) Cloning, Purification, and Partial Characterization of the *Halobacterium* sp. NRC-1 Minichromosome Maintenance (MCM) Helicase. *The Open Microbiology Journal*, **2**, 13-17.
71. Kelman, Z., Lee, J.K. and Hurwitz, J. (1999) The single minichromosome maintenance protein of *Methanobacterium thermoautotrophicum* Δ H contains DNA helicase activity. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 14783-14788.
72. Chong, J.P., Hayashi, M.K., Simon, M.N., Xu, R.M. and Stillman, B. (2000) A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 1530-1535.

73. Shechter, D.F., Ying, C.Y. and Gautier, J. (2000) The intrinsic DNA helicase activity of *Methanobacterium thermoautotrophicum* Δ H minichromosome maintenance protein. *The Journal of biological chemistry*, **275**, 15049-15059.
74. Carpentieri, F., De Felice, M., De Falco, M., Rossi, M. and Pisani, F.M. (2002) Physical and functional interaction between the MCM-like DNA helicase and the single-stranded DNA binding protein from the crenarchaeon *Sulfolobus solfataricus*. *The Journal of biological chemistry*, **277**, 12118-12127.
75. Haugland, G.T., Shin, J.H., Birkeland, N.K. and Kelman, Z. (2006) Stimulation of MCM helicase activity by a Cdc6 protein in the archaeon *Thermoplasma acidophilum*. *Nucleic acids research*, **34**, 6337-6344.
76. Grainge, I., Scaife, S. and Wigley, D.B. (2003) Biochemical analysis of components of the pre-replication complex of *Archaeoglobus fulgidus*. *Nucleic acids research*, **31**, 4888-4898.
77. Shin, J.-H., Mauro, R., Melamud, E. and Kasiviswanathan, R. (2006) Cloning and partial characterization of the *Methanococcoides burtonii* minichromosome maintenance (MCM) helicase. *BIOS*, **77**, 37-41.
78. Yoshimochi, T., Fujikane, R., Kawanami, M., Matsunaga, F. and Ishino, Y. (2008) The GINS complex from *Pyrococcus furiosus* stimulates the MCM helicase activity. *The Journal of biological chemistry*, **283**, 1601-1609.
79. Bae, B., Chen, Y.H., Costa, A., Onesti, S., Brunzelle, J.S., Lin, Y., Cann, I. and Nair, S.K. (2009) Crystal Structure of an Archaeal MCM Homolog Provides Insights into the Architecture of the Replicative Helicase. *Structure*. In press.

80. Liu, B. and Alberts, B.M. (1995) Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex. *Science*, **267**, 1131-1137.
81. Kasiviswanathan, R., Shin, J.H., Melamud, E. and Kelman, Z. (2004) Biochemical Characterization of the *Methanothermobacter thermautotrophicus* Minichromosome Maintenance (MCM) Helicase N-terminal Domains. *The Journal of biological chemistry*, **279**, 28358-28366.
82. Brewster, A.S., Wang, G., Yu, X., Greenleaf, W.B., Carazo, J.M., Tjajadia, M., Klein, M.G. and Chen, X.S. (2008) Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 20191-20196.
83. Yu, X., VanLoock, M.S., Poplawski, A., Kelman, Z., Xiang, T., Tye, B.K. and Egelman, E.H. (2002) The *Methanobacterium thermoautotrophicum* MCM protein can form heptameric rings. *EMBO Rep*, **3**, 792-797.
84. Pape, T., Meka, H., Chen, S., Vicentini, G., van Heel, M. and Onesti, S. (2003) Hexameric ring structure of the full-length archaeal MCM protein complex. *EMBO Rep*, **4**, 1079-1083.
85. Chen, Y.J., Yu, X., Kasiviswanathan, R., Shin, J.H., Kelman, Z. and Egelman, E.H. (2005) Structural Polymorphism of *Methanothermobacter thermautotrophicus* MCM. *J Mol Biol*, **346**, 389-394.
86. Gomez-Llorente, Y., Fletcher, R.J., Chen, X.S., Carazo, J.M. and Martin, C.S. (2005) Polymorphism and Double Hexamer Structure in the Archaeal Minichromosome Maintenance (MCM) Helicase from *Methanobacterium thermoautotrophicum*. *The Journal of biological chemistry*, **280**, 40909-40915.

87. Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A. and Onesti, S. (2006) Structural basis of the *Methanothermobacter thermautotrophicus* MCM helicase activity. *Nucleic acids research*, **34**, 5829-5838.
88. Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A. and Onesti, S. (2006) Structural studies of the archaeal MCM complex in different functional states. *J Struct Biol*, **156**, 210-219.
89. Barry, E.R., McGeoch, A.T., Kelman, Z. and Bell, S.D. (2007) Archaeal MCM has separable processivity, substrate choice and helicase domains. *Nucleic acids research*, **35**, 988-998.
90. McGeoch, A.T., Trakselis, M.A., Laskey, R.A. and Bell, S.D. (2005) Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nature structural & molecular biology*, **12**, 756-762.
91. Rothenberg, E., Trakselis, M.A., Bell, S.D. and Ha, T. (2007) MCM forked substrate specificity involves dynamic interaction with the 5'-tail. *The Journal of biological chemistry*, **282**, 34229-34234.
92. Pucci, B., De Felice, M., Rossi, M., Onesti, S. and Pisani, F.M. (2004) Amino Acids of the *Sulfolobus solfataricus* Mini-chromosome Maintenance-like DNA Helicase Involved in DNA Binding/Remodeling. *The Journal of biological chemistry*, **279**, 49222-49228.
93. Shin, J.H., Santangelo, T.J., Xie, Y., Reeve, J.N. and Kelman, Z. (2007) Archaeal minichromosome maintenance (MCM) helicase can unwind DNA bound by archaeal histones and transcription factors. *The Journal of biological chemistry*, **282**, 4908-4915.

94. Berthon, J., Cortez, D. and Forterre, P. (2008) Genomic context analysis in Archaea suggests previously unrecognized links between DNA replication and translation. *Genome biology*, **9**, R71.
95. Haugland, G.T., Sakakibara, N., Pey, A.L., Rollor, C.R., Birkeland, N.K. and Kelman, Z. (2008) *Thermoplasma acidophilum* Cdc6 protein stimulates MCM helicase activity by regulating its ATPase activity. *Nucleic acids research*, **36**, 5602-5609.
96. Singleton, M.R., Sawaya, M.R., Ellenberger, T. and Wigley, D.B. (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, **101**, 589-600.
97. Enemark, E.J., Chen, G., Vaughn, D.E., Stenlund, A. and Joshua-Tor, L. (2000) Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus. *Molecular cell*, **6**, 149-158.
98. Martin, A., Baker, T.A. and Sauer, R.T. (2005) Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. *Nature*, **437**, 1115-1120.
99. Gai, D., Zhao, R., Li, D., Finkielstein, C.V. and Chen, X.S. (2004) Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell*, **119**, 47-60.
100. Moreau, M.J., McGeoch, A.T., Lowe, A.R., Itzhaki, L.S. and Bell, S.D. (2007) ATPase site architecture and helicase mechanism of an archaeal MCM. *Molecular cell*, **28**, 304-314.
101. Pucci, B., De Felice, M., Rocco, M., Esposito, F., De Falco, M., Esposito, L., Rossi, M. and Pisani, F.M. (2007) Modular organization of the Sulfolobus

- solfataricus mini-Chromosome maintenance protein. *The Journal of biological chemistry*, **282**, 12574-12582.
102. Jenkinson, E.R. and Chong, J.P. (2006) Minichromosome maintenance helicase activity is controlled by N- and C-terminal motifs and requires the ATPase domain helix-2 insert. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 7613-7618.
103. Sakakibara, N., Kasiviswanathan, R., Melamud, E., Han, M., Schwarz, F.P. and Kelman, Z. (2008) Coupling of DNA binding and helicase activity is mediated by a conserved loop in the MCM protein. *Nucleic acids research*, **36**, 1309-1320.
104. Shin, J.H., Grabowski, B., Kasiviswanathan, R., Bell, S.D. and Kelman, Z. (2003) Regulation of minichromosome maintenance helicase activity by Cdc6. *The Journal of biological chemistry*, **278**, 38059-38067.
105. De Felice, M., Esposito, L., Pucci, B., Carpentieri, F., De Falco, M., Rossi, M. and Pisani, F.M. (2003) Biochemical characterization of a CDC6-like protein from the crenarchaeon *Sulfolobus solfataricus*. *The Journal of biological chemistry*, **278**, 46424-46431.
106. De Felice, M., Esposito, L., Pucci, B., De Falco, M., Manco, G., Rossi, M. and Pisani, F.M. (2004) Modular organization of a Cdc6-like protein from the crenarchaeon *Sulfolobus solfataricus*. *Biochem J*, **381**, 645-653.
107. Costa, A., van Duinen, G., Medagli, B., Chong, J., Sakakibara, N., Kelman, Z., Nair, S.K., Patwardhan, A. and Onesti, S. (2008) Cryo-electron microscopy reveals a novel DNA-binding site on the MCM helicase. *The EMBO journal*, **27**, 2250-2258.

