Biochemical studies on the pre-replication complex of Archaea

Neli Ivanova Atanassova

A thesis submitted for the degree of Doctor of Philosophy

at

The University College of London

October 2005

Cancer Research UK London Research Institute,

Clare Hall Laboratories,

South Mimms, Hertfordshire, EN6 3LD

and

University College London,

Gower Street, London, WC1 6BT

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGMENTS

I am indebted to many people for the support they have given me during my time in Clare Hall. First and foremost, I would like to thank my supervisor Dr. Dale Wigley for giving me the opportunity to work in his lab, for his valuable comments and constructive criticism, and for all the support and advice I have received throughout my years as a graduate student.

I am very grateful to Ian Grainge for his patience in teaching me pretty much all I know about biochemistry, his enthusiasm and optimism always made the experiments fun at times they might have been unbearable. I would also like to thank Anja Seybert and Martin Gaudier for their help and advice with the biochemical calculations and Martin Singleton for his help with the numerous computing problems that I have faced.

I would like to thank all past and present members of Molecular Enzymology lab for making my time in the lab enjoyable and wish them good luck in the future.

It is my great pleasure to thank all my friends in Clare Hall for the funny, precious and unforgettable moments we share together. Elena, Munah, Nicky, Alessia and Beatriz, thank you for your friendship and being there for me when I needed you most.

Special thanks to Kyriakos for his unlimited patience, love and understanding, and for waiting for me all those long years.

Finally, I would like to thank my family for their perpetual love and support and most of all to my Mum who always believed in me and to whom I dedicate this thesis.

ABSTRACT

Several proteins play a crucial role in the formation of the pre-replication complex (pre-RC) in eukaryotes: origin recognition complex (ORC), CDC6 and minichromosome maintenance proteins (MCM). In yeast, these proteins associate with DNA in a sequence-specific manner during the cell cycle. Archaea, the third kingdom of life, are believed to replicate DNA in a eukaryote-like fashion. Most archaeal species encode single MCM and two CDC6/ORC proteins. The reduced number of replication proteins suggests that archaea may provide a simplified model for assembly of the machinery required for initiation of eukaryotic DNA replication.

This thesis describes a biochemical analysis of the proteins involved in the formation of pre-replication complex (pre-RC) in Archaea. The yeast two-hybrid system was employed to identify protein-protein interactions within the pre-replication complex of two archaeal species *Archaeoglobus fulgidus* and *Aeropyrum pernix*. A library of truncated MCM proteins was used to identify a region that appears to be responsible for hexamerisation of MCM.

The pre-RC proteins of archaea belong to the AAA+ family of ATPases. AAA+ proteins function as ATP-driven conformational switches that can mediate the assembly, remodelling, or disassembly of protein complexes. Thus, the assembly of the replication initiation complex may involve an ordered sequence of ATP coupled conformational changes. A variety of biochemical techniques were used to investigate the relationship between ATP hydrolysis, DNA binding and protein interactions within the pre-RC of Archaea.

ABBREVIATIONS

Ab	Antibody
AD	Activation domain
ADP	Adenosine diphosphate
A.fulgidus or *Af*	*Archaeoglobus fulgidus*
Amp	Ampicillin
A.pernix or *Ape*	*Aeropyrum pernix*
ATP	Adenosine triphosphate
BD	Binding domain
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
CDC6	Cell division cycle 6 homologue
Da	Daltons
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
DTT	Di-thiothreitol
ECL	Enhanced chemiluminescence
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
GAL	Galactose

HA	Haemagglutinin
HEPES	N- (2-Hydroxyethyl) piperazine-N'- (2-ethanesulfonic acid)
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kan	Kanamycin
LB	Luria broth
LiAc/TE/PEG	Lithium acetate/Tris-EDTA/Polyethylene glycol
MCM	Minichromosome maintenance
M.thermoautotrohicum or *Mth*	*Methanobacterium thermoautotrophicum*
NCBI	National Centre for Biotechnology Information
NTP	Nucleoside triphosphate
OD	Optical density
ORB	Origin binding box
ORC	Origin recognition complex
ORF	Open reading frame
P.aerophilum or *Pae*	*Pyrobaculum aerophilum*
PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
PMSF	Phenylmethanesulfonyl fluoride
PNK	Polynucleotide kinase
RNA	Ribonucleic acid
rpm	Revolutions per minute

S.cerevisiae	*Saccharomyces cerevisiae*
SD/A, H, L, W	Synthetic dropout media lacking adenine, histidine, leucine and tryptophan
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
ssDNA	Single-stranded DNA
S. solfataricus or *Sso*	*Sulfolobus solfataricus*
TBST	Tris-buffered saline with Tween20
TCA	Trichloroacetic acid
Tris	Tris (Hydroxymethyl) aminomethane
Tween20	Polyoxyethylene (20) sorbitan monolaurate
UAS	Upstream activating sequence
UV	Ultraviolet
X-gal	5-bromo-4chloro-3-indolyl β-D-galactoside

CONTENTS

TABLES AND FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 5

CHAPTER 1
INTRODUCTION

> "A phylogenetic analysis based upon ribosomal RNA sequence characterisation reveals that living systems represent one of three aboriginal lines of descent: i) the eubacteria, comprising all typical bacteria; ii) the archaebacteria, containing methanogenic bacteria; and iii) the urkaryotes, now represented in the cytoplasmic component of eukaryotic cells. "
>
> CR Woese and GE Fox, 1977

1.1. Archaea-the third domain of life

These words marked one of the most important scientific discoveries of modern times-the three-domain classification of life. Based on analysis of the small subunit ribosomal RNA nucleotide sequence, Woese and Fox found a new group of prokaryotic microorganisms, which are no more related to bacteria than they are to eukaryotes (Woese and Fox, 1977). Initially, this group was named archaeabacteria, but subsequently this name was shortened to archaea (Woese *et al.*, 1990).

Phylogenetic analysis based on 16S rRNA divided archaea into two major kingdoms: euryarchaeota and crenarchaeota (Woese *et al.*, 1990). Euryarchaeota is the most diverse group and includes hyperthermophiles, methanogens, thermoacidophiles and halophiles. The crenarchaeota are less diverse and include hyperthermophiles, psychrophiles and thermoacidophiles. Unlike euryarchaeota, crenarchaeota lack bacterial cell division proteins, eukaryal-like histones and RPA-like proteins. In addition to the above archaeal groups, two minor kingdoms have been proposed-korarchaeota (Barns *et al.*, 1996) and nanoarchaea (Huber *et al.*, 2002; Huber *et al.*, 2003). The existence of korarchaeota was deduced from small subunit rDNA sequence data, but the isolation and cultivation of korarchaeal organisms has not yet been reported. Nanoarchaea are believed to have the smallest genomes of all known cellular life forms (Fig. 1.1).

Despite their unique peculiarities, phylogenetic analysis of archaeal genomes has revealed that they exhibit a mixture of features from both prokaryotes and eukaryotes. Like bacteria, archaea have a prokaryotic cell morphology showing an

absence of nuclear structures, cytoskeleton and organelles (Edgell and Doolittle, 1997; Olsen and Woese, 1997) and bidirectional chromosomal replication (Koonin *et al.*, 1997; Jain *et al.*, 1999; Forterre *et al.*, 2002). However, sequence information from several archaeal genomes revealed striking differences between archaea and bacteria in the organisation of their information-processing systems. The structures of ribosomes and chromatin, the presence of histones, and sequence similarity between proteins involved in replication, transcription, translation and DNA repair all suggest a closer relationship between archaea and eukaryotes than between either of these and bacteria (Sandman and Reeve, 2000; Kelman, 2000; Bell and Jackson, 2001; MacNeill, 2001; Forterre *et al.*, 2002; Kelman and Kelman, 2003; Reeve, 2003). Despite their mosaic nature, archaea are considered a monophyletic group, with most encoded proteins more similar to other archaeal proteins than those of bacteria and eukarya (Forterre *et al.*, 2002). However, because a subset of eukaryotic proteins are found in archaea, archaea are extensively used as a simple model system for studying complex eukaryotic processes such as DNA replication, transcription and translation.

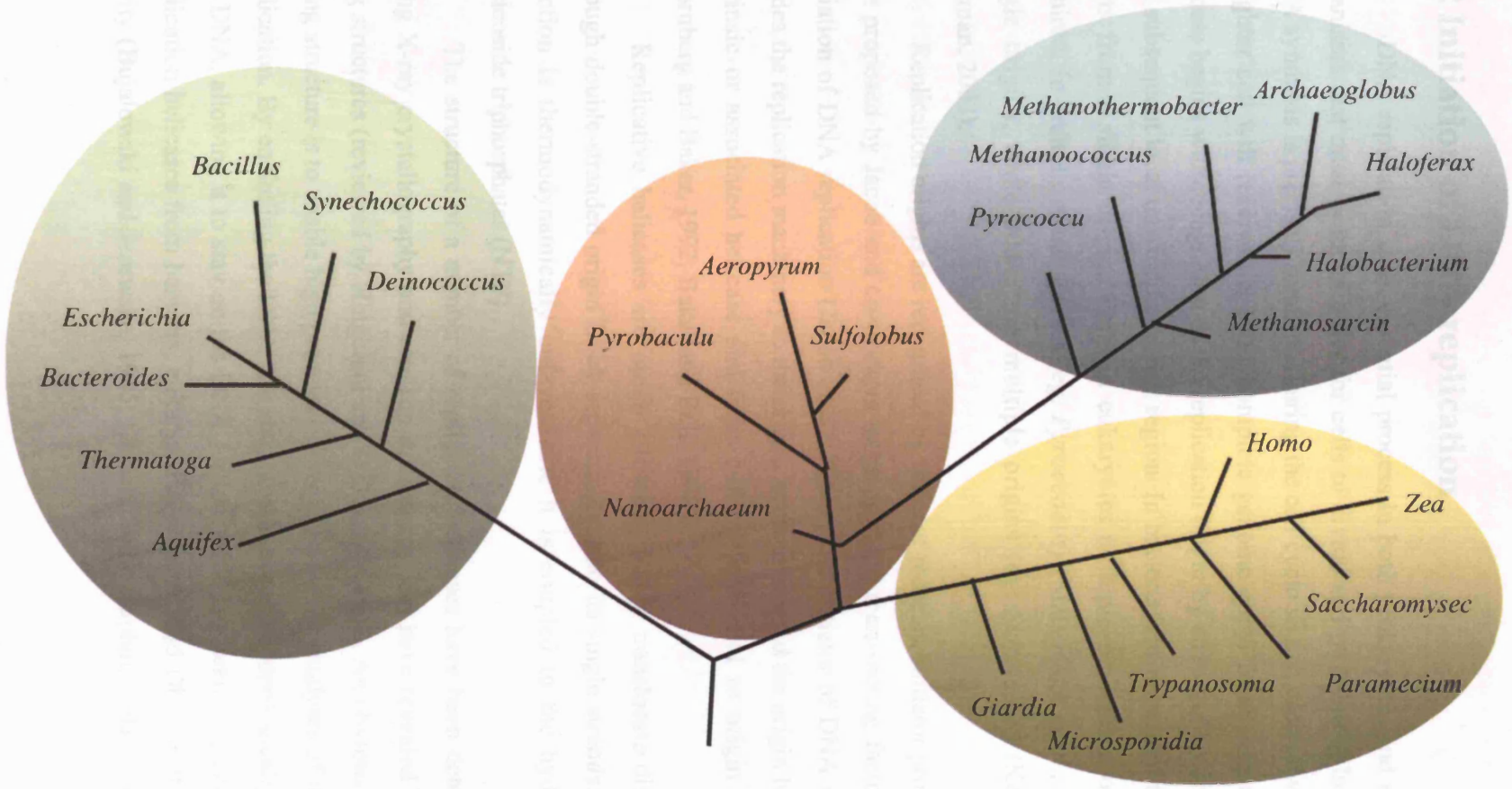


Fig. 1.1 The three domains of life-bacteria (green oval), eukarya (yellow oval), and archaea with its two kingdoms: euryarchaeota (light blue) and crenarchaeota (tan oval).

1.2 Initiation of DNA replication

DNA replication is an essential process in both prokaryotic and eukaryotic organisms that must occur in order for cells to divide and proliferate. Initiation of DNA synthesis is precisely regulated during the cell cycle so that each newly divided daughter cell will receive the same complete genome copy. In all organisms this process begins with recognition of the replication origin by an origin binding protein and subsequent local unwinding of that region. In bacteria, chromosomal replication starts from a single origin, while in eukaryotes multiple replication origins are required. In archaea, some species (eg. *Pyroccocus*) (Matsunaga *et al.*, 2001) use single origins, while others use multiple origins (eg. *Sulfolobus*) (Kelman and Kelman, 2003).

Replication origins are recognised by initiator proteins. Initiator proteins were first proposed by Jacob and co-workers to be essential *trans*-acting factors for the initiation of DNA replication (Jacob *et al.*, 1964). The initiator of DNA replication guides the replication machinery to the origin, helps to unwind the origin by using an intrinsic or associated helicase and links cell cycle control to origin activation (Kornberg and Baker, 1992; Baker and Bell, 1998).

Replicative helicases are motor proteins, which translocate directionally through double-stranded origin DNA separating it into its single strands. Since the reaction is thermodynamically unfavourable it is coupled to the hydrolysis of nucleoside triphosphates (NTP).

The structure of a number of replicative helicases have been determined in using X-ray crystallography and electron microscopy and have revealed hexameric ring structures (reviewed by Hingorani and O'Donnell, 2000). An obvious reason for a ring structure is to unable high processivity required for the catalyses of the genome replication. By encircling the DNA, the ring protein becomes topologically linked to the DNA, allowing it to stay on the DNA for longer. It had been suggested that the replication helicases from bacteria, eukarya and archaea bind DNA in their central cavity (Bujalowski and Jezewska, 1995; Kaplan, 2000; Kaplan, 2003). This binding

of the DNA in the central channel together with the ring structure decrease the possibility of complete dissociation of these helicases from the DNA and ensures good processivity.

The replication process in all organisms begins with the ordered assembly of a multiprotein complex, called pre-replication complex (pre-RC). Formation of the pre-RC complex in both prokaryotes and eukaryotes involves marking of the replication origin by an origin binding protein, followed by loading of the replicative helicase.

Some of the of the prokaryotic and eukaryotic pre-RC proteins belong to the AAA+ family of ATPases. AAA+ proteins function as ATP-driven conformational switches that can mediate the assembly, remodelling, or disassembly of protein complexes. The AAA+ proteins share a common structure, consisting of an amino-terminal mononucleotide-binding fold and a carboxy-terminal helical domain. Each of them contain highly conserved Walker A and Walker B motifs. The Walker A motif, which is also referred to as the P-loop, is involved in binding the triphosphate tail of the NTP (Walker *et al.*, 1982). It also contains the invariant lysine residue, which directly interacts with the phosphatyl group of bound ATP. The Walker A motif also contains a conserved threonine residue which co-ordinates the Mg^{2+} ion required for ATP hydrolysis. The Walker B motif has a hydrophobic structure. The acidic residues within this motif play an essential role in coordinating the Mg^{2+}. It is postulated that Walker B motif is involved in ATP hydrolysis, rather than ATP binding (Pause and Sonenberg, 1992; Brosh and Matson, 1995).

1.3 Origin binding in bacteria

1.3.1 Bacterial replication origins

Among bacteria, *E.coli* has the best characterised initiation process. Initiation of chromosome replication in *E.coli* is a complex process involving several different proteins, the most prominent being the DnaA protein (Kornberg and Baker, 1992). Bacterial initiator protein DnaA recognises and binds DNA in a sequence-specific manner to multiple sites within the single origin (*oriC*) (reviewed by Giraldo, 2003). *In vitro* studies suggest that DnaA is bound to ATP or ADP but only the DnaA-ATP complex is able to initiate replication at *oriC* (Sekimizu *et al.*, 1988). In the *E.coli* chromosome, 308 different recognition sites for DnaA (DnaA boxes) have been found with the following consensus sequence: TTAT(C/A)CA(C/A)A. Each of these boxes exhibits different affinity to DnaA (Schaper and Messer, 1995). Eight of them have been shown to bind preferentially to DnaA. Three of these sites are bound with high to medium affinity (R1, R4 and R2) and five are with lower affinity (R5(M), I2, I3, I1 and R3) (Fuller *et al.*, 1984; Matsui *et al.*, 1985, Grimwade *et al.*, 2000). The R boxes can bind inactive DnaA-ADP and active DnaA-ATP with equal affinity (Sekimizu *et al.*, 1987; Schaper and Messer, 1995). The higher affinity R boxes have a consensus sequence 5' TGTGNAT/AAA. The I boxes differ from the R box consensus at 3-4 bases and exhibit a preference for binding DnaA-ATP (Ryan *et al.*, 2002; McGarry, 2004). The main difference between the I sites and the R boxes therefore is their binding affinity for DnaA-ADP. McGarry and co workers suggested that the third base of either I2 or I3 is the main factor that triggers discrimination between DnaA-ADP and DnaA-ATP forms (McGarry *et al.*, 2004).

The interaction of DnaA with *oriC* is switch-like, with high affinity sites R4, R1 and R2 filled throughout most of the cell cycle and weaker sites R5(M), I1, I2, I3 and R3 only filled immediately before the onset of DNA synthesis (Samitt *et al.*, 1989; Cassier *et al.*, 1995, Ryan *et al.*, 2002). The order of DnaA recognition site

loading *in vitro* is R4,R1>R2>R5(M),I2>I3>I1>R3 (Margulies and Kaguni, 1996; Weigel *et al.*, 1997; Grimwade *et al.*, 2000; Ryan *et al.*, 2002).

ATP binding is not required for the binding of DnaA to the R boxes from *oriC*, but is essential for its specific binding to six secondary sites (named ATP-DnaA boxes) found in the three 13-mer AT-rich repeats adjacent to R1 box (Speck *et al.*, 1999; Speck and Messer, 2001). Co-operative binding of DnaA to the R1 and the ATP-DnaA boxes results in DNA unwinding stabilisation of the single strands by DnaA-ATP (Speck and Messer, 2001) and loading of the DnaB helicase (Davey *et al.*, 2002) (Fig. 1.2).

In most of the bacterial origins, the DnaA boxes are organised in clusters. There are two hypotheses concerning the numbers of DnaA box clusters (Mackiewicz *et al.*, 2004). It is possible that more than one cluster participates in the initiation of replication. For example, the *oriC* of *B.subtilis* is composed of two DnaA clusters that are both required for a functional origin. However, in *Mycobacterium*, only one of the two clusters acts as an origin (Madiraju *et al.*, 1999). The second hypothesis suggests that clusters of DnaA boxes may be responsible for the negative regulation of DNA replication by reducing the levels of free DnaA protein.

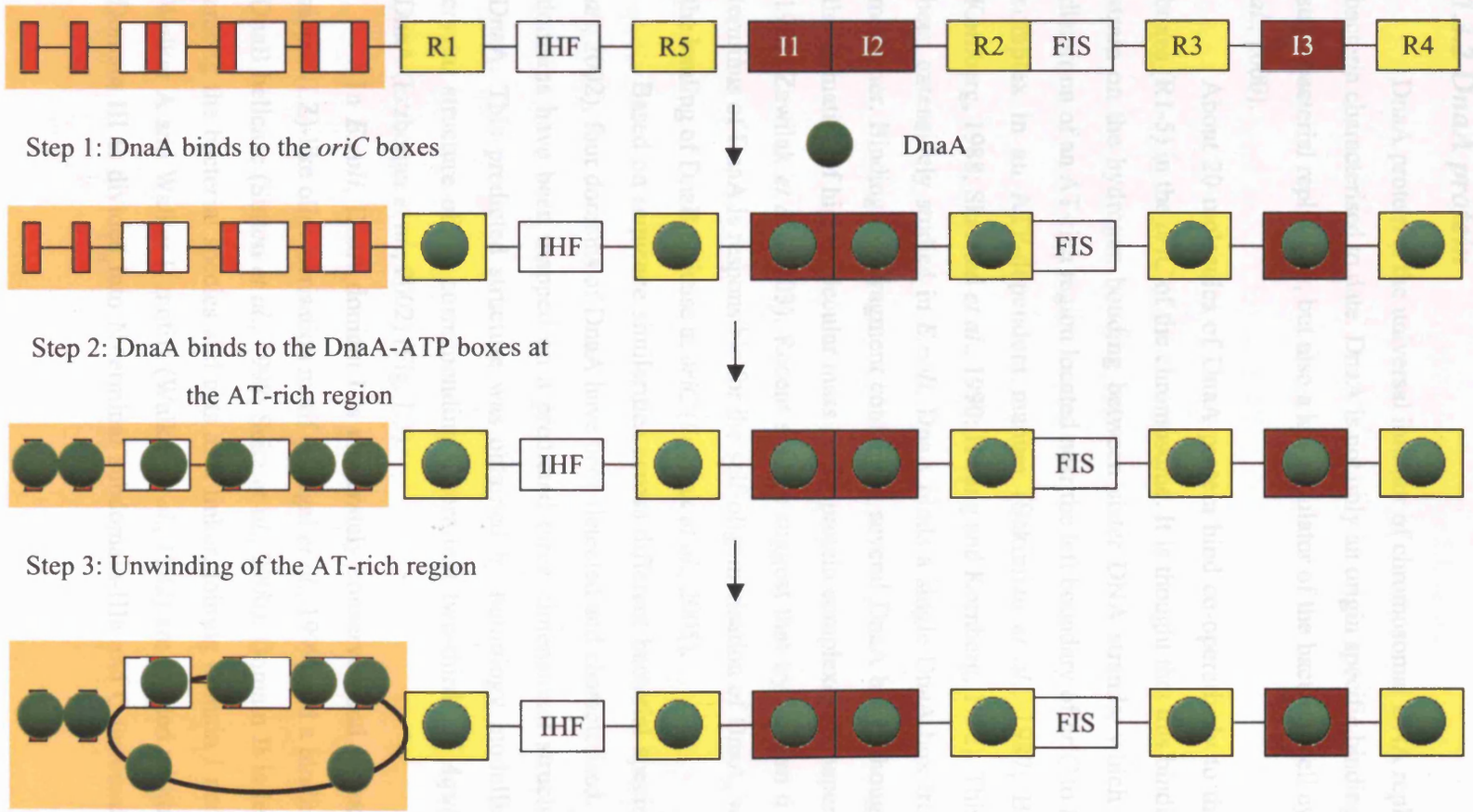


Fig. 1.2 DnaA-dependent unwinding reaction at *oriC*. The yellow colour represents the R boxes, the brown colour-the I boxes. The red colour represents the DnaA-ATP boxes located within the AT-rich regions on the left. The resulting ssDNA from the unwinding reaction will then be bound by the DnaB helicase. The binding sites for IHF and FIS are shown (Ryan *et al.*, 2004).

1.3.2 *DnaA protein*

DnaA protein is the universal initiator of chromosomal DNA replication in all bacteria characterised to date. DnaA is not only an origin specific binding protein that starts bacterial replication, but also a key regulator of the bacterial cell cycle (Boye *et al.*, 2000).

About 20 molecules of DnaA protein bind co-operatively to the five DnaA boxes (R1-5) in the *oriC* of the chromosome. It is thought that this binding induces a strain on the hydrogen bonding between sister DNA strands, which causes local distortion of an AT-rich region located near the left boundary of *oriC* to form an open complex in an ATP-dependent manner (Sekimizu *et al.*, 1987; Bramhill and Kornberg, 1988; Skarstad *et al.*, 1990; Hwang and Kornberg, 1992). This process has been extensively studied in *E.coli*. DnaA binds a single DnaA box from *oriC* as a monomer. Binding to a fragment containing several DnaA boxes though resulted in the formation of high molecular mass nucleoprotein complexes (Schaper and Messer, 1995; Zawilak *et al.*, 2003). Recent studies suggest that tryptophan 6 from the N-terminus of DnaA is responsible for the self-oligomerisation of DnaA, which leads to the loading of DnaB helicase at *oriC* (Felczak *et al.*, 2005).

Based on sequence similarities across different bacterial species (Messer *et al.*, 2002), four domains of DnaA have been detected and characterised. Two of these domains have been mapped on a predicted three-dimensional structure of *E.coli* DnaA. This predicted structure was obtained by homology modelling using the crystal structure of the corresponding C-terminal two-thirds of *Aquifex aeolicus* DnaA (Erzberger *et al.*, 2002) (Fig. 1.3).

In *E.coli*, DnaA domain I is moderately conserved and includes a leucine zipper (LZ)-like oligomerisation motif (Weigel *et al.*, 1999) and a binding site for the DnaB helicase (Sutton *et al.*, 1998; Seitz *et al.*, 2000). Domain II is less conserved among the bacterial species and acts as a linker, joining domain I and domain III. Walker A and Walker B motifs (Walker *et al.*, 1982) are located within domain III. Domain III is divided into N-terminal subdomain-IIIa and C-terminal subdomain-

IIIb. The Walker A motif (the P-loop) contains the GX_4GKT motif and is an ATP binding site. The Walker B motif contains the LLIDD motif and is responsible for ATP hydrolysis (Mizushima *et al.*, 1998). Domain III also contains a conserved motif, the box VII motif, which is shared by the AAA+ family of ATPases (Erzeberger *et al.*, 2002; Koonin, 1993). The box VII motif of DnaA forms a small helix that helps create a bipartite ATP-binding site between adjacent protomers in AAA+ oligomers (Neuwald *et al.*, 1999). Within the box VII motif, a conserved arginine residue reaches into the ATP-binding cleft to assist ATP hydrolysis or to transmit conformational signals between subunits by "sensing" whether hydrolysis has occurred. Mutation of this conserved arginine finger results in destabilisation of DnaA-*oriC* complexes and inactivates initiation at *oriC* (Felczak and Kaguni, 2004). Thus, box VII residues are essential for initiation of replication.

In the DnaA structure, the purine and ribose rings of ATP are held in a cavity formed by residues from the subdomains IIIa and IIIb. Nucleotide binding and hydrolysis result in a change in the position of each subdomain with respect to each other (Erzberger *et al.*, 2002; Wang *et al.*, 2001). Domain III contains residues, which are believed to be an interaction site for DnaB helicase, as well as residues that comprise two potential α-helices that bind acidic phospholipids and alter the conformation of the domain to release ADP (Garner and Crooke, 1996; Hase *et al.*, 1998; Makise *et al.*, 2001; Zheng *et al.*, 2001). Furthermore, the crystal structure of domain III and IV of the *Ae*DnaA reveals that a long α-helix (α12) connects domain III with the C-terminal DnaA binding domain IV (Fig. 1.3).

Domain IV consists of a typical helix-turn-helix motif (HTH) and a basic loop, which possibly makes a contact with the DNA minor groove or the phosphate backbone. The crystal structure of domain IV with a dsDnaA box (R1) shows that DnaA domain IV makes base-specific interactions with the major and minor grooves of the DnaA box (Fig. 1.4a) (Fujikawa *et al.*, 2003). Previous biochemical experiments have suggested that upon binding, DnaA induces bending of the DNA (Schaper and Messer, 1995). Indeed, for opening of the duplex DNA at *oriC*, the origin DNA has to become supercoiled (Kornberg and Baker, 1992). The DnaA

multimer formed at the origin is believed to wrap the *oriC* DNA by DNA bending at several points, thus introducing supercoiling. This DnaA multimer formation with DNA bending may be the basis for the initiation complex formation, which is the crucial step required for duplex opening (Fig.1.4b)

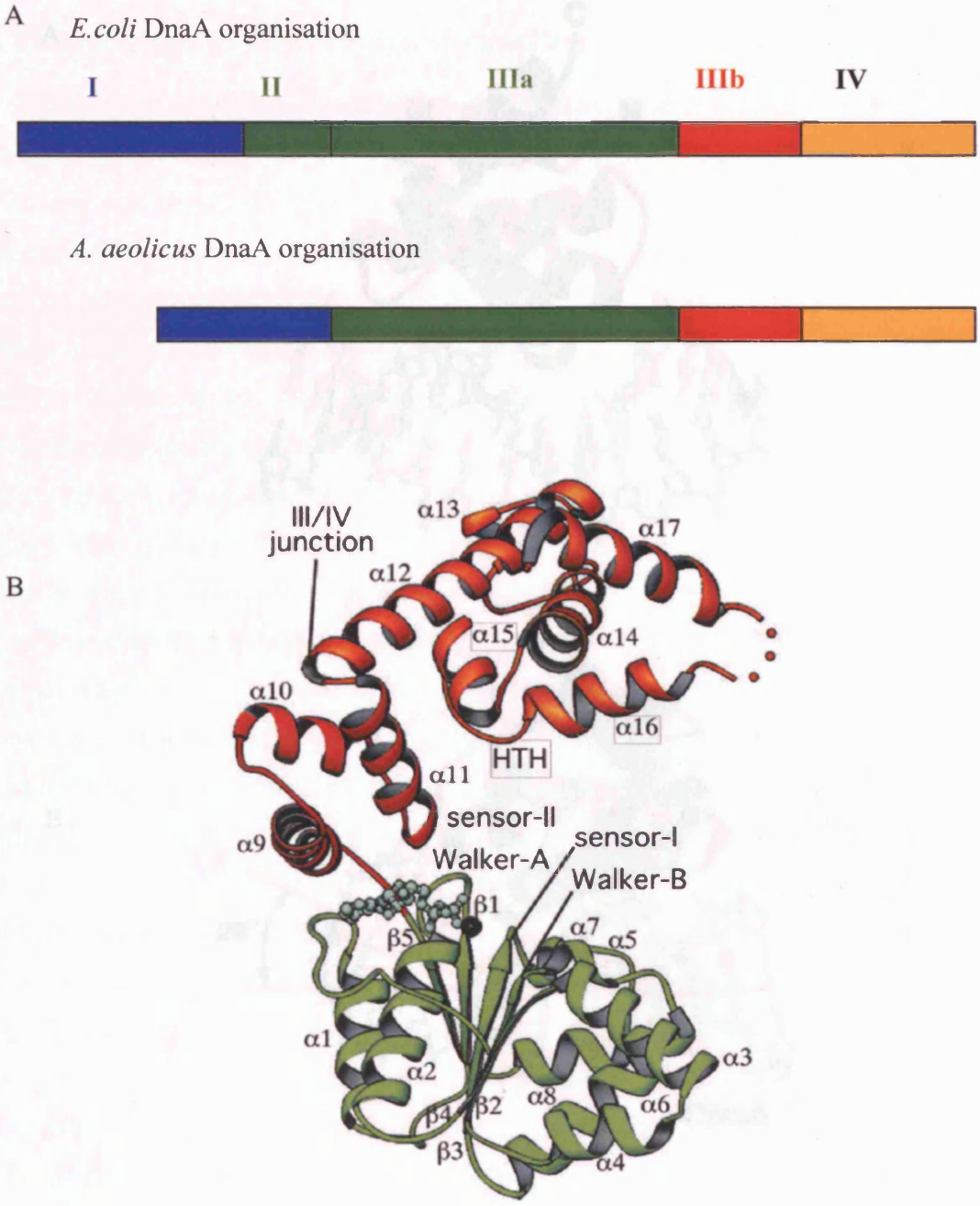


Fig. 1.3 Structure of DnaA. (A) Domain organisation of DnaA from *E.coli* and *A.aeolicus*; (B) Crystal structure of domain III and IV of *A.aeolicus* DnaA. Domains are coloured as follows: I-blue, II-green, IIIa-green, IIIb, red, IV-gold. The bound nucleotide is shown in turquoise (Erzberger *et al.*, 2002).

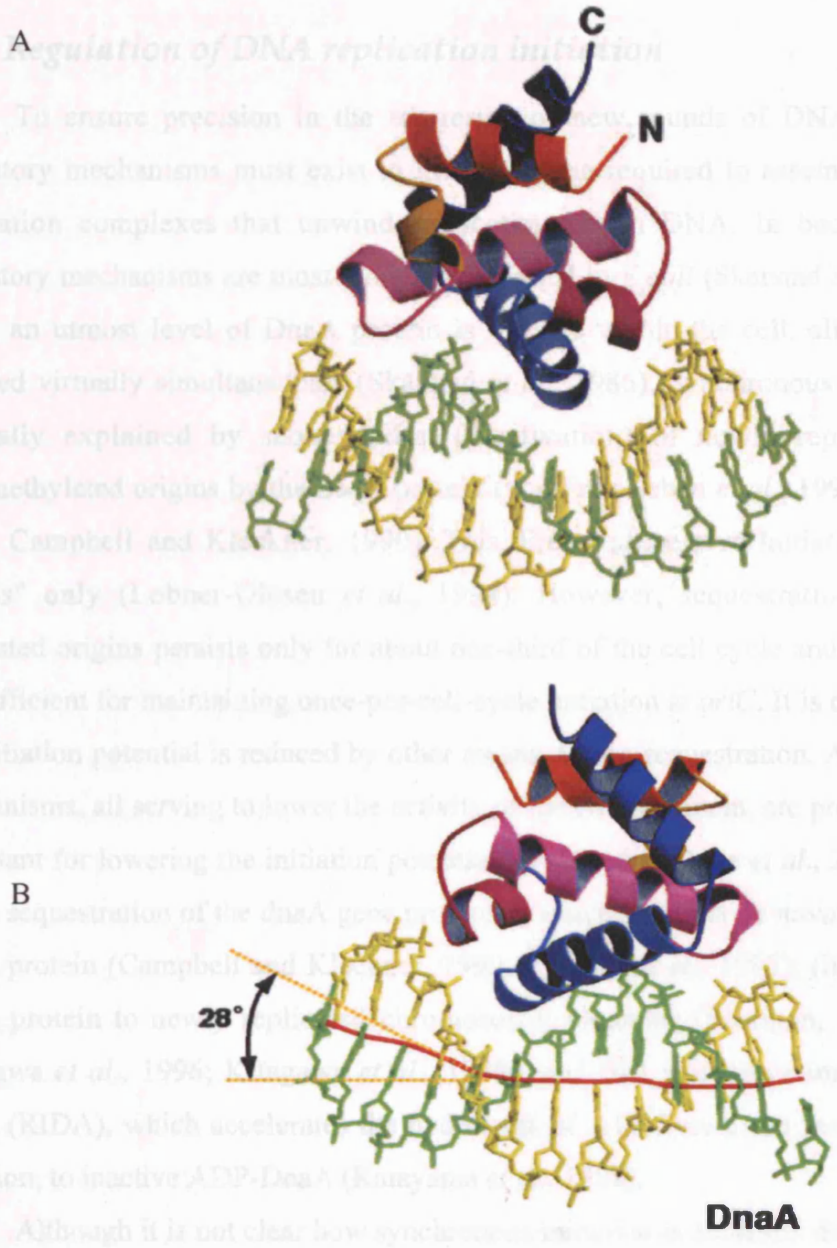


Fig. 1.4 Crystal structure of DnaA domain IV complexed with DnaA box (R1) DNA. (A) Overall structure of DnaA domain IV-DNA complex; (B) DNA bending induced upon DnaA binding by a base-specific interactions and the interactions with the backbone phosphate groups in the major and minor grooves (Fujikawa *et al.*, 2003).

1.3.3 Regulation of DNA replication initiation

To ensure precision in the triggering of new rounds of DNA synthesis, regulatory mechanisms must exist to limit the time required to assemble the pre-replication complexes that unwind replication origin DNA. In bacteria, these regulatory mechanisms are most thoroughly studied in *E.coli* (Skarstad *et al.*, 1986). When an utmost level of DnaA protein is reached within the cell, all origins are initiated virtually simultaneously (Skarstad *et al.*, 1986). Synchronous initiation is generally explained by sequestration (inactivation) of newly replicated and hemimethylated origins by the SeqA protein (von Freiesleben *et al.*, 1994; Lu *et al.*, 1994; Campbell and Kleckner, 1990). This directs successive initiations to "old origins" only (Løbner-Olesen *et al.*, 1994). However, sequestration of newly replicated origins persists only for about one-third of the cell cycle and is therefore not sufficient for maintaining once-per-cell-cycle initiation at *oriC*. It is essential that the initiation potential is reduced by other means during sequestration. At least three mechanisms, all serving to lower the activity of the DnaA protein, are proposed to be important for lowering the initiation potential (reviewed in Boye *et al.*, 2000). These are (i) sequestration of the dnaA gene promoter, which prevents *de novo* synthesis of DnaA protein (Campbell and Kleckner, 1990; Theisen *et al.*, 1993); (ii) titration of DnaA protein to newly replicated chromosomal elements (Morigen, *et al.*, 2001; Kitagawa *et al.*, 1996; Kitagawa *et al.*, 1998); and (iii) regulatory inactivation of DnaA (RIDA), which accelerates the hydrolysis of ATP-DnaA, the form active for initiation, to inactive ADP-DnaA (Katayama *et al.*, 1998).

Although it is not clear how synchronous initiation is achieved during the cell cycle, several *oriC* binding proteins function as precision factors in the pre-RC assembly (Boye *et al.*, 1993; Bahloul *et al.*, 2001; Ryan *et al.*, 2002). Two of these proteins are the DNA bending proteins IHF (Nash *et al.*, 1996) and FIS (Finkel and Johnson, 1992). Cells lacking these proteins experience severe initiation asynchrony during rapid growth (Boye *et al.*, 1993; Ryan *et al.*, 2002).

FIS and IHF proteins affect the DnaA-containing pre-RC differently. FIS inhibits *in vitro* DNA replication from *oriC* (Hiasa and Marians, 1994; Wold *et al.*, 1996; Margulies and Kaguni, 1998). FIS interacts at multiple sites within *oriC*, but the highest affinity binding site (primary site) is located immediately next to the R2 box (Filutowicz *et al.*, 1992) (Fig. 1.2). In living cells, FIS can be detected in its primary site throughout the majority of the cell cycle (Cassier *et al.*, 1995). However, at the time of initiation of DNA synthesis, FIS binding is reduced.

The IHF protein plays a major role in stimulating the DnaA-induced *oriC* unwinding (Hwang and Kornberg, 1992; Grimwade *et al.*, 2000; Ryan *et al.*, 2002). IHF protein binds to a unique site located between R1 and R5 (Polaczek, 1990) (Fig. 1.2) and alters the interaction of DnaA with *oriC in vitro* (Grimwade *et al.*, 2000; Ryan *et al.*, 2002). Particularly, IHF stimulates redistribution of DnaA from high affinity to lower affinity sites (Grimwade *et al.*, 2000). This IHF-induced redistribution of DnaA causes *oriC* unwinding at DnaA concentrations below those needed in the absence of IHF. IHF ability to redistribute DnaA to I2 and I3 sites may be the mechanism for its ability to stimulate unwinding (McGarry *et al.*, 2004). Unlike FIS, IHF interacts with *oriC* during only a very short interval of the cell cycle, coincident with the onset of DNA synthesis (Cassier *et al.*, 1995). It is likely that FIS and IHF act in opposition since they do not bind the *oriC* simultaneously. The behaviour of DnaA, FIS and IHF binding therefore suggest that the assembly of pre-RC is an interactive process during the cell cycle, with each protein contributing, either positively or negatively, to *oriC* unwinding. A dynamic interplay exists among DnaA, FIS and IHF that affects loading of DnaA on *oriC* (Ryan *et al.*, 2004). In the presence of IHF alone, DnaA loading at weak affinity cites is continuous. However, in the presence of FIS and IHF, DnaA loading is suppressed at its weaker affinity sites until a DnaA concentration is reached that could overcome FIS repression and allow filling of all remaining DnaA binding sites simultaneously. FIS suppresses DnaA binding to weaker sites by competing with DnaA for overlapping FIS/DnaA sites. The suppression of DnaA binding at weaker sites allows accumulation of DnaA oligomers at the stronger sites that are necessary for open complex formation

(Simmons *et al.*, 2003). FIS acts as a repressor of IHF binding. When DnaA levels are high enough to displace FIS from its primary sites, IHF can bind quickly and redistribute DnaA from stronger to weaker sites overcoming the FIS repression (Grimwade *et al.*, 2000; Ryan *et al.*, 2002). It has been proposed that DnaA binding to R2 box weakens the ability of FIS to repress IHF binding. R2 is located immediately next to the primary FIS site at *oriC*. R2 box has an intermediate affinity for DnaA, which can lead to displacement of FIS protein at appropriate concentrations of DnaA, and triggering of *oriC* unwinding.

1.4 Origin binding in eukaryotes

1.4.1 Eukaryal replication origins

In eukaryotic cells, DNA replication is initiated from multiple locations referred to as replication origins throughout each chromosome. The sequences required for an origin of replication vary significantly between different eukaryotic organisms.

In the budding yeast *Saccharomyces cerevisiae*, DNA sequences important for replication were first identified by ARS (autonomously replicating sequence) assays that determine whether a given DNA fragment initiates replication when placed on a plasmid (Stinchcomb *et al.*, 1979). Three or four sequences of 10-15 base pairs (bp) spread over a few hundred base pairs in length are sufficient to act as an origin. These sequences include the highly conserved and essential A-element or autonomously replicating consensus sequence (ACS) (Palzkill and Newlon, 1998; Newlon and Theis, 1993), which is necessary but not sufficient for replication function (Marahrens and Stillman, 1992). The ACS is recognised in an ATP-dependent manner by the six-subunit origin recognition complex (ORC) (Bell and Stillman, 1992; Klemm *et al.*, 1997; Lee and Bell, 1997). In addition to the A element, ARS function requires an approximately 100-bp B domain, which is 3' of the T-rich strand of the ACS. Domain B has been found to be composed of subdomains B1, B2 and B3. In *S.cerevisiae*, element B1 contributes to ORC binding (Rao and Stillman, 1995), element B3 binds the replication enhancing protein ABF1 (Walker *et al.*, 1990), while element B2 is believed to function as a DNA unwinding element (DUE). More recent analysis, however, suggests that B2 function requires sequence-specific contacts with the pre-replication complex (Wilmes and Bell, 2002) (Fig. 1.5). Identification of ARS sequences in other eukaryotes has been unsuccessful, suggesting that origins in other eukaryotes are not defined entirely by DNA sequence.

ARS elements have also been identified in the fission yeast *Schizosaccharomyces pombe*, but their sequence requirements are different. A much

larger region of between 800-1000 bp is required to direct initiation in fission yeast (Clyne and Kelly, 1995; Dubey *et al.*, 1994, Dubey *et al.*, 1996). Detailed analysis of this region identified several AT-rich sequences of 20-50bp that are important for origin function (Okuno *et al.*, 1999; Segurado *et al.*, 2003) (Fig. 1.5).

In higher eukaryotes, the organisation of origins appears even more complex and difficult to define. Two-dimensional gel-electrophoresis techniques have revealed that metazoan origins were much larger and heterogeneous in size, ranging from 2kb to 55kb (Hamlin and Dijkwel, 1995; DePamphilis, 1999). These origins were referred to as broad initiation zones with DNA replication occurring at multiple sites within these replication zones. The replication origins of all eukaryotes direct the formation of a number of protein complexes leading to the assembly of two bidirectional DNA replication forks. These events are initiated by the formation of the pre-replicative complex (pre-RC) at the origins during G1. Pre-RC formation involves the ordered assembly of a number of replication factors including ORC, CDC6, CDT1 and MCM 2-7.

1.4.2 Eukaryotic origin recognition complex (ORC)

The origin recognition complex (ORC) is a six-subunit complex that acts as the initiator at eukaryotic origins of replication. Initially, ORC was identified first in *S.cerevisiae* as a complex that binds to the conserved ACS (Bell and Stillman, 1992). Subsequent studies have now found that this complex is conserved in all eukaryotes (Romanowski *et al.*, 1996; Rowles *et al.*, 1996; Carpenter *et al.*, 1996; Landis *et al.*, 1997; Chesnokov *et al.*, 1999; Dhar *et al.*, 2001).

Like the bacterial DnaA, the main role of ORC is to recognise and bind origin DNA. The complex nature of ORC interaction with DNA has made identification of ORC binding sites in the origin and DNA binding domains in ORC difficult.

In *S.cerevisiae*, ORC interacts with both A and B elements of yeast origins in an ATP dependent manner (Bell and Stillman, 1992, Rao and Stillman, 1995; Rowley

et al., 1995; Lee and Bell, 1997). All five *Sc*ORC subunits (ORC1p-ORC5p) are required to recognise DNA, but only four of them (ORC1p, 2p, 4p and 5p) are in close contact with the origin (Lee and Bell, 1997). ORC6p is not required for DNA binding, but is essential for DNA replication and cell viability (Li and Herskowitz, 1993) (Fig. 1.5).

Interestingly, a C-terminal fragment of *Sc*ORC4p (Bell *et al.*, 1995) shares sequence similarity with the plasmid replication initiation protein REP (Giraldo and Diaz-Orejas, 2001). Although this level of sequence identity is not obvious in standard database searches, it could conceal a true structural similarity. Indeed, a structural similarity between REP and *Sc*ORC4p includes common structural elements, such as a winged helix domain and a leucine-zipper (LZ) dimerisation motif.

Unlike *Sc*ORC, ORC derived from *S.pombe* (*Sp*ORC) has a defined DNA binding motif at the N-terminus of *Sp*ORC4p, which contains a repeated AT-hook region that is essential for *S.pombe* viability (Chuang and Kelly, 1999). Studies of this DNA binding motif show that the AT-hook region interacts preferentially with AT-rich DNA similarly to the A-rich nature of *S.cerevisiae* ACS (Chuang and Kelly, 1999; Kong and DePamphilis, 2001; Lee *et al.*, 2001). In agreement with this, *Sp*ORC lacking this domain of *Sp*ORC4p, exhibits reduced affinity for chromatin *in vivo* (Moon *et al.*, 1999) (Fig. 1.5).

In *Drosophila*, the precise ORC binding site within replication origins has not been discovered. *In vivo* and *in vitro* studies indicate that the sequence specificity of *Dm*ORC binding is limited to an AT-rich 80-bp region adjacent to the start site of DNA replication (Austin *et al.*, 1999; Chesnokov *et al.*, 2001; Bielinski *et al.*, 2001). Both DNA binding and DNA replication require all six *Dm*ORC proteins (Chesnokov *et al.*, 2001) (Fig. 1.5).

Targeting ORC to specific chromosomal locations could be accomplished through its interactions with another pre-RC component-CDC6. CDC6 modulates ORC binding by increasing the stability of ORC on chromatin (Harvey and Newport, 2003) and inhibits ORC binding to non-specific DNA (Mizushima *et al.*, 2000).