108. Costa, A. and Onesti, S. (2008) The MCM complex: (just) a replicative helicase? *Biochemical Society transactions*, **36**, 136-140.
109. Hardy, C.F., Dryga, O., Seematter, S., Pahl, P.M. and Sclafani, R.A. (1997) *mcm5/cdc46-bob1* bypasses the requirement for the S phase activator Cdc7p. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 3151-3155.
110. Fletcher, R.J. and Chen, X.S. (2006) Biochemical activities of the *BOBI* mutant in *Methanobacterium thermoautotrophicum* MCM. *Biochemistry*, **45**, 462-467.
111. Barry, E.R. and Bell, S.D. (2006) DNA replication in the archaea. *Microbiol Mol Biol Rev*, **70**, 876-887.
112. Fletcher, R.J., Shen, J., Gomez-Llorente, Y., Martin, C.S., Carazo, J.M. and Chen, X.S. (2005) Double hexamer disruption and biochemical activities of *Methanobacterium thermoautotrophicum* MCM. *The Journal of biological chemistry*, **280**, 42405-42410.
113. Shin, J.H., Heo, G.Y. and Kelman, Z. (2009) The *Methanothermobacter thermautotrophicus* MCM Helicase Is Active as a Hexameric Ring. *The Journal of biological chemistry*, **284**, 540-546.
114. Shin, J.H., Heo, G.Y. and Kelman, Z. (2008) The *Methanothermobacter thermautotrophicus* Cdc6-2 protein, the putative helicase loader, dissociates the minichromosome maintenance helicase. *J Bacteriol*, **190**, 4091-4094.
115. Kelman, Z. and Hurwitz, J. (2003) Structural lessons in DNA replication from the third domain of life. *Nat Struct Biol*, **10**, 148-150.

116. Kelman, Z. and White, M.F. (2005) Archaeal DNA replication and repair. *Curr Opin Microbiol*, **8**, 669-676.
117. Kasiviswanathan, R., Shin, J.H. and Kelman, Z. (2006) DNA Binding by the *Methanothermobacter thermautotrophicus* Cdc6 Protein Is Inhibited by the Minichromosome Maintenance Helicase. *J Bacteriol*, **188**, 4577-4580.
118. Shin, J.H. and Kelman, Z. (2006) The replicative helicases of bacteria, archaea and eukarya can unwind RNA-DNA hybrid substrates. *The Journal of biological chemistry*, **281**, 26914-26921.
119. McEntee, K., Weinstock, G.M. and Lehman, I.R. (1980) recA protein-catalyzed strand assimilation: stimulation by *Escherichia coli* single-stranded DNA-binding protein. *Proceedings of the National Academy of Sciences of the United States of America*, **77**, 857-861.
120. Heyduk, T. and Lee, J.C. (1990) Application of fluorescence energy transfer and polarization to monitor *Escherichia coli* cAMP receptor protein and lac promoter interaction. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 1744-1748.
121. Licata, V.J. and Wowor, A.J. (2008) Applications of Fluorescence Anisotropy to the Study of Protein-DNA Interactions. *Methods in cell biology*, **84**, 243-262.
122. Kirchhoff, W.H. (1993) Exam: A Two-State Thermodynamic Analysis Program. *NIST Tech. Note*, **1401**, 1-103.
123. Kaplan, D.L., Davey, M.J. and O'Donnell, M. (2003) Mcm4,6,7 Uses a "Pump in Ring" Mechanism to Unwind DNA by Steric Exclusion and Actively Translocate along a Duplex. *The Journal of biological chemistry*, **278**, 49171-49182.