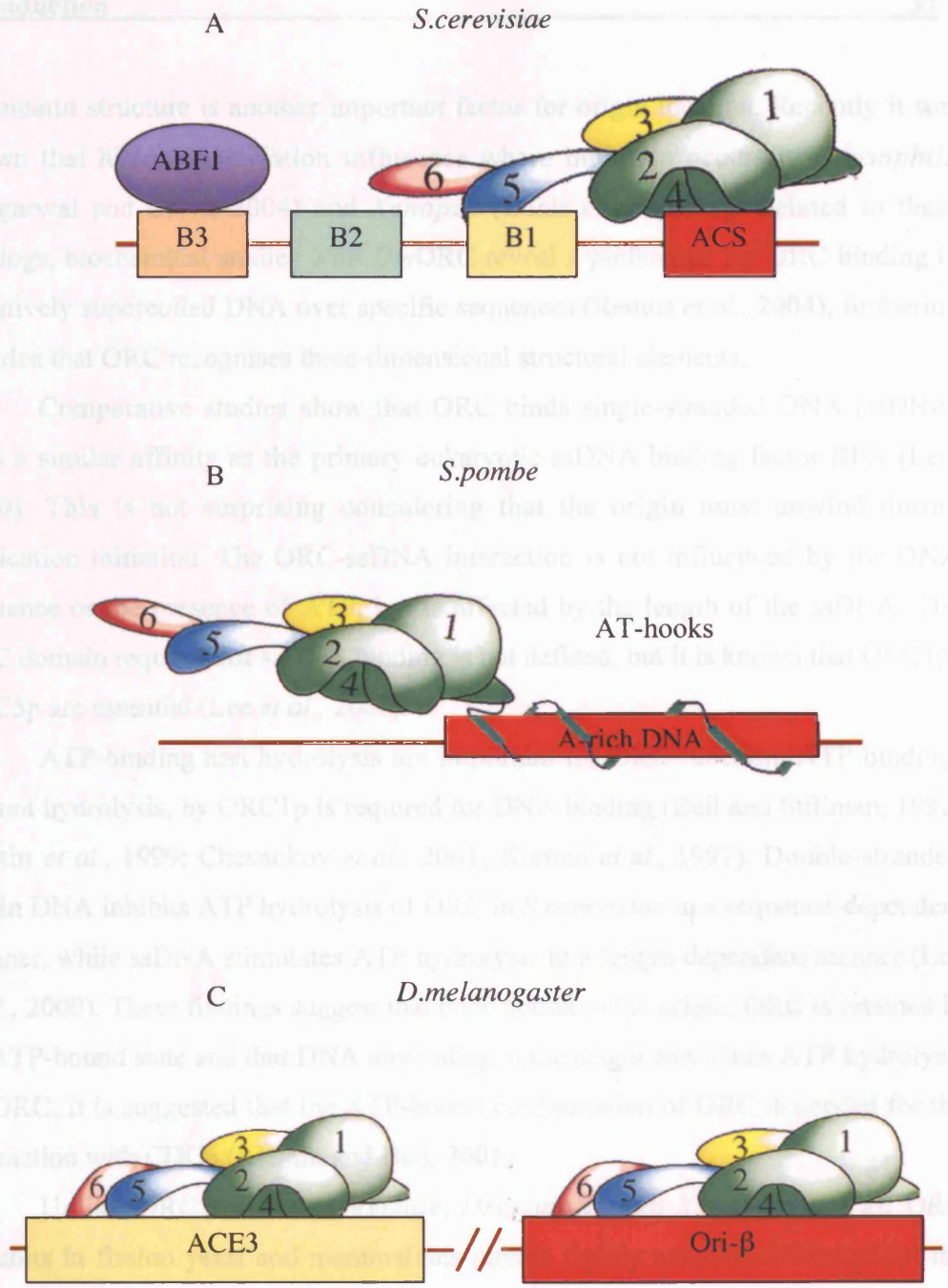


Fig. 1.5 Association of ORC subunits with origins of replication. (A) In *S.cerevisiae* only subunits *Sc*Orc1p, 2p, 4p and 5p interact directly with the ACS and B1 region of the origin; (B) In *S.pombe* an AT hook region of *Sp*ORC4p interacts with the AT-rich region of the origin, (C) In *D.melanogaster* all six subunits of *Dm*ORC are required for binding to both ACE3 and ori-β elements of the third chromosome (Bell, 2002)

Chromatin structure is another important factor for origin location. Recently it was shown that histone acetylation influences where initiation occurs in *Drosophila* (Aggarwal and Calvi, 2004) and *Xenopus* (Danis *et al.*, 2001). Related to these findings, biochemical studies with *Dm*ORC reveal a preference for ORC binding to negatively supercoiled DNA over specific sequences (Remus *et al.*, 2004), furthering the idea that ORC recognises three-dimensional structural elements.

Comparative studies show that ORC binds single-stranded DNA (ssDNA) with a similar affinity as the primary eukaryotic ssDNA binding factor RPA (Lee, 2000). This is not surprising considering that the origin must unwind during replication initiation. The ORC-ssDNA interaction is not influenced by the DNA sequence or the presence of ATP, but is affected by the length of the ssDNA. The ORC domain required for ssDNA binding is not defined, but it is known that ORC1p-ORC5p are essential (Lee *et al.*, 2000).

ATP binding and hydrolysis are important for ORC function. ATP binding, but not hydrolysis, by ORC1p is required for DNA binding (Bell and Stillman, 1992; Austin *et al.*, 1999; Chesnokov *et al.*, 2001; Klemm *et al.*, 1997). Double-stranded origin DNA inhibits ATP hydrolysis of ORC in *S.cerevisiae* in a sequence-dependent manner, while ssDNA stimulates ATP hydrolysis in a length-dependent manner (Lee *et al.*, 2000). These findings suggest that once bound to the origin, ORC is retained in an ATP-bound state and that DNA unwinding at the origin stimulates ATP hydrolysis by ORC. It is suggested that the ATP-bound conformation of ORC is needed for the interaction with CDC6 (Klemm and Bell, 2001).

Unlike ORC from *S.cerevisiae*, *Drosophila* and *Xenopus*, not all ORC subunits in fission yeast and mammalians remain tightly associated throughout the cell cycle (Moon *et al.*, 1999; Natale *et al.*, 2000; Thome *et al.*, 2000, Kreitz *et al.*, 2001; Tatsumi *et al.*, 2000; Saha *et al.*, 1998; Okuno *et al.*, 2001). This suggests the hypothesis that the six-subunit ORC may be assembled on the chromatin only during the G1 and S phase of the cell cycle. ORC binding to chromatin is regulated through the cell cycle in some but not all species. In *S.cerevisiae*, *S.pombe* and *D.melanogaster* ORC is associated with origin sequences throughout the cell cycle

(Ogawa *et al.*, 1999; Aparicio *et al.*, 1997; Santocanale and Diffley, 1996; Tanaka *et al.*, 1997, Pak *et al.*, 1997). However, in *Xenopus* and mammalians, the whole or at least part of the ORC complex is cleared from the chromatin during metaphase (Romanowski *et al.*, 1996; Carpenter *et al.*, 1996; Natale *et al.*, 2000; Kreitz *et al.*, 2001; Tatsumi *et al.*, 2000). The removal of ORC could help prevention of pre-RC formation before the completion of M phase.

1.5 Origin binding in archaea

1.5.1 Archaeal replication origins

The first archaeal origin was predicted by skew analyses in *Pyrococcus abyssi* (Myllykallio *et al.*, 2000). Subsequently, *in silico* and *in vivo* studies and two-dimensional gel electrophoresis confirmed that the *P.abyssi* origin is flanked by a gene encoding the archaeal homologue of the initiation protein ORC/CDC6 (Kelman and Kelman, 2003; Kelman, 2000; Matsunga *et al.*, 2001). The replication origins of two other *Pyrococcus* species (*P.horikoshii*, *P.furiosus*) were identified (Salzberg *et al.*, 1998; Lopez *et al.*, 1999; Myllykallio *et al.*, 2000; Matsunaga *et al.*, 2001).

Bioinformatic analysis of the *Halobacterium* strain NRC-1 genome had suggested multiple replication origins of the large chromosome, coincident with two of the three chromosomal *orc/cdc6* genes (Kennedy *et al.*, 2001). Later on that year, Berquist and DasSarma confirmed functionally one of these replication origins (Berquist and DasSarma, 2003). This single replicating sequence was located near one of the *orc/cdc6* genes on the chromosome. Both *orc/cdc6* genes and an AT-rich region are required for autonomous replication. In addition to *Halobacterium* strain NRC-1, two other halophiles *H.marismortui* and *H.volcanii* also contained this conserved region (Berquist and DasSarma, 2003).

Recently, three replication origins of the crenarchaeal *S.acidocaldarius and S.solfataricus* were revealed (Lundgren *et al.*, 2004; Robinson *et al.*, 2004). The number of replication origins in both *Sulfolobus* species correlates to the number of their ORC/CDC6 homologues (named CDC6-1, 2, 3). Only two of the three origins are located next to the *orc/cdc6-1* replication initiation gene. In two of the *Sulfolobus* origins a number of repeats of a conserved DNA sequence were observed (called Origin of Recognition Boxes-ORB). One of the origin binding proteins of *Sulfolobus* recognises three distinct ORB elements in one of the origins. The second OBP protein binds several different sites within the origin and one site that is bound by the first OBP. All three OBPs proteins bound to the overlapping sites in the second origin

from both *Sulfolobus* species. Therefore, it has been proposed that one of the OBP acts as a repressor of replication, possibly by blocking access to the origins by residual OBP-1 and OBP-3 during the post-replicative G2 phase of the cell cycle and perhaps during the stationary phase (Robinson *et al.*, 2004).

The ORB elements are functionally conserved between species from both the euryarchaeota and crenarchaeota kingdoms (Robinson *et al.*, 2004). Thus, ORB elements may be a general feature of archaeal origins. A parallel may be drawn between archaeal and bacterial origins (Fig. 1.6). In bacteria, the origin binding boxes are also recognisably conserved between species (Messer *et al.*, 2001).

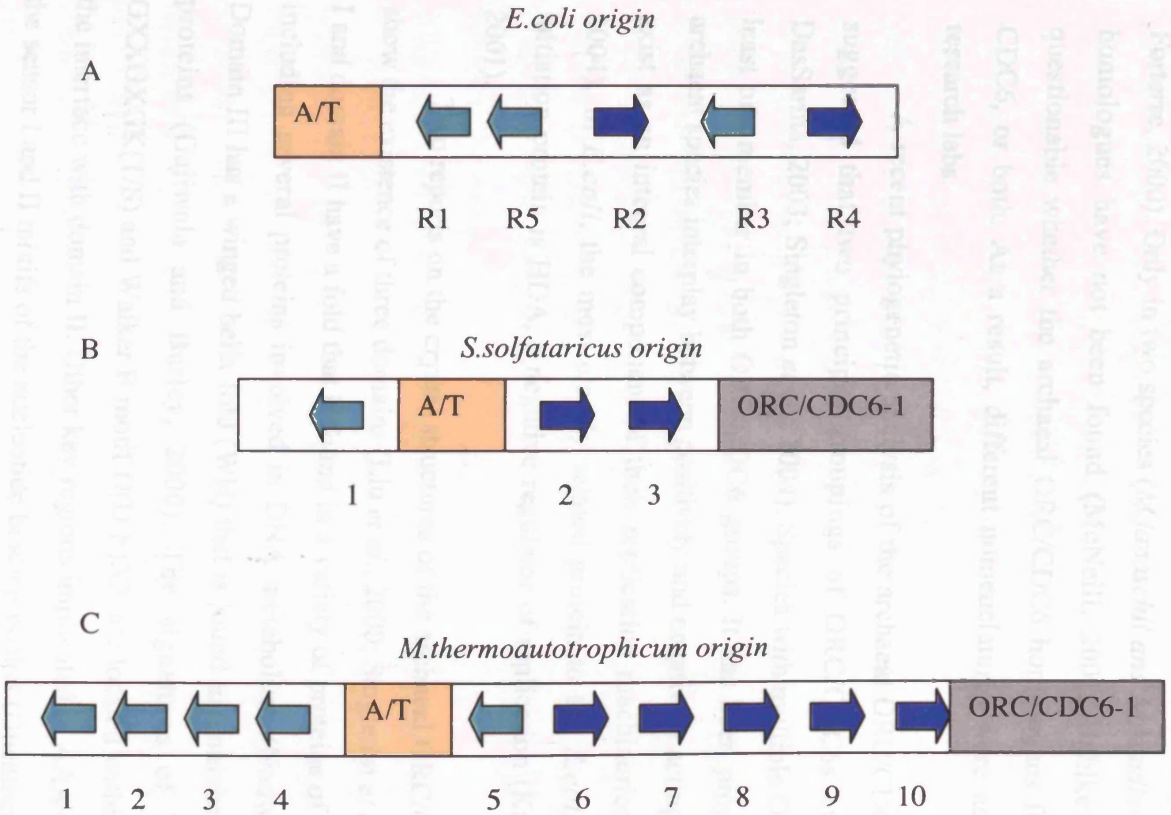


Fig. 1.6 Comparison of bacterial and archaeal origins of replication. (A) Origin structure of *E.coli*; (B) Origin structure of crenarchaeal *S.solfataricus*; (C) Origin structure of euryarchaeal *M.thermoautotrophicum*. The arrows represent the repeats and their direction (Adapted from Cunningham and Berger, 2005).

1.5.2 Archaeal ORC/CDC6 homologues

In most sequenced archaea, at least one homologue of ORC/CDC6 has been identified. Although the number of the ORC/CDC6 homologues varies in the different archaeal species, most of them contain one or two copies (Myllykallio and Forterre, 2000). Only in two species (*M.janaschii and M.kandleri*) clear ORC/CDC6 homologues have not been found (McNeill, 2001; Bohlke *et al.*, 2002). It is questionable whether the archaeal ORC/CDC6 homologues function as ORC or CDC6, or both. As a result, different nomenclatures were adopted by different research labs.

A recent phylogenetic analysis of the archaeal ORC/CDC6 homologues has suggested that two principal groupings of ORC/CDC6s exist (Berquist and DasSarma, 2003; Singleton *et al.*, 2004). Species with multiple ORC/CDC6s have at least one member in both ORC/CDC6 groups. It has been proposed that in some archaeal species interplay between positively and negatively acting ORC/CDC6s may exist as an integral component of their replication machineries (Robinson *et al.*, 2004). In *E.coli*, the most closely related protein to the *E.coli* DnaA replication initiation protein is HDA, a negative regulator of replication (Kato and Katayama, 2001).

Two reports on the crystal structures of the archaeal ORC/CDC6 homologues show the existence of three domains (Liu *et al.*, 2000; Singleton *et al.*, 2004). Domain I and domain II have a fold that is found in a variety of proteins of the AAA+ family, including several proteins involved in DNA metabolism (Neuwald *et al.*, 1999). Domain III has a winged helix fold (WH) that is found in a number of DNA binding proteins (Gajiwala and Burley, 2000). The signatures of Walker A motif GXXGXGK(T/S) and Walker B motif D(D/E)XX are located within domain I, but at the interface with domain II. Other key regions implicated in AAA+ function include the sensor I and II motifs of the nucleotide binding pocket (Guenther *et al.*, 1997) and they are found within domain I and II, respectively.

Although ADP was not added to the protein during crystallisation, both crystal structures of *A.pernix* and *P.aerophilum* ORC/CDC6 proteins show a presence of ADP (Liu *et al.*, 2000; Singleton *et al.*, 2004). Crystal structures of *Ape*ORC/CDC6 in a complex with the non-hydrolysable ATP analogue (ADPNP) exhibit an interesting feature (Singleton *et al.*, 2004). In the ADP complex of the protein, domain III is conformationally mobile, while in the ATP-complex of the protein a single conformation is stabilised (Fig. 1.7).

Domain III contains a winged helix motif (WH), which has been shown to be responsible for the DNA binding of the protein (Grainge *et al.*, 2003; Robinson *et al.*, 2004; Singleton *et al.*, 2004). Winged helix proteins are part of the helix-turn helix (HTH) family and are characterised by an α(β)βααββ fold, in which the later of the middle two helices represents the "recognition helix" (shown in bold) of the HTH motif and the "wing" of the fold arises from a loop which connects the last two β strands of the fold (Harrison, 1991; Pabo and Sauer, 1992; Nelson, 1995). It has been proposed that interaction between WH proteins and their DNA targets is *via* positively charged surfaces that interact with the phosphodiester backbone. Based on that observation, two modes of interaction between DNA and WH proteins have been proposed (Gajiwala and Burley, 2000). In the first mode, DNA binding is achieved by inserting the recognition helix into the major groove of DNA, while the β-hairpin wing interacts with the phosphodiester backbone. This mode of interaction is typical for the canonical WH proteins. However, another subfamily of WH proteins (to which RFX protein belongs) bind DNA differently by inserting the wing into the major groove, while the helix makes a phosphodiester backbone contact (Gajiwala *et al.*, 2000). Based on that and the sequence comparison of ORC/CDC6 homologues of variety of archaeal species, two classes of WH protein were defined. Class I proteins have a conserved positively charged recognition helix, while in class II proteins the number of positively charged residues is greater in the wing region.

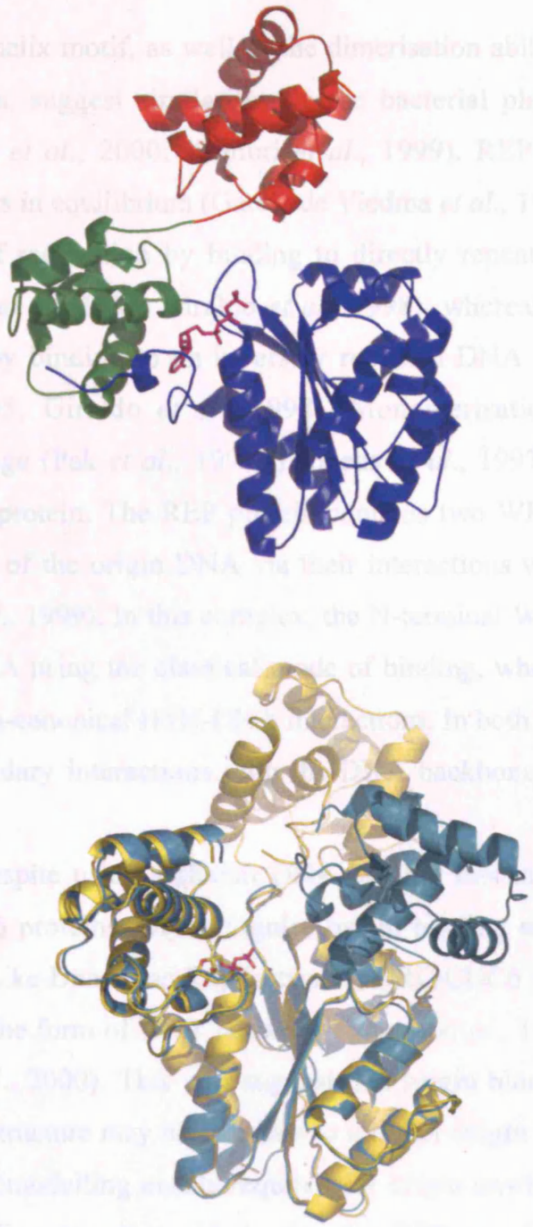


Fig. 1.7 Crystal structure of *A.pernix* ORC-2 protein. (A) An overall fold, domain I is coloured in blue, domain II in green, and domain III in red, the bound ADP-PNP is in magenta, (Singleton *et al.*, 2004); (B) Superimposition of *Ape*ORC-2 upon the structure of ORC-2 from *P.aerophilum.*

The winged helix motif, as well as the dimerisation abilities of the archaeal ORC/CDC6 proteins, suggest similarities to the bacterial plasmid initiator REP (Giraldo, 2003; Liu *et al.*, 2000; Komori *et al.*, 1999). REP molecules exist as monomers and dimers in equilibrium (Garcia de Viedma *et al.*, 1996). The monomers activate initiation of replication by binding to directly repeated DNA sequences (Garcia de Viedma *et al.*, 1995; Giraldo *et al.*, 1998), whereas the dimers repress *repA* transcription by binding to an inversely repeated DNA operator (Garcia de Viedma *et al.*, 1995; Giraldo *et al.*, 1998). Monomerisation is coupled to a conformational change (Pak *et al.*, 1997; Dibbens *et al.*, 1997) in the N-terminal domain of the REP protein. The REP protein contains two WH domains that bind together to one face of the origin DNA via their interactions with the DNA major groves (Komori *et al.*, 1999). In this complex, the N-terminal WH domain binds the major groove of DNA using the classical mode of binding, whereas the C-terminal WH domain uses non-canonical HTH-DNA interactions. In both domains, the "wing" loops mediate secondary interactions with the DNA backbone and minor grooves (Fig. 1.8).

Therefore, despite using different DNA binding mechanisms, REP and the archaeal ORC/CDC6 proteins may recognise origin binding sites in a structurally analogous manner. Like DnaA and Rep initiators, ORC/CDC6 protein can probably cause distortions in the form of DNA bending (Komori *et al.*, 1999; Fujikawa *et al.*, 2003; Gajiwala *et al.*, 2000). This will suggest that origin binding protein-induced alterations of DNA structure may be common to initiator-origin interactions, perhaps as a preface to the remodelling events required for origin unwinding and replisome assembly. These findings together with the fact that REP proteins exhibits structural similarities with *Sc*ORC4p protein (WH domains and LZ motifs) suggest that prokaryotic (bacterial and archaeal) and eukaryotic initiation complexes are very similar (Giraldo and Diaz-Orejas, 2001).

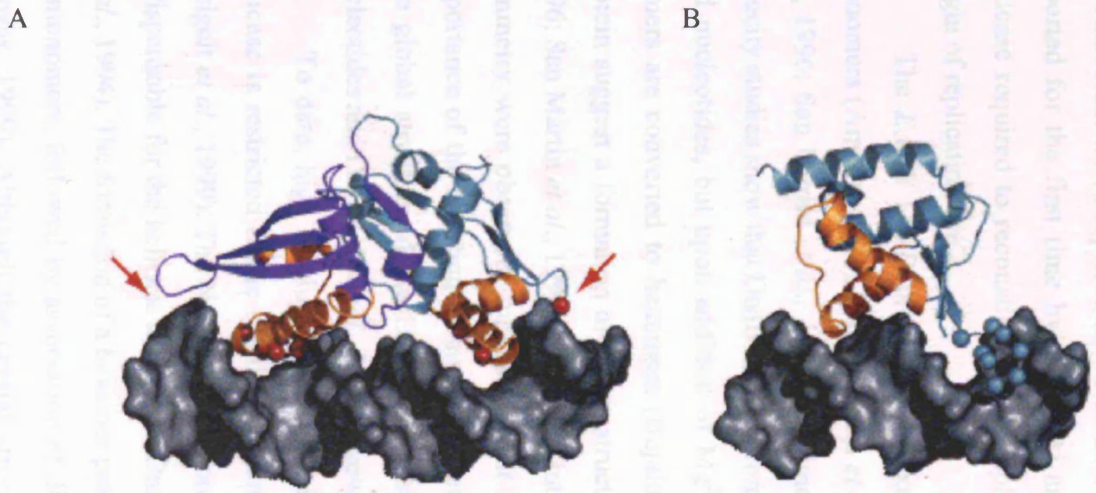


Fig. 1.8 Two models for origin binding protein-DNA interactions. (A) The *E.coli* RepE dimer binds DNA with its WHD1 (cyan) and WHD2 (purple). The protein binds the DNA major groove predominantly by inserting a recognition helix. Secondary interactions with the DNA backbone and minor groove are mediated by the wing loop (arrows); (B) A *Ape*ORC-2 binds with DNA by inserting a recognition wing into the major groove, while the helix makes a phosphobackbone contact (Singleton *et al.*, 2004; Cunningham and Burger, 2005).

1.6 Helicase loading in bacteria

1.6.1 DnaB

The *E.coli* DnaB protein is an essential component of the bacterial replication machinery that is involved in both the initiation and elongation stages of chromosomal DNA replication.. The DNA helicase activity of this protein was reported for the first time by LeBowitz and Macken in 1986. DnaB is the only helicase required to reconstitute DNA replication *in vitro* from the chromosomal origin of replication *oriC*.

The *E.coli* DnaB protein is a hexameric protein comprising six 52 kDa monomers (Arai *et al.*, 1981; Bujalowski *et al.*, 1994; San Martin *et al.*, 1995; Yu *et al.*, 1996; San Martin *et al.*, 1998). Sedimentation equilibrium and sedimentation velocity studies show that DnaB helicase forms mainly trimers in the absence of Mg^{2+} and nucleotides, but upon addition of Mg^{2+} (or a high concentration of Na^{+}) the trimers are converted to hexamers (Bujalowski *et al.*, 1994). EM studies on the protein suggest a formation of a ring structure (San Martin *et al.*, 1995; Yu *et al.*, 1996; San Martin *et al.*, 1998). Structures intermediate between the three and six-fold symmetry were observed, suggesting that the two forms interconvert. The global importance of the two ring forms in the helicase function of the protein is unclear. The global structure of the DnaB hexamers changes in the presence of different nucleotides and DNA (Yu *et al.*, 1996; Jezewska and Bujalowski, 1996).

To date, high resolution structural information about the DnaB hexameric helicase is restricted to the small N-terminal domain of DnaB (Fass *et al.*, 1999; Weigelt *et al.*, 1999). The N-terminal region of DnaB proteins has been shown to be indispensable for the helicase activity of DnaB as well as its hexamerisation (Biswas *et al.*, 1994). The formation of a hexamer proceeds through LZ-mediated dimerisation of monomers, followed by association of dimers to form the hexamer (Biswas and Biswas, 1999). Although the crystal structure of the protein remains unknown,

protein mapping has defined specific domains involved in nucleotide hydrolysis, DNA binding and oligomerisation (Biswas and Biswas, 1999).

It has been observed that DnaB binds one strand of the DNA more tightly than the other (Higorani and Patel, 1993; Jezewska *et al.*, 1998) in a sequence-independent manner. Protein-DNA photo-crosslinking studies in the presence of non-hydrolysable nucleotide analogues show that only one subunit of the DnaB efficiently cross-linked to ssDNA (Bujalowski and Jezewska, 1995; Yu *et al.*, 1996; Jezewska *et al.*, 1996). Therefore, the ssDNA binds predominately to a single subunit of the hexamer (Bujalowski and Jezewska, 1995; Jezewska and Bujalowski, 1996). *E.coli* DnaB is believed to encircle ssDNA. FRET studies showed that DNA is bound at the central channel of the ring with a defined polarity (Jezewska *et al.*, 1998).

The *E.coli* DnaB helicase binds to ssDNA in an ATP-dependent manner. However, this activity alone is not sufficient for helicase loading at a replication origin. Thus, the entry of DnaB helicase into the unwound *oriC* depends on additional protein factors.

1.6.2 Helicase loaders

A physical interaction between DnaA and DnaB is essential for delivering of DnaB helicase to *oriC* (Marszalek and Kaguni, 1994). The loading of DnaB depends not only on DnaA binding to the *oriC* boxes, but also on DnaA binding to the open region of the origin, which is then stabilised for subsequent helicase loading (Speck and Messer, 2001).

In addition to the interaction between DnaB and DnaA, the helicase is transferred to the replication origin by the helicase loading factor DnaC (Konieczny and Helinski, 1997). DnaC is a member of the AAA+ family and it has been generally assumed that its ATPase activity plays a role in the helicase loading process (Kobori and Kornberg, 1982; Wahle *et al.*, 1989). Recent reports suggests that DnaC is a dual ATP/ADP switch protein, where ATP hydrolysis by DnaC is triggered by DnaB and ssDNA (Davey *et al.*, 2002). Surprisingly, ATP is not required for DnaC to

load DnaB onto ssDNA. DnaC exists in two states: DnaC-ATP (which expands the replication bubble formed at the origin) and DnaC-ADP (which allows DnaB to unwind DNA). The transition from the ATP-bound to the ADP-bound state depends on the substrates, ssDNA and DnaB. Each of the two states are required in turn for replication of *oriC*. DnaC-ATP is required to bind ssDNA for DNA melting at *oriC*, and DnaC-ADP is required to relieve the inhibition of DnaB by DnaC-ATP. It has been observed that ATP binding, but not hydrolysis induces a conformation of DnaC that inhibits helicase activity when it remains associated with DnaB (Davey *et al.*, 2002).

The first step of DnaB loading onto the origin is the formation of DnaB-DnaC complex. In this step, DnaC probably exist in an ATP-bound state (Neuhard and Nygaardm, 1987). At this stage, DnaC can not hydrolyse ATP, because its is not yet associated with the ssDNA from the origin. Consequently, DnaB-DnaC complex is directed to the DnaA-induced open complex at *oriC*, probably by the DnaB interaction with the origin-bound-DnaA (Marszalek and Kaguni, 1994). It has been suggested that DnaC within the DnaB-DnaC complex enlarges the replication bubble formed at the origin (Davey *et al.*, 2002). This process is a result either of the DnaB-DnaC modification of DnaA function or by direct binding of DnaC to ssDNA. In the proposed model, ATP induces DnaC to bind ssDNA tighter than in the absence of ATP. Therefore, ssDNA binding by DnaC requires ATP. At the second step of the model, DnaC assembles DnaB onto the expanded bubble. Since, DnaA is known to interact with DnaB (Marszalek *et al.*, 1994) it is possible that this interaction helps the assembly of the DnaB onto the *oriC*. It has been observed that the N-terminal deletion of DnaA prevents the DnaC-dependent assembly of DnaB onto DNA (Sutton *et al.*, 1998) (Fig. 1.9).

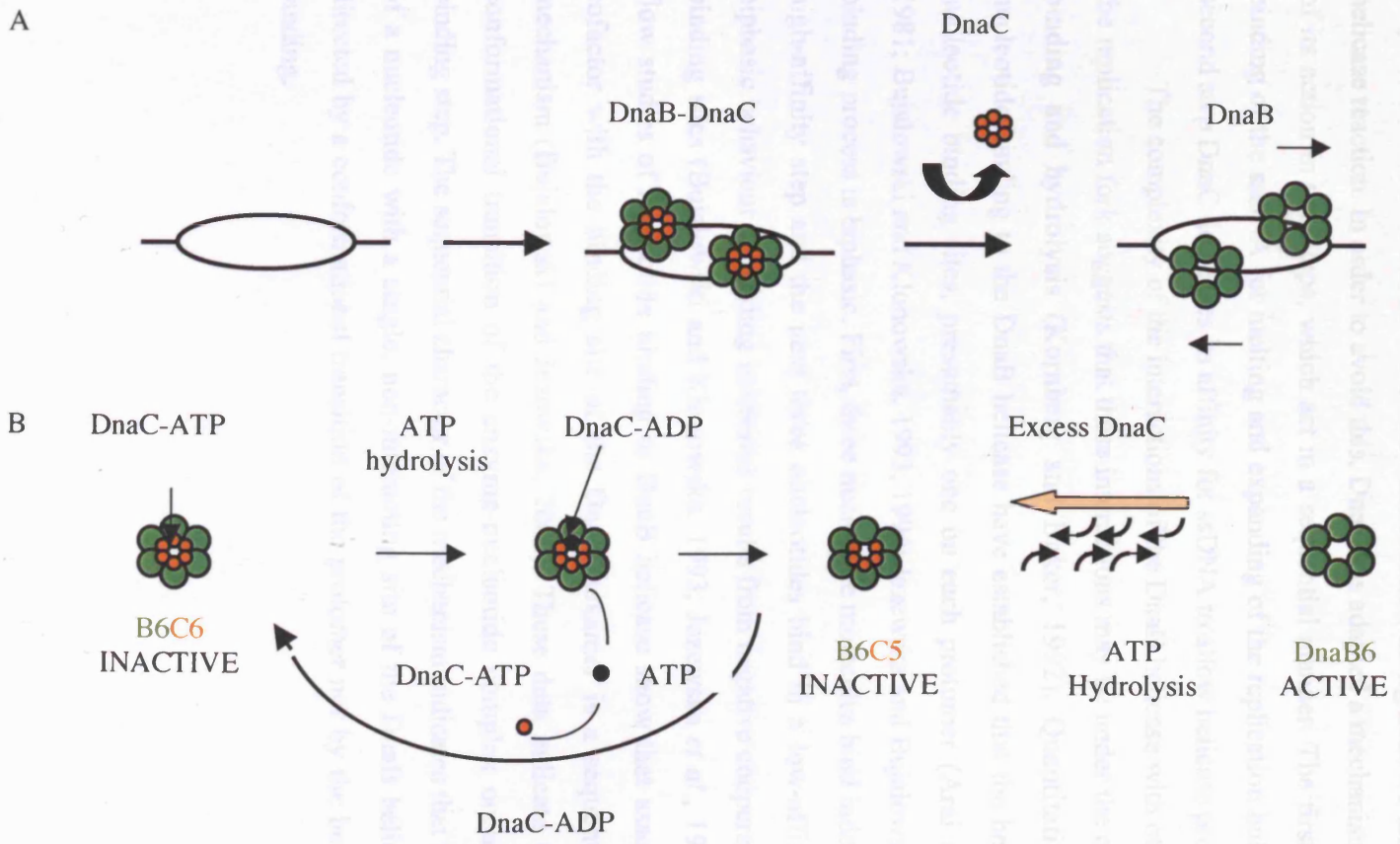


Fig. 1.9 Models of DnaC action. (A) Loading of DnaB helicase on *oriC* by the DnaB–DnaC complex. Once the helicase is recruited to the origin, DnaC lives the complex; (B) Dynamic inhibition of DnaB by DnaC is controlled by the dual ATP/ADP switch in DnaC.

DnaC uses ATP to regulate its affinity for ssDNA. The binding of ssDNA may result in inhibition of the DnaB translocation along ssDNA and prevents the helicase reaction. In order to avoid this, DnaC has adapted a mechanism of separation of its action in two steps, which act in a sequential manner. The first step involves binding of the ssDNA for melting and expanding of the replication bubble, while the second step DnaC decreases its affinity for ssDNA to allow helicase progression.

The complexity of the interactions of the DnaB helicase with other proteins in the replication fork suggests that these interactions may be under the control of ATP binding and hydrolysis (Kornberg and Baker, 1992). Quantitative studies of nucleotide binding to the DnaB helicase have established that the hexamer has six nucleotide binding sites, presumably one on each protomer (Arai and Kornberg, 1981; Bujalowski and Klonowska, 1993, 1994; Jezewska and Bujalowski, 1997). The binding process is biphasic. First, three nucleotide molecules bind independently in a high-affinity step and the next three nucleotides bind in a low-affinity step. The biphasic behaviour of binding isotherms results from negative cooperativity between binding sites (Bujalowski and Klonowska, 1993, Jezewska *et al.*, 1996). Stopped-flow studies of nucleotide binding to DnaB helicase show that association of the cofactor with the binding site of the DnaB hexamer is a sequential, multistep mechanism (Bujalowski and Jezewska, 2000). These data indicate that the major conformational transition of the enzyme-nucleotide complex occurs in the first binding step. The sequential character of the mechanism indicates that the association of a nucleotide with a single, non-interacting site of the DnaB helicase is neither directed by a conformational transition of the protomer nor by the hexamer prior to binding.

1.6.3 Helicase activity

DnaB translocates along ssDNA and unwinds duplex DNA with a distinct polarity in the 5'-3' direction (LeBowitz and McMacken, 1986; Richardson and Nossal, 1989; Kaplan, 2000). It has been demonstrated that the DnaB protein first binds to the 5' ss tail region with the DNA positioned inside the central channel. With the 3' tail positioned outside the central channel, the helicase then translocates with 5' to 3' polaity and unwinds the dsDNA. Translocation along ssDNA accomplishes unwinding if the second strand is positioned on the outside of the protein ring, even though DnaB does not contact this second strand (Kaplan, 2000) (Fig. 1.10). When DnaB tracks along ssDNA and encounters a duplex with no 3' tail, the DnaB moves off the ssDNA in order to encircle both strands and translocates along the dsDNA (Kaplan, 2000). Thus, when the 3' tail is long or bulky, the strand passes outside of DnaB ring and unwinding occurs, but when the 3' tail is absent, DnaB slides up over both strands and they can pass through the central channel of DnaB.

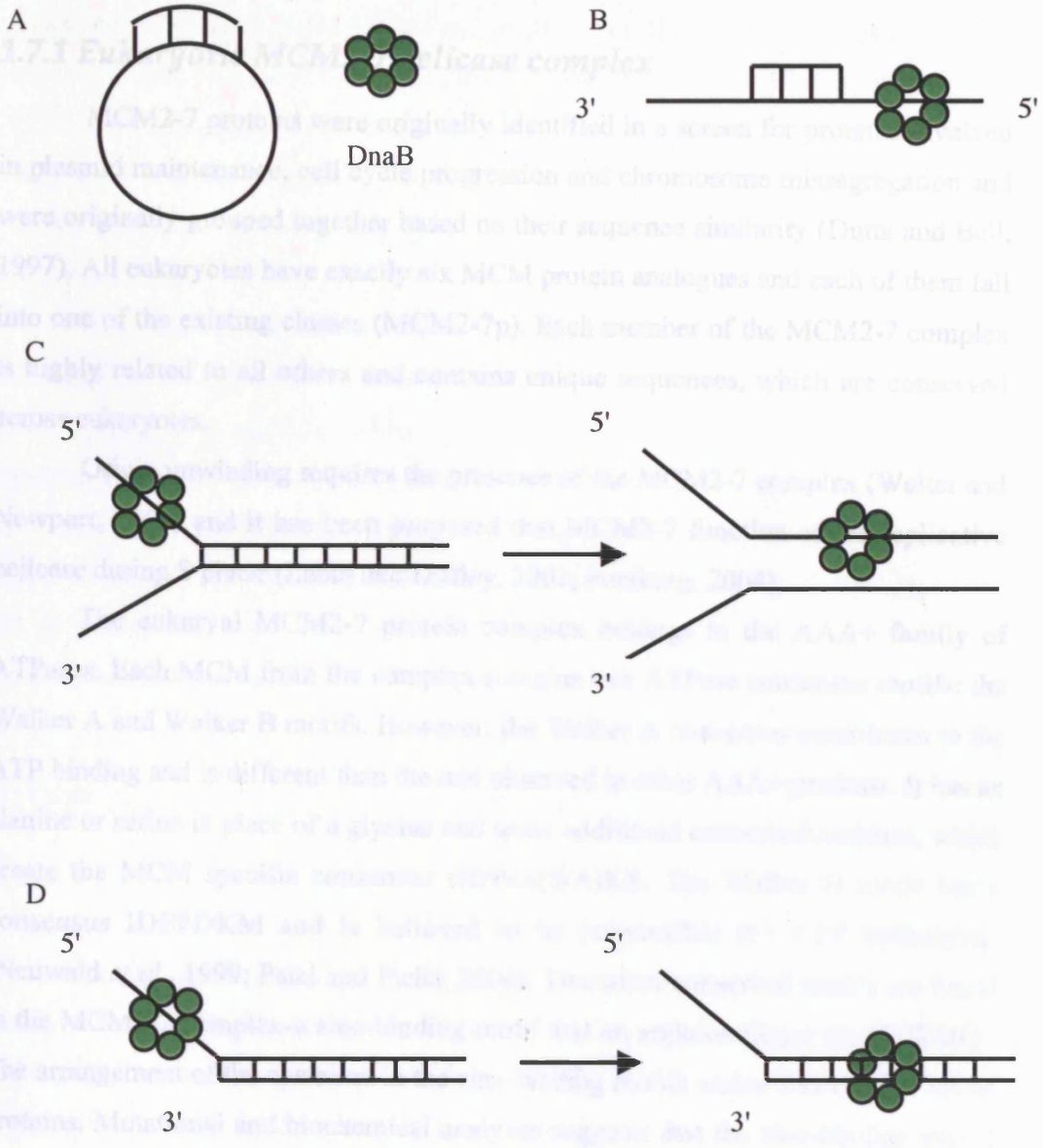


Fig. 1.10 Schematic representation of DnaB helicase behaviour on different DNA substrates. In the absence of a 5' and 3' tail, DnaB can not unwind the duplex DNA (Kaplan, 2000).

1.7 Helicase loading in eukaryotes

1.7.1 Eukaryotic MCM2-7 helicase complex

MCM2-7 proteins were originally identified in a screen for proteins involved in plasmid maintenance, cell cycle progression and chromosome missegregation and were originally grouped together based on their sequence similarity (Dutta and Bell, 1997). All eukaryotes have exactly six MCM protein analogues and each of them fall into one of the existing classes (MCM2-7p). Each member of the MCM2-7 complex is highly related to all others and contains unique sequences, which are conserved across eukaryotes.

Origin unwinding requires the presence of the MCM2-7 complex (Walter and Newport, 2000) and it has been proposed that MCM2-7 function as the replicative helicase during S phase (Labib and Diffley, 2001; Forsburg, 2004).

The eukaryal MCM2-7 protein complex belongs to the AAA+ family of ATPases. Each MCM from the complex contains two ATPase consensus motifs: the Walker A and Walker B motifs. However, the Walker A consensus contributes to the ATP binding and is different than the one observed in other AAA+ proteins. It has an alanine or serine in place of a glycine and some additional conserved residues, which create the MCM specific consensus GDPxx(S/A)KS. The Walker B motif has a consensus IDEFDKM and is believed to be responsible for ATP hydrolysis. (Neuwald *et al.*, 1999; Patel and Picha, 2000). Two other conserved motifs are found in the MCM2-7 complex-a zinc-binding motif and an arginine finger motif (SFRD). The arrangement of the cysteines in the zinc-binding motifs varies between the MCM proteins. Mutational and biochemical analyses suggests that the zinc-binding motifs are required for viability and contribute to the complex assembly (Forsburg *et al.*, 1997; Sherman *et al.*, 1998; You *et al.*, 2002).

MCMs are associated *in vivo* in a heterohexamer, although there are likely to be small amounts of single MCMs and MCM subcomplexes in the cell (Chong *et al.*, 1995; Lei *et al.*, 1996; Ishimi, 1997; Coue *et al.*, 1998; Sherman *et al.*, 1998; Lee and

Hurwitz, 2000; Schwacha and Bell, 2000; Davey *et al.*, 2002). Different MCM subunits have different relative affinities for one other and can be arranged in a series of subcomplexes (Lei *et al.*, 1996; Coue *et al.*, 1998; Sherman *et al.*, 1998; Lee and Hurwitz, 2000; Schwacha and Bell, 2000; Davey *et al.*, 2002). MCM4. MCM6 and MCM7 subunits have been found to bind tightly together and form a trimeric complex. It has been suggested that MCM2, MCM3 and MCM5 negatively regulate the active MCM4,6,7 complex (You *et al.*, 1998; Ishimi *et al.*, 1998; Lee and Hurwitz, 2001).

EM studies of the whole eukaryotic MCM complex and MCM4,6,7 subcomplex reveal a hexameric doughnut-like structure with a central cavity (Adachi *et al.*,1997; Sato *et al.*,2000. When ssDNA was added to the MCM4,6,7 structure, it resulted in ssDNA passing through the central cavity of complex. These findings raise the question of how the MCM protein ring is loaded onto DNA. It is possible that the helicase activity of the MCM4,6,7 complex relative to the inactivity of the MCM2-7 complex may be because the former complex can fluctuate between an open and closed ring to allow loading onto DNA.

1.7.2 Helicase loaders

Little is known about the mechanism of loading of the eukaryotic MCM2-7 complex to the multiple origins of replication, but the available evidence indicates that ORC, CDC6 and CDT1 act together as a clamp loader to open the MCM2-7 ring and load it onto DNA (Mendez and Stillman, 2003). Alternatively, the ORC, CDC6 and CDT1 molecules might facilitate the assembly of MCM2-7 hexamer from many subcomplexes of MCM proteins. Once MCM2-7 is loaded onto chromatin, the other components of pre-RC are dispensable for replication initiation (Donovan *et al.*, 1997; Hua and Newport, 1998; Rowles *et al.*, 1999; Maiorano *et al.*, 2000; Harvey and Newport, 2003). Therefore, the primary function of ORC, CDC6 and CDT1 in replication is to load the MCM2-7 complex.

1.7.2.1 CDC6 protein

CDC6 protein was first identified in the original screen for *S.cerevisiae* mutants with changes in the cell division cycle (Hartwell, 1973). CDC6 plays a crucial role in the assembly of the pre-RC after ORC binds the replication origin and before MCM2-7 proteins are recruited. CDC6 has been found to be highly related to ORC1p and less related to ORC4p, ORC5p and MCM2-7 (Neuwald *et al.*, 1999). As a member of AAA+ family, CDC6 contains conserved bipartite ATP binding and hydrolysis motifs described by Walker and colleagues (Walker *et al.*, 1982). Studies of mutations in the *S.cerevisiae* CDC6 ATP-binding motif suggest that ATP binding and hydrolysis are required at distinct stages during pre-RC assembly. Mutation of a conserved lysine in the Walker A motif result in a non functional protein which is unable to support cell growth (Perkins and Diffley, 1998; Wang *et al.*, 1999; Weinreich *et al.*, 1999). This mutant protein also reduces CDC6-chromatin association, prevents the formation of the pre-RC and fails to load MCM proteins onto chromatin (Perkins and Diffley, 1998; Weinreich *et al.*, 1999). In contrast to a Walker A motif mutation, glycine, but not alanine, substitution in the Walker B causes lethality when this protein is overexpressed (Perkins and Diffley, 1998). Analysis of pre-RC formation in the presence of this Walker B mutant shows that loading of MCM5p onto chromatin is blocked (Weinreich *et al.*, 1999). These studies suggest a model in which CDC6 acts as a loading factor for MCM and requires ATP binding to form an interaction with ORC and MCM at the origin (Perkins and Diffley, 1998; Weinreich *et al.*, 1999). ORC1, 4 and 5 subunits also belong to AAA+ ATPase family and electron micrographs of *S.cerevisiae* ORC complex have shown that it undergoes a conformational shift in the presence of ssDNA and ATP (Lee *at al.*, 2000), similar to what is seen with some clamp loaders (Shiomi *et al.*, 2000). Therefore, ORC and CDC6 may function together as a clamp loader complex for opening and closing MCM at origins.