124. Manly, S.P., Matthews, K.S. and Sturtevant, J.M. (1985) Thermal denaturation of the core protein of lac repressor. *Biochemistry*, **24**, 3842-3846.
125. Grabowski, B. and Kelman, Z. (2001) Autophosphorylation of the archaeal Cdc6 homologues is regulated by DNA. *J Bacteriol*, **183**, 5459-5464.
126. DeLano, W.L. (2002). DeLano Scientific, San Carlos, CA, USA. .
127. Fletcher, R.J., Shen, J., Holden, L.G. and Chen, X.S. (2008) Identification of Amino Acids Important for the Biochemical Activity of *Methanothermobacter thermautotrophicus* MCM. *Biochemistry*.
128. Duggin, I.G. and Bell, S.D. (2006) The chromosome replication machinery of the archaeon *Sulfolobus solfataricus*. *The Journal of biological chemistry*, **281**, 15029-15032.
129. Schwarz, F.P. (1988) Interaction of cytidine 3'-monophosphate and uridine 3'-monophosphate with ribonuclease a at the denaturation temperature. *Biochemistry*, **27**, 8429-8436.
130. Wittinghofer, A. (1997) Signaling mechanistics: aluminum fluoride for molecule of the year. *Curr Biol*, **7**, R682-685.
131. Fisher, A.J., Smith, C.A., Thoden, J.B., Smith, R., Sutoh, K., Holden, H.M. and Rayment, I. (1995) X-ray structures of the myosin motor domain of *Dictyostelium discoideum* complexed with MgADP-BeFx and MgADP-AlF₄. *Biochemistry*, **34**, 8960-8972.
132. Enemark, E.J. and Joshua-Tor, L. (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature*, **442**, 270-275.

133. Snyder, M., He, W. and Zhang, J.J. (2005) The DNA replication factor MCM5 is essential for Stat1-mediated transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 14539-14544.
134. Ishimi, Y., Ichinose, S., Omori, A., Sato, K. and Kimura, H. (1996) Binding of human minichromosome maintenance proteins with histone H3. *The Journal of biological chemistry*, **271**, 24115-24122.
135. Burke, T.W., Cook, J.G., Asano, M. and Nevins, J.R. (2001) Replication factors MCM2 and ORC1 interact with the histone acetyltransferase HBO1. *The Journal of biological chemistry*, **276**, 15397-15408.
136. Pietrokovski, S. (1994) Conserved sequence features of inteins (protein introns) and their use in identifying new inteins and related proteins. *Protein Sci*, **3**, 2340-2350.
137. Pietrokovski, S. (1998) Modular organization of inteins and C-terminal autocatalytic domains. *Protein Sci*, **7**, 64-71.
138. Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO journal*, **1**, 945-951.
139. Kennedy, S.P., Ng, W.V., Salzberg, S.L., Hood, L. and DasSarma, S. (2001) Understanding the adaptation of *Halobacterium* species NRC-1 to its extreme environment through computational analysis of its genome sequence. *Genome Res*, **11**, 1641-1650.

140. Saleh, L. and Perler, F.B. (2006) Protein splicing in cis and in trans. *Chem Rec*, **6**, 183-193.
141. Davey, M.J. and O'Donnell, M. (2003) Replicative helicase loaders: ring breakers and ring makers. *Curr Biol*, **13**, R594-R596.
142. Matsunaga, F., Norais, C., Forterre, P. and Myllykallio, H. (2003) Identification of short 'eukaryotic' Okazaki fragments synthesized from a prokaryotic replication origin. *EMBO Reports*, **4**, 154-158.
143. Takahashi, T.S., Wigley, D.B. and Walter, J.C. (2005) Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase. *Trends Biochem Sci*, **30**, 437-444.
144. van Holde, K.E., Johnson, W.C. and Ho, P.S. (1998) *Principles of Physical Biochemistry*. Prentice-Hall, Inc., Upper Saddle River, New Jersey.
145. Leatherbarrow, R. (2001) *GraFit Version 5*. Erithacus Software Limited, Horley, U.K.
146. Bruylants, G., Wouters, J. and Michaux, C. (2005) Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design. *Current medicinal chemistry*, **12**, 2011-2020.
147. Sanchez, R.M., Alegret, M., Adzet, T., Merlos, M. and Laguna, J.C. (1992) Differential inhibition of long-chain acyl-CoA hydrolases by hypolipidemic drugs in vitro. *Biochemical pharmacology*, **43**, 639-644.
148. Gomez, J., Hilser, V.J., Xie, D. and Freire, E. (1995) The heat capacity of proteins. *Proteins*, **22**, 404-412.

149. Lakowicz, J.R. (1999) *Principles of Fluorescence Spectroscopy*. 2nd ed. Klumer Academic Plenum Publishers, New York.