1.7.2.2 CDT1 protein

CDC6 and CDT1 cooperatively promote the association of MCM2-7 to the replication origin. CDT1 from *S.pombe* was originally identified as a gene that was regulated by the *Sp*CDC10 transcription factor (Hofmann and Beach, 1994) and subsequently has been implicated as a key factor in the pre-RC assembly. *Sp*CDT1 associate with the C-terminus of *Sp*CDC6 and cooperatively promote the association of MCM protein with chromatin (Nishitani *et al.*, 2000). CDT1 is conserved in other eukaryotes including *Xenopus, Drosophila* and humans (Maiorano *et al.*, 2000; Whittaker *et al.*, 2000; Wohlschlegel *et al.*, 2000). In *S.pombe* and *X.laevis* CDC6 and CDT1 are required independently for MCM chromatin association. In *Xenopus* egg extracts ORC, but not CDC6 is required for chromatin association of CDT1 (Maiorano *et al.*, 2000). In *Drosophila,* CDT1 co-localises with ORC at sites of DNA amplification and requires ORC for its localisation (Whittaker *et al.*, 2000). CDT1 bears no sequence homology to any known protein. However, the crystal structure of a fragment of CDT1 reveals a domain that bears structural resemblance to a bacterial replication terminator protein that is believed to physically interact with the bacterial replicative helicase DnaB (Lee *et al.*, 2004). The function of CDT1 in recruitment of MCM to the chromatin remains unclear, although it is clear that CDT1 is the target of the replication inhibitor geminin.

1.7.3 Helicase activity

The six subunit MCM2-7 complex lacks helicase activity *in vitro*, while a subcomplex of MCM4,6,7 has a weak, non-processive, intrinsic 3'-5' helicase activity (Ishimi, 1997; You *et al.*,1999; Lee and Hurwitz, 2001). When a DNA with only a single exposed end was used as a helicase substrate, the MCM4,6,7 exhibited low processivity (Lee and Hurwitz, 2000). However, when a template with exposed 5' and 3' tails was used with excess of *Sp*MCM4,6,7 and *E.coli* ssDNA binding protein, more than 500bp of DNA were displaced (Lee and Hurwitz, 2001). This finding raises the possibility that the MCM complex requires a fork-like structure for

processive helicase activity. The addition of MCM2 or MCM3,5 to the MCM4,6,7 subcomplex inhibited its helicase activity (Ishimi, 1997; Lee and Hurwitz, 2000; Sato *et al.*, 2000). These facts suggest that MCM4,6,7 constitutes the core helicase, while MCM2/3/5 have a regulatory function. This model is supported by studies showing that the ATPase activity of the MCM2-7 complex requires the coordinate action of all six subunits with MCM4,6,7 involved in ATP hydrolysis and MCM2/3/5 regulating its activity (Schwacha and Bell, 2001).

Two models have been proposed for MCM4,6,7 helicase action based on the possibility of MCM ring complexes to translocate along ssDNA or dsDNA.

The first model derives from the fact that two tails (3' and 5') are required for processive unwinding and protein cross-linking indicates that two MCM4,6,7 complexes bind a forked DNA (Lee and Hurwitz, 2001). In this model, one protein ring is positioned on the 5'-tail and one ring is positioned at the 3'-tail (Kelman and Hurwitz, 2003). This suggests that the ring positioned at the 5'-tail will translocate away from the duplex region, because it only translocates with 3' to 5' polarity (Fig. 1.11). However, *in vivo* replication occurs mostly at internal regions of DNA (such as a bubble) rather than at the end of the forked linear molecule. Therefore, MCM4,6,7 will form a ring on each ssDNA strand of the bubble, then two rings will translocate in opposite directions to form two forks for bidirectional replication.

In the second model for eukaryotic MCM action (called "pump in ring" mechanism), the protein rings encircled two DNA strands and actively translocate along the duplex (Laskey and Madine, 2003; Kaplan *et al.*, 2003). This may alter the twisting of the DNA, resulting in unwinding at a distance from the MCM protein complex. However, MCM4,6,7 directly catalyses the DNA unwinding if one strand passes through the central channel and one strand passes outside the central channel. This is a more likely model for replication fork unwinding *in vivo*, because it will result in a physical separation of the two DNA strands, which can allow ssDNA to be delivered directly to the replication fork polymerase (Fig. 1.11).

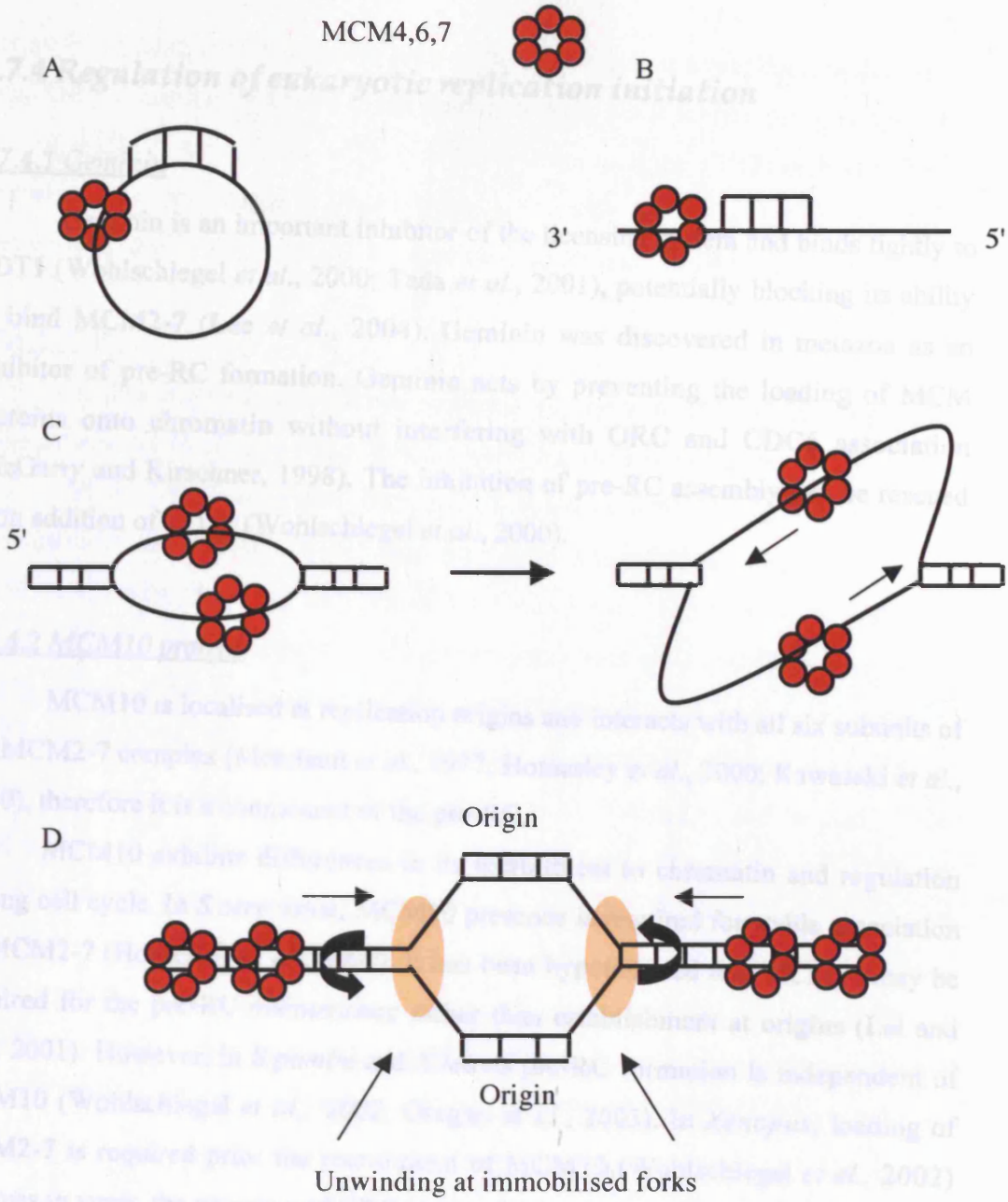


Fig. 1.11 Schematic representation of MCM4,6,7 action on different DNA substrates. (A) Substrate without ss tails; (B) Double-stranded substrate with ss tails on both sites; (C) First model for helicase activity-strict 3'-5' polarity suggest that two rings will move in separate directions; (D) "Pump in ring" model-MCM ring translocates along duplex DNA and twisted it, causing unwinding at a distant replication fork (Kaplan et al., 2003; Laskey and Madine, 2003)

1.7.4 Regulation of eukaryotic replication initiation

1.7.4.1 Geminin

Geminin is an important inhibitor of the licensing system and binds tightly to CDT1 (Wohlschlegel *et al.*, 2000; Tada *et al.*, 2001), potentially blocking its ability to bind MCM2-7 (Lee *et al.*, 2004). Geminin was discovered in metazoa as an inhibitor of pre-RC formation. Geminin acts by preventing the loading of MCM proteins onto chromatin without interfering with ORC and CDC6 association (McGarry and Kirschner, 1998). The inhibition of pre-RC assembly can be rescued upon addition of CDT1 (Wohlschlegel *et al.*, 2000).

1.7.4.2 MCM10 protein

MCM10 is localised at replication origins and interacts with all six subunits of the MCM2-7 complex (Merchant *et al.*, 1997; Homesley *et al.*, 2000; Kawasaki *et al.*, 2000), therefore it is a component of the pre-RC.

MCM10 exhibits differences in its recruitment to chromatin and regulation during cell cycle. In *S.cerevisiae*, MCM10 presence is required for stable association of MCM2-7 (Homesley *et al.*, 2000). It has been hypothesised that MCM10 may be required for the pre-RC maintenance rather than establishment at origins (Lei and Tye, 2001). However, in *S.pombe* and *X.leavis* pre-RC formation is independent of MCM10 (Wohlschlegel *et al.*, 2002; Gregan *et al.*, 2003). In *Xenopus*, loading of MCM2-7 is required prior the recruitment of MCM10 (Wohlschlegel *et al.*, 2002) whereas in yeast, the presence of ORC on the chromatin is crucial and MCM2-7 does not affect the MCM10 recruitment (Homesley *et al.*, 2000; Gregan *et al.*, 2003). In mammalian cells, the regulation of MCM10 is cell cycle dependent with the highest protein levels and chromatin association in S phase (Izumi *et al.*, 2000; 2001). However, in yeast the MCM10 protein level is constant and MCM10 is bound to the chromatin in all phase of the cell cycle (Homesley *et al.*, 2000; Kawasaki *et al.*, 2000; Gregan *et al.*, 2003). Interestingly, the differences in cell cycle regulation between

eukaryotes match the differences seen with ORC binding to replication origins. In yeast ORC is constantly bound to replication origins, while in mammalian cells, ORC is dynamically associated with the replication origins (DePamphilis, 2003). Therefore, the differences in MCM10 regulation may be crucial for the replication mechanism, timing and selection between unicellular and multicellular eukaryotes.

1.7.4.3 CDC45

CDC45 interaction with components of the pre-RC and the replicative polymerases provides a physical link between replication initiation and elongation (Hopwood and Dalton, 1996; Dalton and Hopwood, 1997, Zou *et al.*, 1997, Kukimoto *et al.*, 1999). CDC45 was originally identified in a screen for cold-sensitive mutants in cell cycle progression (Moir *et al.*, 1982). *In vitro* studies in *Xenopus* egg extracts indicate that CDC45 interacts with DNA polymerase α (pol α) and is required for its loading on chromatin (Mimura and Takisawa, 1998). In addition to that, CDC45 has been found to associate with ORC, DNA pol ε, RPA and the MCM proteins (Saha *et al.*, 1998; Kukimoto *et al.*, 1999; Zou and Stillman, 2000; Kamimura *et al.*, 2001; Uchiyama *et al.*, 2001). The interaction between CDC45 and MCM complex occurs sometime at the beginning of S phase, just before the migration of the MCM complex away from the replication origin (Zou and Stillman 1998; Aparicio *et al.*, 1999). Formation of a tight complex between CDC45 and MCM2-7 correlates closely with the helicase activation and it has been hypothesised that CDC45 function activates the helicase activity of MCM2-7 (Masuda *et al.*, 2003).

1.7.4.4 CDC7/DBF4 kinase (DDK)

Initiation of DNA replication requires the activity of two cell cycle-regulated kinases one of which is the CDC7/DBF4 kinase (DDK). CDC7 kinase activation requires binding to the non-catalytic subunit known as DBF4. The regulation of DDK activity is accomplished by the periodic stabilisation and destruction of DBF4 protein during the cell cycle (Brown and Kelly, 1999; Cheng *et al.*, 1999, Weinreich and Stillman, 1999; Ferreira *et al.*, 2000).

An abundance of genetic and biochemical data indicates that MCM2-7 proteins are the primary candidates for the target of the DDK that triggers initiation of DNA replication. *In vivo* and in vitro studies indicate that five of the six MCM proteins are phosphorylated, but the cell cycle regulation of this phosphorylation is not well characterised (Lei *et al.*, 1997; Sato *et al.*, 1997; Brown and Kelly, 1998; Kumagai *et al.*, 1999; Takeda *et al.*, 1999; 2001; Jares and Blow, 2000; Kihara *et al.*, 2000). Studies in *S.pombe* and mammalians revealed that MCM2p is the preferred substrate for DDK phosphorylation (Jiang and Hunter, 1997; Brown and Kelly, 1998; Masai *et al.*, 2000). MCM5p is the only MCM subunit that has not been observed to be a phosphoprotein.

1.7.4.5 Cyclin-dependent kinases (CDK)

Cyclin dependent kinase (CDK) not only plays a major role in driving the cell cycle, but also is required for the initiation of DNA replication (Kelly and Brown, 2000; Bell and Dutta, 2002; Masai and Arai, 2002; Kim *et al.*, 2003). Cyclin-dependent kinases have been reported to physically associate with components of the pre-RC. A series of observations suggest that CDKs are present at origins. Interactions observed between ORC, CDC6 and CDK suggested that CDK is recruited to the origin of replication. It has been reported that CDK interacts directly with ORC and CDC6 (Adams *et al.*, 1996; Chen *et al.*, 1996; Saha *et al.*, 1998; Jiang *et al.*, 1999; Peteren *et al.*, 1999; Herbig *et al.*, 2000; Romanowski *et al.*, 2000).

Many replication factors have been shown to be phosphorylated by CDK *in*

vivo and *in vitro* (Kelly and Brown, 2000). Although ORC, CDC6 and MCM proteins are all targeted by CDKs, studies of *S.pombe* ORC2 and *Xenopus* MCM4 phosphorylation suggested the existence of CDK-independent phosphorylation as cells enter S phase (Coue *et al.*, 1996; Vas *et al.*, 2001).

Even though it is clear that both CDK and DDK are absolutely required for the initiation of replication and tight association of CDC45p with chromatin, the order of function of these kinases is controversial. In *S.cerevisiae*, reciprocal shift studies suggest that CDK activity is required before DDK (Nougarede *et al.*, 2000). In contrast, depletion of DDK activity or inhibition of CDK activity in *Xenopus* egg extracts, leaded to the opposite conclusion that DDK activity is required prior to CDK activity (Jares and Blow, 2000; Walter, 2000).

1.8 Loading of the replicative helicase in archaea

1.8.1 Archaeal MCM helicase

Almost all of the archaea for which the genome is known contain a single MCM homologue (Myllykallio and Forterre, 2000). The only exceptions are *M.janaschii*, with four MCM homologues, and *M.kandleri* and *Methanosarrina acetivorans*, with two.

Like their eukaryotic homologues, the archaeal MCMs contain Walker A and Walker B as well as the zinc-binding and arginine finger (SFRD) motifs. All of the archaeal MCM homologues are more related to eukaryotic MCM4 than any of the other MCMs. Generally, each archaeal MCM protein consists of two main portions. The N-terminal region participates in protein multimerisation and ssDNA binding and may also have a regulatory function. The C-terminal portion of the protein contains the helicase catalytic domain(s) (Chong *et al.*, 2000; Fletcher *et al.*, 2003).

The oligomeric structure of the archaeal MCM complex is still not clear. Whereas the recombinant *M.thermoautotrophicum* MCM was reported to form dodecamers (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Poplawski *et al.*, 2001), the MCM from *S.solfataricus* (Carpentieri *et al.*, 2002) and *A.fulgidus* (Grainge *et al.*, 2003) were both shown to behave as homohexamers in solution. Electron microscopy (EM) analyses of *MthMCM* revealed ring-shaped hexamers (Chong *et al.*, 2000) or heptameric assemblies (Yu *et al.*, 2002) (Fig. 1.12). Recent EM analyses of different fragments of the *Mth*MCM revealed formation of single hexamers only (Chen *et al.*, 2004). Thus, it appears that only the full-length protein may be capable of forming heptamers. Although most processive helicases appear to be hexameric, it is possible that hexamers and heptamers may form under different conditions, and the ability to form rings with different symmetries may be a general property of helicases and other ring-forming proteins (Yu *et al.*, 2002). For example, the *Thermus thermophilus* RuvB protein also appears to be heptameric when not bound to DNA (Miyata *et al.*, 2000). Upon assembly around the DNA the

heptameric ring of the *Mth*MCM may lose one subunit, encircle DNA, and then close to form a ring with six subunits, becoming the active enzyme (Yu *et al.*, 2002; Kelman and Kelman, 2003). A toroidal hexameric structure was also observed for the MCM4,6,7 (Sato *et al.*, 2000) and MCM2,4,6,7 (Yabuta *et al.*, 2003) complexes purified from HeLa cells.

The crystallographic structure of the *MthMCM* N-terminal portion (residues 2-286) revealed a dodecameric architecture, with two hexameric rings juxtaposed in a head-to-head configuration (Fletcher *et al.*, 2003). This is reminiscent to the SV40 large T antigen helicase, which forms double hexamers on the origin of a SV40 DNA (Mastrangelo *et al.*, 1989). The eukaryotic MCM4,6,7 complex was also suggested to form double hexamers, but only with forked DNA structures (Lee and Hurwitz, 2001) (Fig. 1.13).

According to the crystal structure of the dodecameric *Mth*MCM, each monomer is folded into three distinct domains A, B and C (Fletcher *et al.*, 2003) (Fig. 1.12). Domain A is needed for dsDNA translocation and is believed to have a regulatory function, while domain C is necessary and sufficient for MCM multimerisation and helicase activity (Kasiviswanathan *et al.*, 2004). Domain B is the major contact with ssDNA and contains the zinc-binding motif. The zinc-binding motif is required for the head-to-head assembly of the two hexameric rings (Fletcher *et al.*, 2003). This motif is represented by three antiparallel β-strands and a further strand held in place by a zinc atom. Therefore, it does not fold into a classical zinc-finger structure, specialised for DNA binding, as was previously thought. The number of the cysteines in the zinc-binding motif varies between the MCM homologues from euryarchaeota and crenarchaeota. The euryarchaeal MCMs contain a C_4 type ($CXXCX_nCXXC$) zinc-bearing motif similar to the one found in eukaryotic MCMs. In contrast, crenarchaeal MCMs have a zinc-binding motif from a different type (HC*3*) (Poplawski *et al.*, 2001; Carpentieri *et al.*, 2002). Studies with the euryarchaeal *M.thermoautotrophicum* MCM demonstrated that the C4-type zinc-bearing motif is needed for efficient ssDNA binding, stimulation of ATPase activity by DNA, and helicase activity (Poplawski *et al.*, 2001). The role of the putative zinc-

binding motif in crenarchaeal MCMs has not been determined (Carpentieri *et al.*, 2002).

A remarkable feature of the *Mth*MCM N-terminal domain crystal structure is the presence of a long central channel whose surface has a considerably high positive charge. A three-dimensional reconstruction of the full sized *Mth*MCM structure by electron microscopy suggests that this central channel runs throughout the entire *Mth*MCM molecule (Pape *et al.*, 2003). The shape and electrostatic character of the channel strongly suggests a role in DNA binding. Since its size is much wider than the channels seen in other known helicases, it is believed to be able to accomodate both single- and double-stranded DNA (Pape *et al.*, 2003; Fletcher *et al.*, 2003). The central channel is constructed in the middle by six β-hairpin "fingers" pointing towards the center and closely linked with the ATPase domain. These fingers contain several arginine and lysine residues that are likely to participate in the interaction with the nucleic acids. Mutations of the two positively charged residues located on the tip of the hairpins in *Mth*MCM abolished binding to both ss and ds DNA. Thus, it was proposed that the ATP hydrolysis-driven shifting of the six fingers may supply a mechanism by which DNA is remodeled and translocated through the central channel during the helicase function (Fletcher *et al.*, 2003; Pucci *et al.*, 2004).

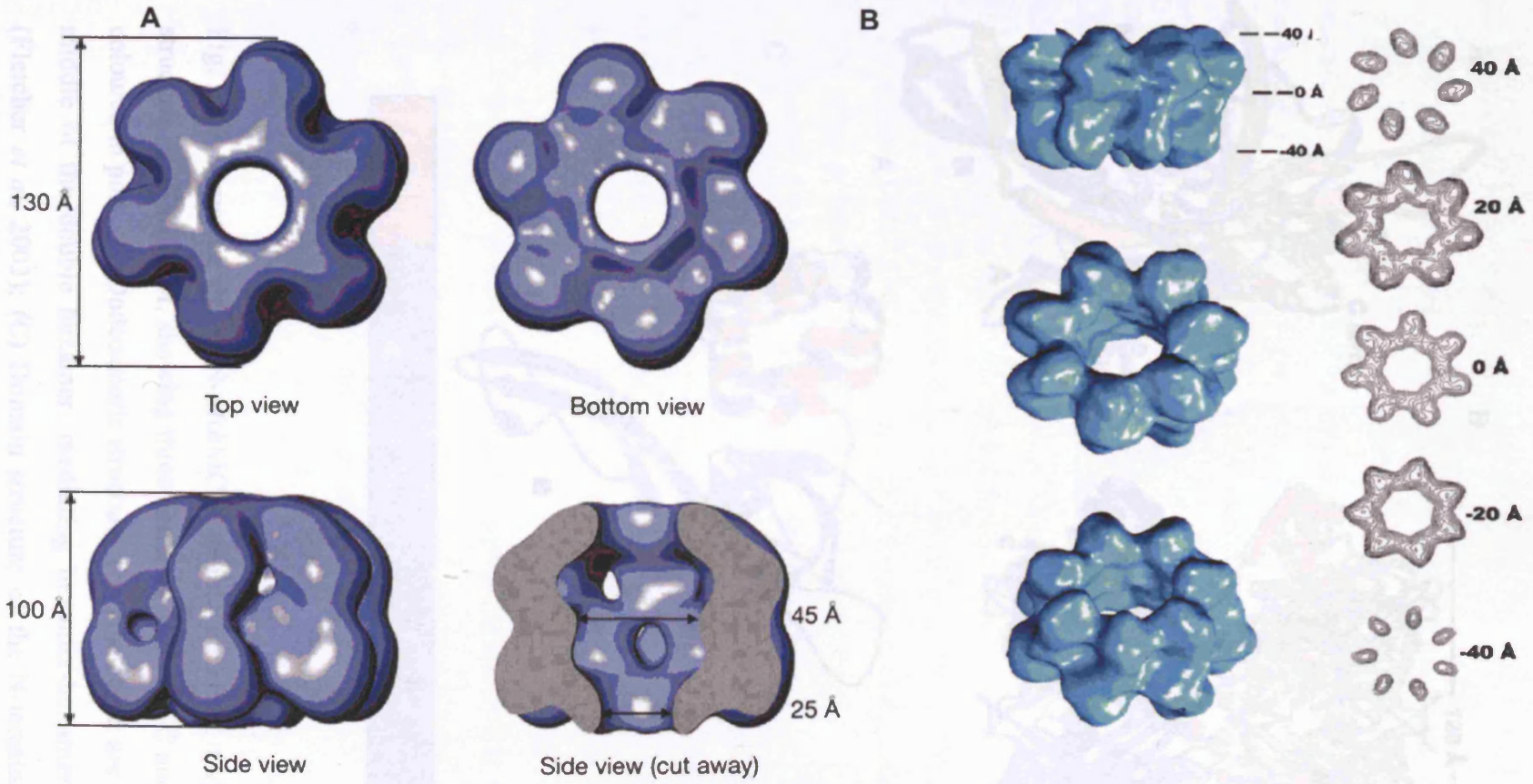


Fig. 1.12 Three-dimensional reconstruction of *Mth*MCM (Pape *et al*, 2003). (A) The protein monomers assemble into a hexameric ring around a central channel (Pape *et al*, 2003); (B) A 3D reconstruction of *Mth*MCM showing a sevenfold arrangement around a central axis (Yu *et al*, 2002)

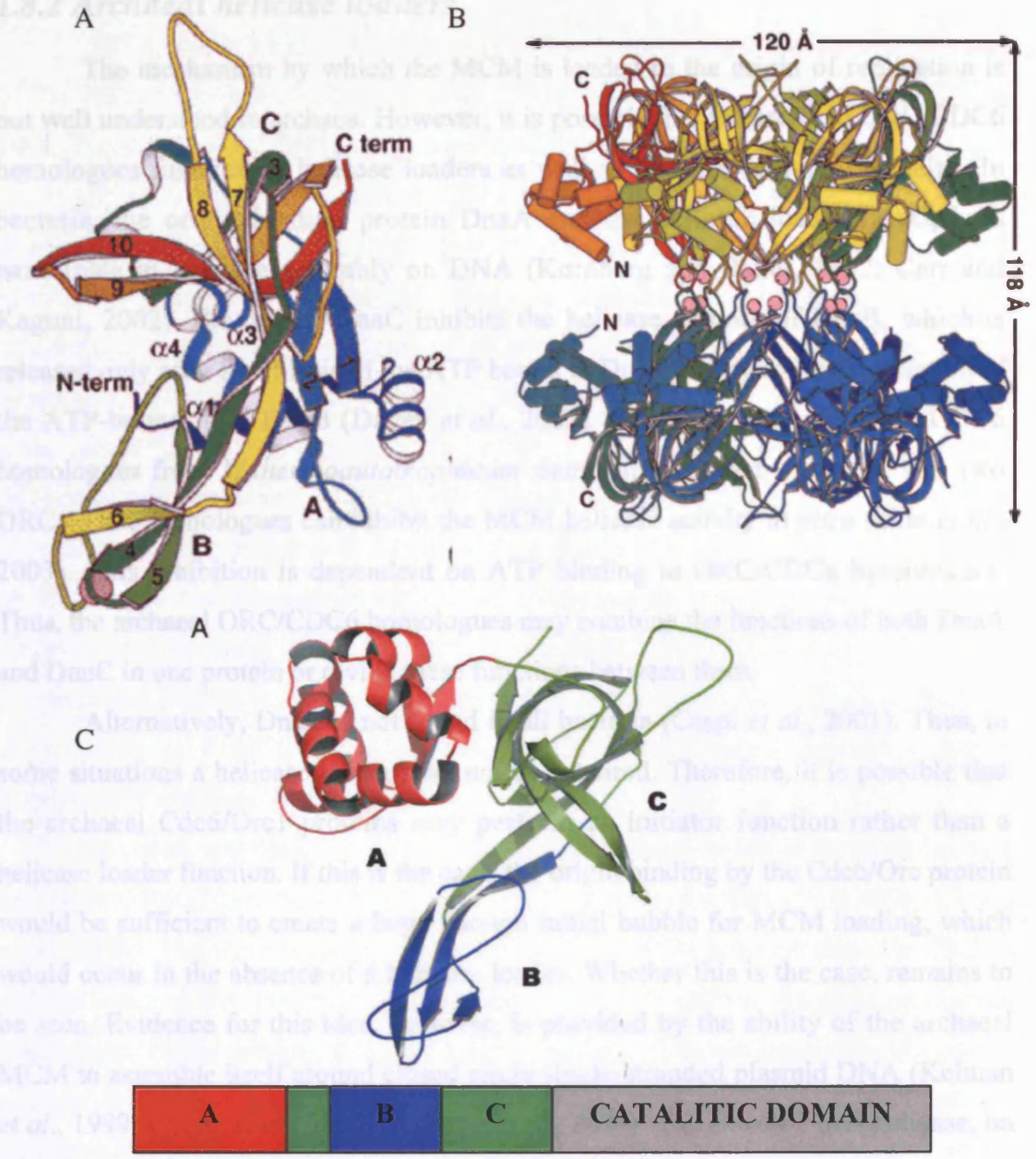


Fig. 1.13 Structure of of N-*Mth*MCM double hexamer. (A) Monomeric structure of N-*Mt*MCM, showing three domains A, B and C and a zinc atom coloured in pink, (B) Dodecameric structure. The zinc atoms are located in the middle of the double hexamer, mediating hexamer-hexamer interactions (Fletcher *et al*, 2003); (C) Domain structure of the N-terminal portion of *Mth*MCM (Kasiviswanathan *et a*l., 2004)

1.8.2 Archaeal helicase loaders

The mechanism by which the MCM is loaded to the origin of replication is not well understood in archaea. However, it is possible that the archaeal ORC/CDC6 homologues function as helicase loaders as well as origin recognition proteins. In bacteria, the origin binding protein DnaA and the helicase loader DnaC, both participate in helicase assembly on DNA (Kornberg and Baker, 1992; Carr and Kaguni, 2002). Binding of DnaC inhibits the helicase activity of DnaB, which is released only after hydrolysis of the ATP bound to DnaC, followed by dissociation of the ATP-bound from DnaB (Davey *et al.*, 2002). Recent reports with ORC/CDC6 homologues from *M.thermoautotrophicum* demonstrated that either of the two ORC/CDC6 homologues can inhibit the MCM helicase activity *in vitro* (Shin *et al.*, 2003). This inhibition is dependent on ATP binding to ORC/CDC6 homologues. Thus, the archaeal ORC/CDC6 homologues may combine the functions of both DnaA and DnaC in one protein or divide these functions between them.

Alternatively, DnaC is not found in all bacteria (Caspi *et al.*, 2001). Thus, in some situations a helicase loader may not be required. Therefore, it is possible that the archaeal Cdc6/Orc1 proteins may perform an initiator function rather than a helicase loader function. If this is the case, the origin binding by the Cdc6/Orc protein would be sufficient to create a large enough initial bubble for MCM loading, which would occur in the absence of a helicase loader. Whether this is the case, remains to be seen. Evidence for this idea, however, is provided by the ability of the archaeal MCM to assemble itself around closed circle single-stranded plasmid DNA (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000). The E. coli DnaB helicase, on the other hand, does not have this ability (Davey *et al.*, 2002).

1.8.3. MCM helicase activity

The archaeal MCM helicases possess biochemical properies similar to those of the eukaryotic enzyme (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Poplawski *et al.*, 2001; Carpentieri *et al.*, 2002).

In vitro studies whow that MCM is a stable protein, which is expressed constitutively throughout the cell cycle (Matsunaga *et al.*, 2001). Studies on the biochemical properties of euryarchaeal and crenarchaeal MCMs demonstrated similarities and important differences (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Carpentieri *et al.*, 2002; Grainge *et al.*, 2003). The biochemical data from euryarchaeal MCMs comes from the characterisation of *M.thermoautotrophicum* MCM (*MthMCM*) and *A.fulgidus* MCM (*Af*MCM). It has been observed that the ss or ds DNA binding ability of these proteins is not affected by the presence or the absence of ATP or divalent metal ions (Fletcher *et al.*, 2003; Kelman and Hurwitz, 2003; Grainge *et al.*, 2003). The ATPase activity of the euryarchaeal MCMs is stimulated by ss and ds DNA, while the ATPase activity of the crenarchaeal SsMCM is not stimulated by DNA, but is much higher than that of *Mth*MCM (Carpentieri *et al.*, 2002).

Archaeal and eukaryal MCM, unlike the bacterial helicase DnaB (Kornberg and Baker, 1992), unwind ds DNA with a 3'-5' polarity (Kelman *et al.*, 1999; Chong *et al.*, 2000). The mechanism by which the archaeal MCM helicase unwinds DNA is not well understood. However, it is possible that MCM utilises the same unwinding mechanism as the simian virus 40 large T antigen. The simian virus 40 large T antigen is another homohexameric ring helicase that works in a head-to-head structure and is required for viral replication (Simmons, 2000). The entire structure was solved, showing that the hexamer has two domains: one forming a smaller "head" that corresponds to the N-terminal domain solved for the *Mth*MCM, and a larger "body" that corresponds to the ATPase domain. These two domains rotate around one another to open and close the central channel (Li *et al.*, 2003). Whether such a model applies to eukaryal chromosomes with multiple origins, as well as to circular archaeal chromosomes with one or multiple origins, remains to be

determined. The differences in processivity and structure of the archaeal and eukaryal MCMs may reflect not only the different requirements of replicating their chromosome but also the flexibility of the MCM family of helicases under different conditions.

1.8.4 Regulation of archaeal DNA replication initiation

Besides structural and functional differences among DNA replication systems, it is almost certain that there will be differences in regulatory mechanisms and pathways as well. The similarities in archaeal and bacterial chromosomal structure suggest that they may have similar forms of regulation. However, archaea with more than one origin may provide some insights into the coordination of replication from multiple origins in eukarya.

One way to regulate initiation of replication is to block access of the initiator protein to the origin. Recent work on ORC/CDC6 homologues from *S.solfataricus* imply that one of the homologues can act as a repressor of replication (Robinson *et al.*, 2004). Temporal fluctuations of one of the CDC6/ORC initiator levels suggest that two initiator paralogues, CDC6/ORC-1 and CDC6/ORC-3 (both expressed in G1 and S phase), initiate replication at origin sites upstream of their gene sequence, whereas a third paralogue, CDC6/ORC-2 (expressed only in G2), may act as a negative regulator (Robinson *et al.*, 2004). In addition, the CDC6/ORC-2 binding sites overlap the binding sites for CDC6/ORC-1 and CDC6/ORC-3 at both functional origins, consistent with its putative role as a repressor of replication.

Although cyclin-dependent kinases do not appear to be present in archaea, additional control mechanisms involving post-translational modifications, such as phosphorylation, have been implicated in the regulation of archaeal CDC6/ORC initiators as well (Bell and Dutta, 2002; Grabowski and Kelman, 2001; Felice *et al.*, 2004).

1.9 Goals:

Archaea are not only interesting from an evolutionary point of view, but can also provide a useful model to elucidate essential conserved aspects of replication initiation. The reduced number of fundamental replication proteins suggests that archaea are an excellent model system for understanding the enzymology of eukaryotic DNA replication. Moreover, the enhanced stability of many archaeal proteins makes them an excellent subject for structural studies.

When this work started, the crystal structure of archaeal MCM was not available. The initial attempt in the lab to crystallise the full length of MCM failed. Therefore, my first goal was to find the region(s) involved in MCM hexamerisation. Knowledge of the minimal hexamerisation region may help in identification of a fragment of the MCM protein suitable for structural studies.

To date, only homologues of eukaryotic MCM, ORC and CDC6 have been found in archaeal species, but only limited information on the interaction between them is available. Homologues of essential replication factors such as MCM10, CDC45 and CDT1 are absent. However, it seems likely that orthologues of some of these proteins might exist. Thus, my second goal was to investigate how the known archaeal pre-RC proteins interact and to discover orthologues of eukaryal replication factors which have not been found.

My third goal was to examine the biochemical properties of the archaeal MCM homologue. Like their eukaryotic homologues, the archaeal MCMs contain Walker A and Walker B as well as the zinc-binding and arginine finger (SFRD) motifs. Although a lot of information exists on the main activities of the archaeal MCM helicase, the importance of the conserved motifs is not well documented.

To date, only limited studies have been reported on the biochemical characteristics of the archaeal ORC/CDC6. It has not yet been determined whether these homologues function as ORC or CDC6 or both. Therefore my fourth goal was to examine their biochemical properties and resolve this problem.

CHAPTER 2
MATERIALS AND METHODS

2.1 Chemicals

Unless stated otherwise all chemicals were obtained from Sigma company Ltd., BDH Ltd., or Aldrich. $\gamma[^{32}P]$ deoxyadenosine triphosphate, $\alpha[^{35}S]$ deoxyadenosine triphosphate and all purification column materials were from Amersham Pharmacia Biotech Ltd. Restriction enzymes and their buffers, and DNA modification enzymes were from New England Biolabs. Synthetic oligonucleotides were prepared by the Oligosynthesis department (Cancer Research UK) or by Oswel Ltd and Sigma Genosys Ltd.

2.2 Specialised computer software

CLUSTALW (Higgins, 1994) was used for multiple sequence alignments. BLAST (Altschul *et al.*, 1990) was used for comparing sequence similarity and database searching. GraphPad Prism was used for fitting steady-state kinetic data to the Michaelis-Menten equation. PredProt (Rost *et al.*, 1994) was used for secondary structure prediction.

2.3 Microbiological techniques

2.3.1 Bacterial strains

Strain	*Genotype*	*Reference*
E.coli XL1-Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F'proAB lacIqZDM15 Tn10 (Tetr)]	Bullock *et al.*, 1987
BL21 codon 2^+ Δ cat	F' ompT hsdS (r_β-m_β-) dcm gal endA Hte (argU ileX) pRI952	Weiner *et al.*, 1994

2.3.2 Yeast strains

Strain	*Genotype*	*Reference*
AH109	MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2:: GAL1 $_{UAS}$ -GAL1 $_{TATA}$ -HIS3, GAL2 $_{UAS}$ -GAL2 $_{TATA}$ -ADE2, URA3:: MEL1 $_{UAS}$ -MEL1 $_{TATA}$ -lacZ	James *et al.*, 1996; A. Holtz, unpublished
Y187	MATα, ura3- 52, his3- 200, ade2- 101, trp1- 901, leu2- 3, 112, gal4Δ, met –, gal80Δ, URA3 : : GAL1 $_{UAS}$ -GAL1 $_{TATA}$ - lacZ	Harper *et al.*, 1993

2.3.3 Cloning vectors

Vector	*Features*	*Reference*
pGADT7-AD	The pGADT7 vector expresses proteins fused to amino acids 768-881 of the GAL4 activation domain (AD). pGADT7 also contains the T7 promoter, an HA epitope tag, and a MCS. pGADT7 replicates autonomously in both *E. coli* and *S. cerevisiae* from the pUC and 2 μ ori, respectively. The vector carries an ampicillin resistance gene for selection in *E. coli* and the *LEU2* nutritional marker for selection in yeast.	Chien *et al.*, 1991
pGBKT7-BD	The pGBKT7 vector expresses proteins fused to amino acids 1–147 of the GAL4 DNA binding domain (DNA-BD). This vector replicates autonomously in both *E. coli* and *S. cerevisiae* from the pUC and 2 m ori, respectively. The vector carries the kanamycin resistance gene for selection in *E. coli* and the *TRP1* nutritional marker for selection in yeast. The T7 promoter is used for *in vitro* transcription and translation of the epitope tagged fusion protein.	Harper *et al.*, 1993

Vector	*Features*	*Reference*
pET 22b (Novagen)	5493bp expression vector for subcloning and induction. Contains IPTG inducible T7 promoter and carries ampicillin resistance.	Studier and Moffat, 1986
pET 28b (Novagen)	5368bp expression vector for subcloning and induction. Contains IPTG inducible T7 promoter and carries kanamycin resistance.	Studier and Moffat, 1986

2.3.4 Media

2.3.4.1 Bacterial media

Luria Broth (LB) was used for all cell growth (Sambrook and Russel, 2001). For solid LB medium 15g/L of agar was added. Antibiotics were used for plasmid selection.

2.3.4.2 Yeast Media

YPD broth and selective dropout (SD) media were used for yeast growth. YPD (20g/l yeast extract, 40g/l peptone and 40g/l dextrose) is a medium that supports yeast cell growth of all strains. In the selective dropout medium (6.7g/l amino acid-free yeast nitrogen base, 20g agar, 2% glucose, 100ml/l of appropriate 10Xdropout solution) a particular ingredient is lacking (Sambrook and Russel, 2001). Solid media were made by adding 20g/L agar to the liquid media.

2.4 Molecular biology

2.4.1 *Agarose gel electrophoresis*

Agarose gel electrophoresis was carried out in TBE buffer (Sambrook *et al.*, 1989). DNA was extracted and purified from agarose gels using the Qiagen gel extraction kit.

2.4.2 *DNA concentration determination*

DNA concentration was estimated spectrophotometrically at 260nm in quartz cuvettes. An optical density (OD) of 1 was taken as equal to 50μg/ml for double-stranded DNA and 33μg/ ml for single-stranded DNA.

2.4.3 *DNA end-labelling*

Radioactive labelling at the 5' end of the DNA was performed using γ^{32}P ATP (Amersham-Pharmacia) and polynucleotide kinase (New England Biolabs) according to manufacturer's instructions.

2.4.4 *DNA sequencing*

DNA sequencing was performed using the Big Dye Terminator Sequencing kit v. 3.1 (Applied Biosystems). Gel electrophoresis and analysis of the sequencing gels was carried out by the Sequencing services (Cancer Research UK).

2.4.5 *Ligation reactions*

Ligations (sticky and blunt-end) were performed with T4 DNA ligase (New England Biolabs) in the manufacturer's buffer at 16°C overnight.

2.4.6 *Plasmid purification*

Plasmid purification was performed using Qiagen mini- or maxi-prep kits according to manufacturer's instructions.

2.4.7 *Polyacrylamide gel electrophoresis (PAGE)*

Protein samples were routinely mixed with an equal volume of SDS sample buffer (4% SDS, 125mM Tris, pH 6.8, 10% β-mercaptoethanol, 20% glycerol and 0.05% bromophenol blue), heated to 95° for 5 min and loaded onto SDS-PAGE gels (Laemmli, 1970). Pre-stained broad range SDS-standard markers (New England Biolabs) were used in each instance and gels run at 220 V using the Mini-Protean gel apparatus (BioRad). Gels were stained either by using Coomassie blue or the Silver Stain Plus kit (BioRad), or were transferred onto nitrocellulose filters for further Western blotting analysis.

2.4.8 *Restriction digest*

All restriction digests were performed according to the manufacturer's instructions at 37°C for 2 hours.

2.4.9 *Site-directed mutagenesis by PCR*

All mutations were introduced using QuickChange^{TM} Site-Directed Mutagenesis Kit. PCR reactions for site-directed mutagenesis were performed according to the manufacturer's instructions (Stratagene) (Fig 2.1) with either one or two base changes being introduced with one mutagenic primer. The primers were homologous to the target sequence (apart from the introduced mutations). Sense and antisensense primers are shown in coding triplets with mutated bases underlined in Table. 2.1.

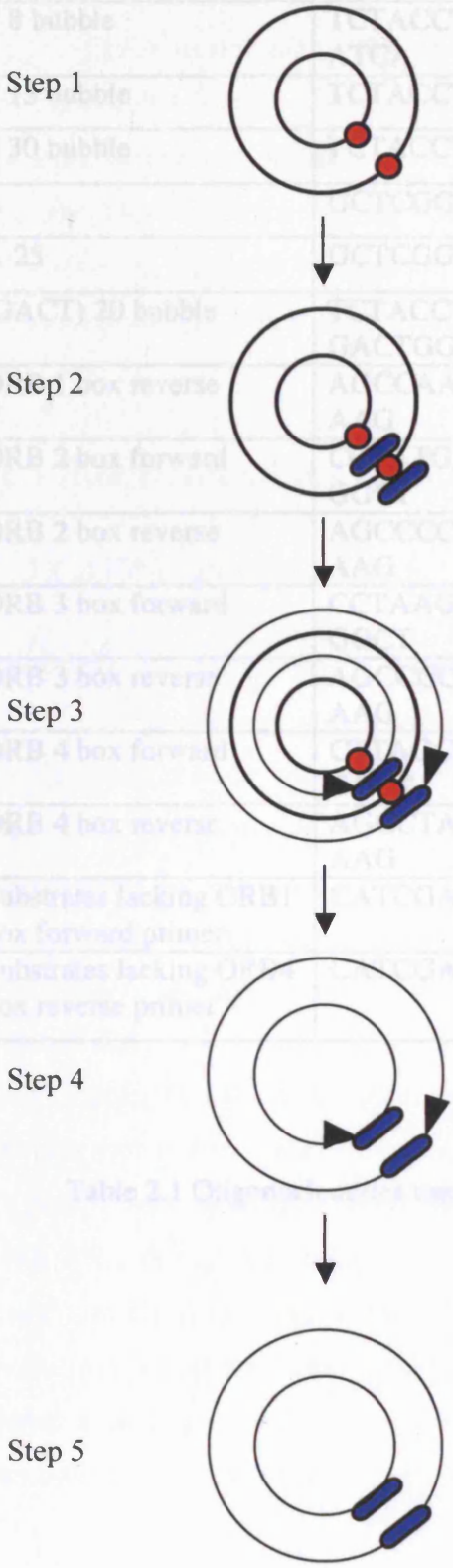


Fig. 2.1 Principle of the QuickChange Site-Directed Mutagenesis method.

Step 1: Plasmid preparation generates a gene with a target site for mutation within the plasmid.

Step 2: Temperature cycling to denature the plasmid and anneal the oligonucleotide primers containing the desired mutation.

Step 3: PCR using Pfu Turbo DNA polymerase to incorporate the mutagenic primers, resulting in nicked circular strands.

Step 4: Digestion of parental DNA template with DpnI.

Step 5: Transformation of nicked circular dsDNA into XL1-Blue supercompetent cells; the cells repair the nicks in the mutated primer. (Adapted from the Stratagene Instruction Manual)

DNA substrate	Top strands
T 8 bubble	TCTACCTGGACGACCGGGT$_{(8)}$GGGCCAGCAGGTCC ATCA
T 15 bubble	TCTACCTGGACGACCGGGT$_{(15)}$GGGCCAGCAGGT
T 30 bubble	TCTACCTGGACGACCGGGT$_{(30)}$GGGCCAGCAGGT
A	GCTCGGTACCCGGGGATCCTTAGA
A-25	GCTCGGTACCCGGGGATCCTTAGAT$_{(25)}$
(GACT) 20 bubble	TCTACCTGGACGACCGGGGACTGACTGACTGACT GACTGGGCCAGCAGGTCCATCA
ORB 1 box reverse	AGCCAAGGACCCCTCCGTTTCCTGTGGAGCATCCT AAG
ORB 2 box forward	CCTATGATGCTCCAGAGGAACCTGGGGGGTGTGG GGCT
ORB 2 box reverse	AGCCCCACACCCCCCAGGTTCCTCTGGAGCATCAT AAG
ORB 3 box forward	CCTAAGAGGATCCAGCGGAAAACAGGGGGTGTGG GGCT
ORB 3 box reverse	AGCCCCACACCCCCTGTTTTCCGCTGGATCCTCTT AAG
ORB 4 box forward	CTTAGCCGTCTCCACAGGAAACGGAGGGGTCGTA GGCT
ORB 4 box reverse	AGCCTACGACCCCTCCGTTTCCTGTGGAGACGGCT AAG
Substrates lacking ORB1 box forward primer	CATCGAATTCCGGCAACCACGCCACG
Substrates lacking ORB4 box reverse primer	CATCGAATTCCGGCGAGCTGACACACTAAG

Table 2.1 Oligonucleotides used for preparation of DNA binding substrates.

2.4.10 Transformation by electroporation

Electrocompetent cells were prepared by Nicola Cook as described by Sambrook and Russel, 2001. Typically, 2µl of a10µl ligation reaction, or 0.5µl of a plasmid midiprep was added to 20µl of electrocompetent cells in an electroporation cuvette. This was chilled on ice before a single 2500 volt shock was applied across the cuvette using Celljet electroporation equipment (Flowgen). 1ml of LB was added to the cells immediately, which were placed at 37°C for 1 hour before plating.

2.5 Biochemical techniques

2.5.1 ATP hydrolysis assay

2.5.1.1 Ammonium molybdate-malachite green assay

ATP hydrolysis was determined by using an ammonium molybdate-malachite green based method for assaying nanomolar quantities of inorganic phosphate as described previously (Lanzetta *et al.*, 1979). Immediately before performing experiments the malachite green indicator solution was prepared by mixing 3 volumes of 0.045% (w/v) malachite green with 1 volume of 5% (w/v) ammonium molybdate in 4M H_2SO_4. The mixture was stirred for 20 minutes and then passed through a 0.22-micron filter before use. ATPase buffer containing different concentration of $MgCl_2$ (see Chapter 4) and 25mM Hepes, pH 7.5 was used. The reactions were carried out at various protein concentrations (see Chapter 5) and saturated amounts of single-stranded or double-stranded DNA. 50µl samples from the reaction were removed and terminated with 0.9ml of indicator solution at various time points. After 1 minute 100µl of 34% (w/v) sodium citrate was added. The colour was allowed to develop for 20 minutes before measuring absorbance at 660nm. This was correlated to phosphate concentration using a standard curve. Appropriate controls with no protein and no ATP were performed and experimental measurements altered accordingly.

2.5.1.2 Thin layer chromatography (TLC)

The ATPase rates of *A.pernix* ORC-1 and ORC-2 proteins were determined by analysing the turnover of radioactively labelled ATP. The 20µl reaction mixture, contained 25mM Hepes, pH 7.5, 25mM $MgCl_2$, 6µM [γ^{32}P] ATP and 3µM *A.pernix* ORC-1 or ORC-2 proteins. Where indicated, reactions contained 3µM protein and/or 6µM single-stranded or double-stranded ORB1 DNA. After incremental times of incubation at 75°C, 1µl aliquots were spotted onto polyethyleneimine-cellulose thin-layer plates (Merk), that were developed in 0.5M LiCl and 1M formic acid (pH 3.0) as the liquid phase and analysed by phosphorimaging.

2.5.2 *DNA binding*

2.5.2.1 Electrophoretic mobility shift assay (EMSA)

Analysis of single-stranded or double-stranded DNA binding was performed using the electrophoretic mobility shift assay technique. Oligonucleotides (Oswel or Cancer Research UK oligonucleotide synthesis service) were labelled using T4 polynucleotide kinase (New England Biolabs) and [γ^{32}P] ATP (Amersham). Unincorporated nucleotide was removed using a Bio-Rad P6 spin column. For double-stranded substrates the labelled oligonucleotide was annealed to a 2-fold molar excess of a non-labelled complementary strand by cooling at 1°C/min from 95°C to 25°C in a PCR machine. Unless stated otherwise, indicated protein concentrations (see Results and Discussions) were incubated with 2.5nM radioactively labelled DNA in a buffer (25mM Hepes, pH 7.5, and 25mM $MgCl_2$) at room temperature for 10 min. The samples were then run on a 6% native polyacrylamide gel to separate the protein-DNA complex from the free DNA. Gels were dried and analysed with a Storm phosphorimager and ImageQuant software. DNA substrates used for binding experiments are shown on Table 2.1.

Ori1 DNA was amplified by PCR from *A. pernix* genomic DNA using the following oligonucleotides: forward 5'TTATTAAGCTAGAAAGCGGGTCCAG3'

and reverse 5'TGGGGGTGGATGCGCTTC3'. This DNA was cloned into the SmaI site of pUC19 to give the plasmid pOri1 (constructed by Ian Grainge). Radiolabelled *Ori1* DNA was produced by PCR using one of the above oligonucleotides 5' end-labelled with T4 PNK and γ ^{32}P ATP and the corresponding non-labelled oligonucleotide with pOri1 as a template. The extended product was run on a 6% acrylamide gel in 1xTBE (90mM Tris borate, pH 7.5, 1mM EDTA) and the wet gel exposed to film. The extension product was excised and eluted at 30°C in 10mM Tris, pH 7.5, 1mM EDTA, 100mM NaCl, and subsequently passed through a 0.22μm filter to remove acrylamide fragments, followed by ethanol precipitation.

2.5.2.2 Analysis of binding data

DNA binding experiments were repeated in triplicate and the average and standart deviations from the three data sets were ploted. The data were analysed using GraphPad Prism, and where appropriate were fitted to a simple independent binding sites model (Clark and Engel, 1996):

$$\alpha=\alpha_M(([P]_T+[DNA]_T+K_{d(app)})-(([P]_T+[DNA]_T+K_{d(app)})^2$$

$$-4.[P]_T.[DNA]_T)^{0.5})/(2.[DNA]_T),$$

where α is the percentage of DNA bound, α_M is the percentage of DNA bound at saturation, $.[P]_T$ is the total protein concentration, $.[DNA]_T$ is the total concentration of DNA and $K_{d(app)}$ is the apparent dissociation constant. This equation describes the association of a protein with a ligand at equilibrium when each binding event is entirely independent and in such cases can be used to derive the dissociation constant for the interaction.

Sigmoidal curves were fitted to the Hill equation and used to determine the Hill coefficient *n*: $\alpha=\alpha_M(([P]^n_T)/([P]^n_T)+K)$ (Clark and Engel, 1996).

2.5.3 *Strand displacement (helicase) assay*

Helicase time course reactions were carried out at 55°C in a reaction buffer containing 25mM Hepes, pH 7.5, 5mM $MgCl_2$, 10mM ATP and 10% glycerol using 1nM DNA substrate and different protein concentrations (see Chapter 4). After 2min pre-incubation at 55°C, the reactions were started by adding ATP and later stopped by adding stop buffer (0.1% w/v SDS, 40mM EDTA, 8% v/v glycerol, 0.1% w/v bromophenol blue). Protein titration experiments were performed exactly as above but using various quantities of protein. The effect of *A.pernix* ORC-1 and ORC-2 proteins and their deletion mutants was investigated by adding these proteins to the reaction mixture in the concentrations indicated in the text (see Chapter 5).

Displaced oligonucleotide was separated from annealed oligonucleotide by electrophoresis through a 10% non denaturing polyacrylamide mini-gel at a constant voltage at 150V. Gels were then dried and exposed to a phosphorimage cassette. Quantitive analysis was performed with a phosphorimager and ImageQuant software. Strand displacement is measured as a percentage by determining the counts associated with displaced oligonucleotide and the total counts associated with both annealed and displaced oligonucleotide in each lane.

2.5.4 *DnaseI footprinting*

Footprinting reactions were carried out in 20µl volume reactions containing 20mM Tris-acetate (pH7.5), 10mM magnesium acetate, 100mM NaCl, 0.1mg/ml BSA, 5µg/ml salmon sperm DNA, 1mM ATP and 50fmol labelled substrate DNA. Protein was added to concentrations stated and allowed to bind for 5 minutes at 37°C. 0.1U of DnaseI was then added and the incubation continued for 70 seconds at 37°C. Reactions were stopped by addition of 2µl 1% SDS, 200mM EDTA and then extracted with phenol/chlorophorm followed by ethanol precipitation. DNA was recovered by centrifugation, washed with 70% EtOH, dried and resuspended in

loading dye (8M urea, 0.01% bromphenol blue, 0.01% xylene cyanol). DNA was then electrophoresed in 6% acrylamide gels containing 8M Urea at 75W.

Maxam and Gilbert (A/G) sequencing ladders were produced in a variation of the method described by Sambrook (Sambrook and Russell, 2001). Labelled PCR product was taken and diluted to 10μl volume. 4μl of 2mg/ml salmon sperm DNA was then added followed by 3μl of 88% formic acid, and the reaction incubated for 5 minutes at 37°C. 100μl of 10% piperidine (v/v) was added and incubated at 90°C for 18 minutes. Reactions were then extracted twice with 1ml water-saturated butanol followed by precipitation with 1ml 100% butanol. DNA was then dried and resuspended in urea loading dye.

2.5.5 *Phosphorylation assay*

Phosphorylation assays for *A.pernix* ORC-2 protein were performed at 75°C for 10min and for ORC-1 protein-for 5 hours. The proteins were incubated in a reaction mixture (10μl volume) containing 1.4μM [γ-^{32}P] ATP in 25mM Hepes, pH 7.5, 25mM $MgCl_2$, in the presence or absence of 5.6μM single-stranded or double-stranded DNA (ORB1). Unless otherwise stated, *A.pernix* MCM protein was used at concentration 0.46μM/hexamer, *A.pernix* ORC-2 and ORC-1 proteins and their deletions mutants were used at concentration 2.8μM per reaction. The proteins were then separated by 12% SDS PAGE gel and the ^{32}P-labelled bands were detected using a phosphorimager.

To detect possible amino acid residues of *A.pernix* MCM protein phosphorylated by *A.pernix* ORC-2, an MCM peptide membrane array was used (prepared by Nicola O'Reilly, Protein and Peptide Chemistry Laboratory, CR UK). The phosphorylation reaction was done as described above in total volume of 3ml. The membrane containing the *A.pernix* MCM peptides was covered with the reaction mixture and incubated for 30 minutes in a 75°C water bath. The membrane was washed five times with 10mM EDTA, followed by washing with 3% SDS and then

with 1.5M NaCl. To remove any non-specifically bound ATP, the membrane was washed overnight at room temperature with 500mM phosphoric acid. The membrane was then washed first with water, then with ethanol, dried, and exposed overnight to a phosphorimager cassette.

2.6 Protein purification

2.6.1 A.fulgidus MCM protein and its deletions

2.6.1.1 Expression and purification of wt A. fulgidus MCM

The *A.fulgidus* mcm gene (AF0517) was amplified by PCR from genomic DNA and ligated into pET22b vetor (Novagen) by Ian Grainge. The construct was transformed into *E.coli* BL21 $2^{+}\Delta$cat electrocompetent cells, bearing a plasmid encoding rare t-RNAs (pSJS1240, gift from S. J. Sandler). Individual colonies were selected on LB plates with added ampicillin and streptomycin. A single colony was used to inoculate three litres of LB culture, containing 100μg/ml ampicillin and 100μg/ml streptomycin. The culture was grown with shaking at 37°C until OD_{600} reached 0.5 and then protein expression was induced with 1mM IPTG for 4 hours. Cells were harvested by centrifugation at 5,000g. The pellets were resuspended in 30ml of resuspension buffer (50mM Tris, pH7.5, 1mM EDTA, 20mM NaCl, 20% sucrose, 1mM DTT) and frozen at -80°C until required. The cells were thawed and lysed by passaging through a homogeniser (Stansted Fluid Instruments) at 15,000psi (1psi≈6.9kPa) in the presence of 140ng/ml pepstatin, 400ng/ml leupeptin, 1mM PMSF. Cell debris was removed by centrifugation at 20,000g at 4°C. The protein in the supernatant was heat precipitated at 55°C for 10min and collected by centrifugation. The supernatant was precipitated by the addition of 0.44g/ml ammonium sulphate, and collected by centrifugation at 20,000g for 20min at 4°C. The pellet was resuspended in a buffer containing 50mM Tris, pH 7.5, 1mM EDTA 1mM DTT, so that the conductivity was 10mS/cm and loaded onto a Cibacron Blue

Sepharose column. The column was washed with buffer A (50mM Tris, pH 7.5, 50mM NaCl, 1mM EDTA, 1mM DTT) and the protein was eluted in buffer B (50mM Tris, pH 7.5, 500M NaCl, 1mM EDTA, 1mM DTT). The eluant was diluted to a conductivity of 10mS/cm and loaded onto a 5ml HiTrap heparin column (Pharmacia). Protein was loaded in buffer A and eluted with a linear gradient of buffer B. The eluant was run on a HiLoad 16/60 Superdex 200 gel filtration column (Pharmacia) equilibrated in 50mM Tris, pH 7.5, 150mM NaCl, 1mM EDTA, 1mM DTT. The protein fractions from the resulting peak were collected together and diluted to a conductivity 10mS/cm. The protein was then loaded onto a 10ml MonoQ column (Pharmacia) in buffer A and eluted with a linear gradient to buffer B. The purified MCM protein was supplemented with glycerol to 10% and stored in aliquots at –70°C. Yields were typically 2–3mg of protein per 1L of culture.

2.6.1.2 Expression and purification of 1-399 A.fulgidusMCM

Deletion mutants of 1-399 amino acids were amplified by PCR and cloned into the NdeI-BamHI sites of pET22b vector using the following oligonucleotides: NdeI-5'GGGAATCATCATATGGGTATAAGCAGTCCGGCAC3' and BamHI-5'AGCCCGGATCCCTATTAAAGTTTGCTTACCAATTTG3'. The plasmid was transformed into *E.coli* BL21 2^{+} Δcat electrocompetent cells and expressed using the same strategy as used for the full-length *A.fulgidus* MCM. The protein was loaded onto a HiTrap Chelating HP column, charged with $ZnCl_2$. The column was washed with buffer A (see above) and eluted with step gradient of buffer I (50mM Tris pH7.5, 50mM NaCl, 500mM Imidazole). The sample was next loaded onto a Hydroxylapatite column at a conductivity of 10mS/cm. The column was washed and the protein eluted with a gradient of buffer K (1M $KHPO_4$, pH 6.8, 1mM DTT). The protein was concentrated to 3ml and loaded onto a Superdex 200 gel filtration column (Pharmacia) as described above for the full-length.

2.6.2 Expression and purification of A. pernix proteins

2.6.2.1 Expression and purification of A. pernix MCM protein and its mutants

The *A.pernix* gene (APE0188) was amplified by PCR from genomic DNA and ligated into pET28b vetor (Novagen) by Ian Grainge. The construct was transformed into *E.coli* BL21 2^+ Δcat electrocompetent cells. Individual colonies were selected on LB plates with added ampicillin and streptomycin. A single colony was used to inoculate 3L of LB culture, containing 100μg/ml ampicillin and 100μg/ml streptomycin. The culture was grown with shaking at 37°C until OD_{600} reached 0.5 and protein expression was induced with 1mM IPTG for 4 hours. Cells were harvested by centrifugation at 5,000g. The pellets were resuspended in 30ml of resuspension buffer (50mM Tris, pH 7.5, 1mM EDTA, 20mM NaCl, 20% sucrose, 1mM DTT) and frozen at -80°C until required. The cells were thawed and lysed by passaging through a homogeniser (Stansted Fluid Instruments) at 15,000psi in the presence of 140ng/ml pepstatin, 400ng/ml leupeptin, 1mM PMSF. Cell debris was removed by centrifugation at 20,000g. The protein in the supernatant was heat precipitated at 55°C for 10min and centrifuged. The supernatant was precipitated by the addition of 0.44g/ml ammonium sulphate, and collected by centrifugation at 20,000g for 20min at 4°C The pellet was resuspended in a buffer containing 50mM Tris, pH 7.5, 1mM EDTA 1mM DTT, so that the conductivity was 10mS/cm and loaded onto a Source30Q column equilibrated with buffer A (50mM Tris, pH 7.5, 50mM NaCl, 1mM EDTA, 1mM DTT). *A.pernix* MCM protein was eluted with a linear gradient of buffer B (50mM Tris, pH 7.5, 1M NaCl, 1mM EDTA, 1mM DTT). Saturated ammonium sulphate solution was added to the eluant to bring the final concentration of ammonium sulphate to 1M. The protein was then loaded onto a Source Isopropyl column (Pharmacia) in 50mM Tris, pH 7.5, 1mM EDTA, 1mM DTT, 1M ammonium sulphate, and eluted with a linear gradient to buffer B. Eluate was run on a HiLoad 16/60 Superdex S200 gel-filtration column (Pharmacia) equilibrated in a buffer containing 50mM Tris, pH 7.5, 500mM NaCl, 1mM EDTA, 1mM DTT.

pET28b constructs carrying C152A, K349E, S350E, E408A and R476A mutations of the mcm gene were transformed into *E.coli* BL21 Δcat 2+ electrocompetent cells and purified the same way as the wt A.*pernix* MCM protein.

PET28b construct containing N-terminal deletion of 256 amino acids of A.*pernix* MCM was made by Ian Grainge. The same expression and purification procedure was followed as for the wild type protein.

2.6.2.2 Expression and purification of A.pernix ORC-2 protein and its domains

ORC-2 and its domains and apo forms were prepared as described previously (Singleton *et al.*, 2004).

2.6.2.3 Expression and purification of A.pernix ORC-1 protein and its domains

i) A.pernix ORC-1 protein

pET22b plasmid carrying ORC-1 (construct made by Ian Grainge) was transformed into *E. coli* strain BL21 (DE3). Cells were grown in LB at 37°C for 3 hours to an OD_{600} of 0.5 and the protein expression was induced with 1mM IPTG at 30°C for 4 hours. Cells were harvested by centrifugation and resuspended in a buffer containing 50mM Tris, pH 7.5, 500mM NaCl, 1mM EDTA, 140ng/ml pepstatin, 400ng/ml leupeptin, 1mM PMSF, 1mM DTT. Cells were disrupted by passaging through a homogeniser (Stanstead Fluid Instruments) at 15,000psi. The lysate was clarified by centrifugation at 20,000g for 20min, followed by heat treatment at 55°C for 10min and another centrifugation step. 0.4g/ml ammonium sulphate was added to the resulting supernatant and the centrifugation was repeated as above. The pellet was then resuspended in a buffer containing 50mM Tris, pH 7.5, 1mM EDTA, 1mM DTT so that the conductivity was around 25mS/cm. The protein was then loaded onto a Cibacron blue Sepharose column equilibrated with buffer A (50mM Tris, pH 7.5, 250mM NaCl, 1mM EDTA, 1mM DTT), washed with several column volumes of a

buffer A and then eluted with buffer B (50mM Tris, pH 7.5, 1M NaCl, 1mM EDTA, 1mM DTT). The resultant protein was diluted in buffer A to a conductivity of 25mS/cm and loaded onto a 5ml HiTrap heparin column (Pharmacia) equilibrated in a buffer A. The column was washed with several volumes of buffer A and then eluted with a linear gradient of buffer B. The eluate was passed through a HiLoad 16/60 Superdex 75 gel-filtration column (Pharmacia) equilibrated in 50mM Tris, pH 7.5, 1M NaCl, 1mM EDTA, 1mM DTT. Yield was approximately 3mg of protein per 1litre of cells grown up.

The apo form of the protein was prepared as described for ORC-2 (Singleton *et al*, 2004) by denaturing in 6M guanidine chloride, re-folding in 1M magnesium chloride, followed by binding to a heparin column and elution with a linear gradient to 1M NaCl.

ii) Domain I+II ORC-1

Domain I+II was amplified by PCR using as a forward primer 5'GGGAATTACCATATGGCGGATCCCTTGGAGG3' and 5'AGCCCGAATTCCTATTACCTATCCCTCTCAATCTCAGC3'as a reverse primer.

The product was cloned and expressed as described for the wild type (see above). Purification proceeded as for the wild type up to the purification step on a Cibacron blue-Sepharose column. Saturated ammonium sulphate solution was added to the protein to bring the final concentration of ammonium sulphate to 1M. The protein was then loaded onto a low-substitution phenyl column (Pharmacia) in 50mM Tris, pH 7.5, 1mM EDTA, 1mM DTT, 1M ammonium sulphate, and eluted with a linear gradient of buffer B. Eluate was run on a HiLoad 16/60 Superdex 75 gel-filtration column (Pharmacia) equilibrated in 50mM Tris, pH7.5, 500mM NaCl, 1mM EDTA, 1mM DTT.

iii) Domain III ORC-1

Domain III of ORC-1 was cloned into the NdeI and EcoRI sites of pET28a vector to yield an N-terminal His-tag fusion protein (construct made by Ian Grainge). Growth and expression was carried out as for the full-length protein. Domain III protein was purified on a 5ml HiTrap chelating column charged with Ni^{2+} (Pharmacia). The column was washed in buffer I (50mM sodium phosphate, pH 6.8, 500mM NaCl and 10mM imidazole) and eluted with a linear gradient of buffer Q (50mM sodium phosphate, pH 6.8, 500mM NaCl and 250mM imidazole). The pooled fractions were then diluted in 10mM sodium phosphate, pH 6.8 to a conductivity of 15mS/cm and then loaded onto a HiTrap heparin column in buffer R (25mM sodium phosphate, pH 6.8, 50mM NaCl, 1mM EDTA) and eluted with a linear gradient of buffer R + 1M NaCl. Pooled fractions were then run on a S75 gel filtration column in 25mM sodium phosphate, pH 6.8, 1mM EDTA, 1mM DTT, 500mM NaCl. The yield was approximately 5mg protein per litre of culture.

iiii) R198A ORC-1

Site directed mutagenesis was carried out by PCR, using QuickChange™ Site-Directed Mutagenesis Kit (Stratagene). pET22b vector carrying ORC-1 was used to introduce a substitution into the coding sequence of ORC-1. The primers for the codon substitution underlined in each case were as follows: for Arginine 198 to alanine (R198A), 5'GAAAACCTGGAGCCCGCGGTCAAGAGCAGCCTC3' and 5'GAGGCTGCTCTTGACCGCGGGCTCCAGGTTTTC3'. The PCR program used for first and second step reactions was: 18 cycles (95°C for 0.30 minutes, 55°C for 1 minute, 68°C for 7 minutes) and a final extension step at 68°C for 10 minutes. All reaction conditions were as described by the manufacturer (Stratagene). The R198A mutant of ORC-1 was expressed and purified as described for the wild type.

2.7. Yeast two hybrid analysis

2.7.1. Lithium acetate method for yeast transformation

A single colony from the AH109 yeast strain (or Y187) was used to inoculate 100ml YPD medium. The culture was grown at 30°C with vigorous shaking until the OD_{600} reached 0.5. The culture was pelleted by centrifugation at 4,000rpm after it reached the desired optical density. The pellet was washed twice with sterile water and resuspended in 100µl LiAc/TE (as suggested by the Clontech Matchmaker GAL4 Two-hybrid System manual). 100µl of this was aliquoted in eppendorf tubes. Transforming DNA was added at a concentration of 1µg per reaction tube, with 3µl carrier DNA and 700µl freshly prepared LiAc/TE/PEG solution. The tubes were incubated at 30°C without shaking for 30min. After the incubation, the cells were heat shocked at 42°C for 15min. 100µl from each tube was plated on selective plates and incubated at 30°C for 2-3 days until the colonies appeared.

2.7.2. Genomic library screen

Yeast strains were maintained at 30°C on YPD plates. The yeast two-hybrid analysis was performed as described in the Clontech Matchmaker System Manual and Yeast Protocol Handbook (Fig. 2.2 and Fig. 2.3). The bait plasmid encoding the first 147 amino acids of GAL4 DNA-BD was fused to *A.fulgidus* MCM, *A.fulgidus* ORC-1 or ORC-2 proteins. An *A.fulgidus* genomic DNA library, containing $3x10^6$ independent clones was prepared by N. Cook. The library was cloned into pGADT7 to create GAL4-AD fusions and transformed into the Y187 yeast strain. The yeast strain AH109 was transformed with pGBKT7-BD/MCM (ORC-1 or ORC-2) by the lithium acetate method described above. Transformants were selected on synthetic dropout plates lacking tryptophan (SD/W). Transformants were tested for transcriptional activation by plating colonies on SD/W, X-α-gal, SD/W, H, X-α-gal and SD/W, A, X-α-gal plates followed by incubation at 30°C for 5 days. Transformants were tested for leaky *HIS3* expression by plating colonies on a series

of SD/W, H plates containing 3-amino-1, 2, 2-triazole (3-AT) in the range of 0-25mM (0, 2.5, 5, 7.5, 10, 12.5, 15, 20 and 25mM). The plates were incubated at 30°C for 6 days. A cell toxicity test was also performed comparing the growth rates in liquid cultures (YPD) of the transformed bait and an "empty" pGBKT7--BD vector.

A culture of one large colony (2-3mm in diameter) of the AH109 strain carrying pGBKT7-BD/MCM (or ORC-1 or ORC-2) was grown overnight at 30°C in 50ml of SD/W. Cells were recovered by centrifugation at 1,000g for 5min and resuspended in 5ml of residual liquid by vortexing. These cells were mixed (mated) with 1ml of Y187 yeast strain carrying pGADT7-AD/*A.fulgidus* genomic DNA library and 45ml of 2xYPD containing 0.003% adenine and 10mg/l kanamycin (2xYPDA/Kan) and incubated at 30°C for 24 hours with shaking (30rpm). The cells were harvested by centrifugation (1,000g for 10min) and resuspended in 10ml of 0.5xYPDA/Kan. 100µl of a 1:10 000, 1:1000, 1:100 or 1:10 dilution of the mating mixture was plated out on SD/L, SD/W and SD/L, W plates to control mating efficiency. Mating efficiency was calculated as (#cfu/ml of diploids/#cfu/ml of limiting partner) x100=% diploid, where the limiting partner is the Y187 (library) strain. 50% of the mating mixture was plated on SD/W, L, H, A and 50% on SD/W, L, H plates. All plates contained X-α-gal at a concentration of 0.5µg per 20ml plate. Plates were incubated at 30°C for 8-15 days, until the colonies had grown bigger than 2 mm in diameter. The surviving colonies from SD/W, L, H plates were replica plated onto SD/W, L, H, A plates and incubated for 3 days at 30°C (Fig. 2.2).

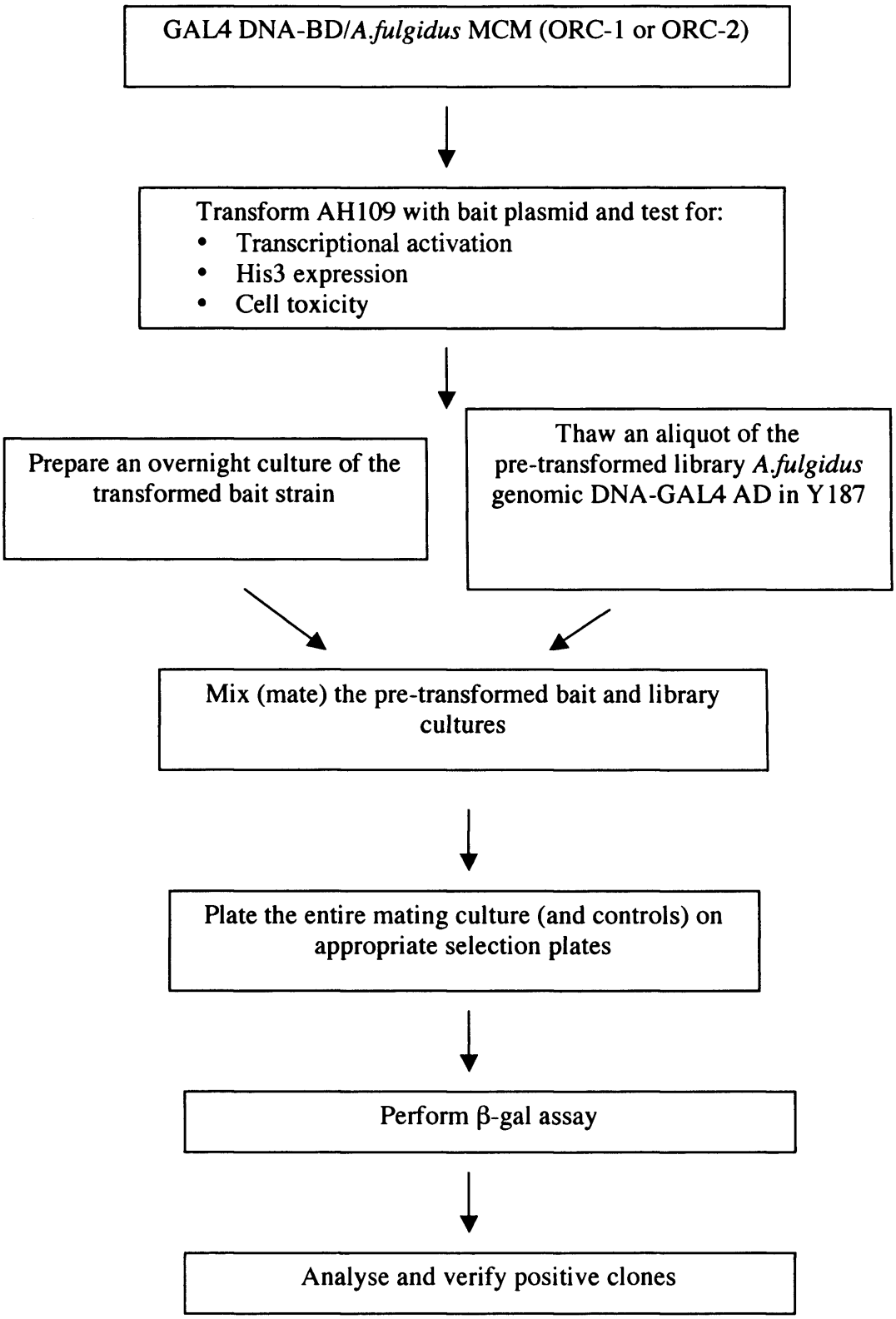


Fig. 2.2 Yeast two-hybrid library screening procedure.

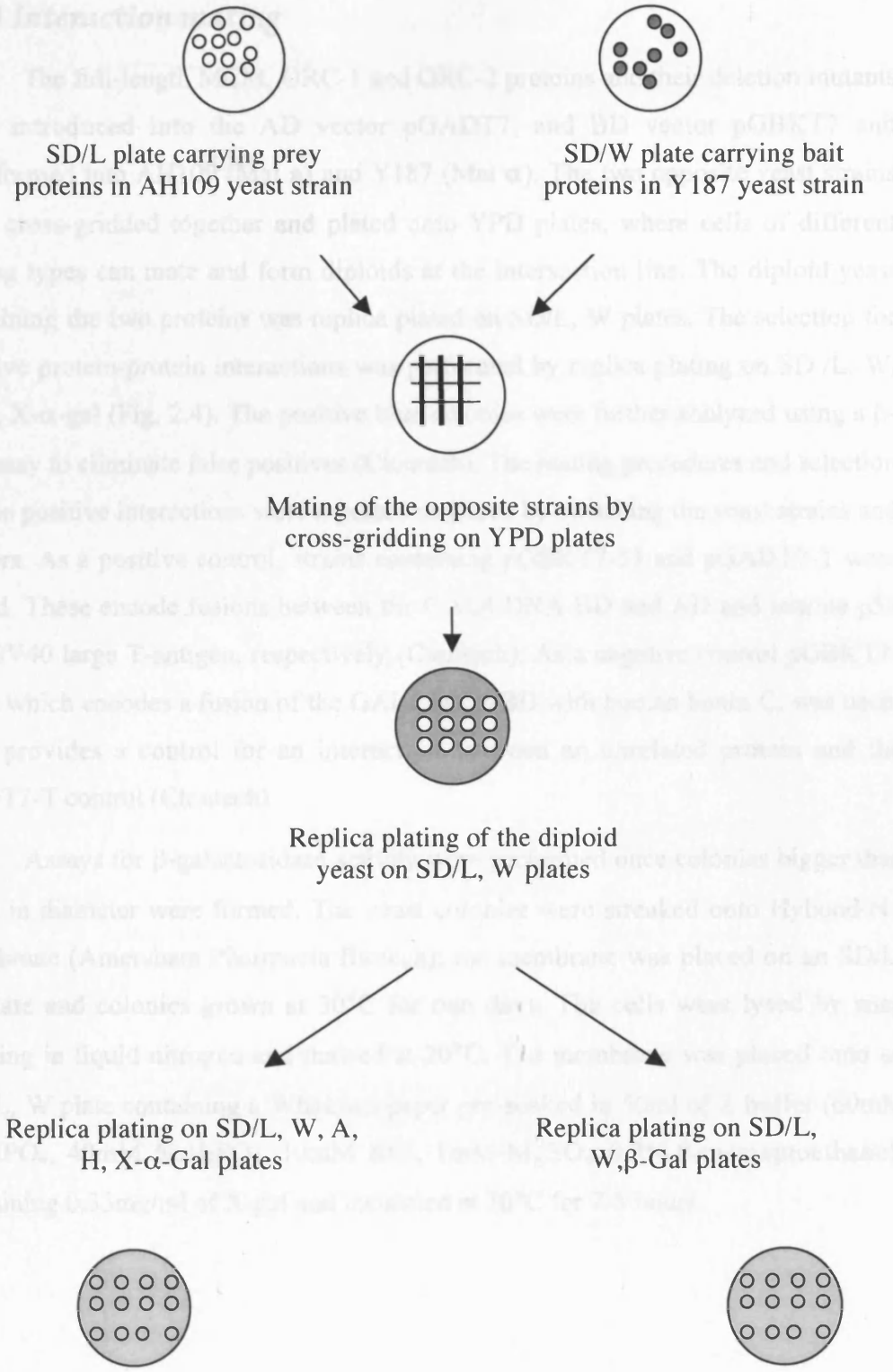


Fig. 2.3 Interaction mating by cross-gridding and β-galactosidase assay.

2.7.3 *Interaction mating*

The full-length MCM, ORC-1 and ORC-2 proteins and their deletion mutants were introduced into the AD vector pGADT7, and BD vector pGBKT7 and transformed into AH109 (Mat **a**) and Y187 (Mat α). The two opposite yeast strains were cross-gridded together and plated onto YPD plates, where cells of different mating types can mate and form diploids at the intersection line. The diploid yeast containing the two proteins was replica plated on SD/L, W plates. The selection for positive protein-protein interactions was performed by replica plating on SD /L, W, A, H, X-α-gal (Fig. 2.4). The positive blue colonies were further analyzed using a β-gal assay to eliminate false positives (Clontech). The mating procedures and selection for the positive interactions were repeated as above by switching the yeast strains and vectors. As a positive control, strains containing pGBKT7-53 and pGADT7-T were mated. These encode fusions between the GAL4 DNA-BD and AD and murine p53 and SV40 large T-antigen, respectively (Clontech). As a negative control pGBKT7-Lam, which encodes a fusion of the GAL4 DNA-BD with human lamin C, was used. This provides a control for an interaction between an unrelated protein and the GADT7-T control (Clontech)

Assays for β-galactosidase activity were performed once colonies bigger than 2mm in diameter were formed. The yeast colonies were streaked onto Hybond-N+ membrane (Amersham Pharmacia Biotech); the membrane was placed on an SD/L, W plate and colonies grown at 30°C for two days. The cells were lysed by snap freezing in liquid nitrogen and thawed at 20°C. The membrane was placed onto an SD/ L, W plate containing a Whatman paper pre-soaked in 50ml of Z buffer (60mM Na_2HPO_4, 40mM NaH_2PO_4, 10mM KCl, 1mM $MgSO_4$, 0.3% β-mercaptoethanol) containing 0.33mg/ml of X-gal and incubated at 30°C for 7-8 hours.

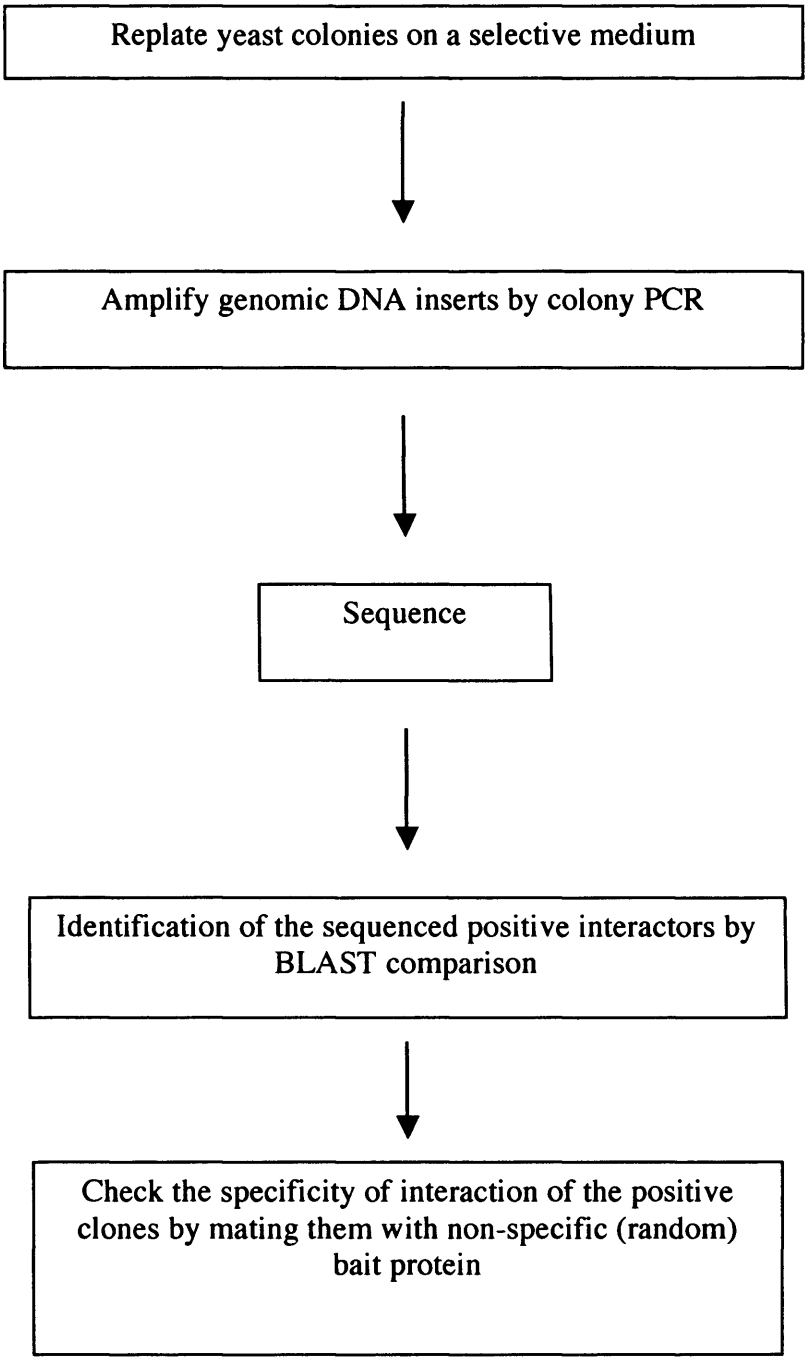


Fig. 2.4 Analysis and verification of positive clones.

2.7.4 *Identifying positive interactors by yeast colony PCR*

The positive interactors from the genomic library screen were retrieved by PCR amplification directly from the yeast colonies. One day before the PCR reactions the yeast strains were replica plated on YPD plates using a 96-well plate as a guide. To PCR the encoding sequences the following primers were used: T7 forward 5'TAATACGACTCACTATAGGGC3' and a 3'activation domain sequencing primer 5'AGATGGTGCACGATGCACAG3'. On the day of the experiment, 50µl PCR mixtures were prepared according to the manufacturer's instructions (Stratagene) in the presence of 100ng of the forward and reverse primers. 50µl aliquots were transferred into each well of the 96-well plate and kept on ice. Yeast cells from the positive colonies were picked using sterile toothpicks and resuspended in 10µl 20mM NaOH. These yeast suspensions were micro waved for 30 seconds and 3µl of each of them was added into each of the PCR mixtures. The PCRs were run according to the following programme: 94°C for 1min, followed by 35 cycles of 94°C-1min, 56°C-1.5min, 72°C-3min and one cycle of 72°C for 5min. 5µl of each PCR reaction were run on a 1% agarose gel. The PCR products were purified and sequenced.

2.7.5 *Preparation of yeast protein extracts by the TCA method*

Each deletion of MCM was cloned in to AD and BD vectors and transformed in AH109 and Y187 yeast strains. Overnight cultures of these transformations were prepared. Next day, 50ml aliquots of YPD medium were inoculated with the entire overnight cultures and incubated at 30°C with shaking until the OD_{600} reached 0.5. The OD_{600} units of each sample were calculated by multiplying the OD_{600} of each sample by the culture volume of the sample. (For example 0.5x50=30 total OD_{600} units). The cultures were then centrifuged at 1000g for 5min at 4°C. The pellets were resuspended in 50ml of ice-cold water and pelleted again at the same conditions. Each cell pellet was resuspended in 100µl of ice-cold TCA buffer (20mM Tris, pH 8.0, 50mM ammonium acetate; 2mM EDTA, 50µl/ml protease inhibitor solutions) per 7.5 OD_{600} units of cells. Each cell suspension was

transferred to a microcentrifuge tube, containing glass beads and ice cold 20%TCA (Sambrook and Russell, 2001). The cells were disrupted by vortexing for five times in 30 seconds sessions, the tubes were placed on ice for 30sec in between sessions. The supernatants above the settled glass beads were transferred to new 1.5ml tubes and placed on ice. Proteins were pelleted by centrifuging at 14,000rpm for 10min at 4°C. Each pellet was then resuspended in 10µl TCA-Laemmli loading buffer (SDS/glycerol stock solution, Tris/EDTA stock solution, β-mercaptoethanol, protease inhibitor) per OD_{600} unit of cells. The samples were boiled for 10min and centrifuged at 14,000rpm for 10min at room temperature. The supernatants were transferred into new tubes and loaded on a SDS-gel.

2.7.6 Western blot analysis

The gel-separated proteins were transferred to nitrocellulose membranes using standard Western blot technique. The membranes were treated for one hour at room temperature with phosphate-buffered saline (PBS), pH 7.5 containing 5% (w/v) milk powder. The membranes were then incubated with primary antibodies for one hour, followed by incubation with secondary antibodies for 45min. Antibodies made against the c-Myc (BD) and HA (AD) epitope tags were used. Protein-antibody complexes were visualized by an enhanced chemoluminescence Western blotting detection system (Amersham).

2.7.7 Analysis with in vitro translated proteins

Co-immunoprecipitation of full-length MCM, ORC-1 and ORC-2 proteins and their deletion mutants with each other was performed using the Matchmaker Co-IP Kit according to the manufacturer's instructions (Clontech). The protein couples were co-transcribed and co-translated *in vitro* using TnT® T7 Coupled Reticulocyte Lysate System in the presence of [^{35}S] methionine, following the manufacturer's instructions (Promega). Each clone was transcribed and translated separately. 5µl of each of the 50µl translation reactions were

analysed by 10% SDS-PAGE gels. 22.5µl of each of the co-translations and 10µl of each of the single translations were incubated for 1h at 20°C with 1µg hemagglutinin (HA) polyclonal antibody. 3µl of protein A beads were equilibrated with PBS and added to each reaction. The beads were then incubated for 1h at 20°C with agitation. The beads were washed seven times with 500µl of wash buffer (25mM Tris, pH 7.5, 300mM NaCl, 0.2% NP40, 0.1% SDS, 10% glycerol) and protein complexes eluted with 20µl of SDS sample buffer. 10µl of each elution was analysed by 10% SDS-PAGE and labeled polypeptides were detected by phosphorimaging.

CHAPTER 3
PROTEIN-PROTEIN INTERACTIONS WITHIN THE PRE-RC OF *A.fulgidus*

3.1. *Af*MCM hexamerisation domain

3.1.1 Introduction

Minichromosome maintenance (MCM) helicases are ring-shaped complexes that play an essential role in archaeal and eukaryal DNA replication. Whereas eukaryotes all have six different, albeit homologous MCM proteins, their number in archaea is variable, with several species having only a single paralog.

The structure of the archaeal MCM complex is unclear. The MCM homologues of *S.solfataricus* (Carpentieri *et al.*, 2002) and *A.fulgidus* (Grainge *et al.*, 2003) form hexamers in solution, but *M.thermoautotrophocum* MCM appears to form dodecamers (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; and Poplawski *et al.*, 2001). Electron microscopy (EM) analyses of *Mth*MCM revealed ring-shaped hexamer (Chong *et al.*, 2000) or heptameric assemblies (Yu *et al.*, 2002). More recent three-dimensional reconstructions of different fragments of the *Mth*MCM revealed formation of single hexamers only (Chen *et al.*, 2005). Thus, it appears that only the full-length protein may be capable of forming heptamers. Although most processive helicases appear to be hexameric, it is possible that hexamers and heptamers may form under different conditions, and the ability to form rings with different symmetries may be a general property of helicases and other ring-forming proteins (Yu *et al.*, 2002). It is possible that upon assembly around the DNA, the heptameric ring of the *Mth*MCM may loose one subunit, encircle DNA, and then close to form a ring with six subunits, becoming the active enzyme (Yu *et al.*, 2002; Kelman and Kelman, 2003).

The crystallographic structure of *MthMCM* N-terminal portion (residues 2-286) revealed a dodecameric architecture, with two hexameric rings juxtaposed in a head-to-head configuration (Fletcher *et al.*, 2003). According to this crystal structure, each monomer is folded into three distinct domains A, B and C (Fletcher *et al.*, 2003). Domain A is needed for dsDNA translocation and is believed to have a regulatory function, while domain C is necessary and sufficient for MCM multimerisation and helicase activity (Kasiviswanathan *et al.*, 2004). Domain B is the major contact with ssDNA and contains the zinc-binding motif. The zinc-binding motif is required for the head-to-head assembly of the two hexameric rings (Fletcher *et al.*, 2003). The three-dimensional crystal structure suggests that the zinc-binding site plays a role in double

hexamer formation. However, biochemical analysis of a zinc-binding mutant showed that the mutant protein is impaired in single-stranded DNA binding but not double-hexamer formation (Poplawski *et al.*, 2001). In addition, studies with the *A.fulgidus* MCM suggested that domain B, which contains the zinc-binding site, is not needed for multimerisation (Grainge *et al.*, 2003) but probably for single-stranded DNA binding. Further analysis on the N-terminal part of *Mth*MCM revealed that domain C is necessary and sufficient for MCM multimerisation and is essential for helicase activity (Kasiviswanathan *et al.*, 2004).

3.1.2 Aim of the project

This work was focused on the characterization of the oligomerisation domain of MCM homologue from the euryarchaeon *Archaeoglobus fulgidus* (*Af*MCM). The GAL4 based yeast two-hybrid system and *in vitro* analysis were used for detecting the minimal hexamerisation sequence of *Af*MCM. Knowledge of the minimal hexamerisation region may help in identification of a fragment of the MCM protein suitable for structural studies.

3.1.3 Two-hybrid approach for delimiting the MCM-MCM interacting domain (s)

Secondary structure prediction of *Af*MCM protein was performed using the protein prediction server http://www.embl-heidelberg.de/predictprotein. Subsequently, six sequential 100 amino acid deletions from the C-terminus (residues 1-599, 1-499, 1-399, 1-299, 1-199 and 1-99) and six sequential deletions of 100 amino acids from the N-terminus of the *Af*MCM protein (residues 101-699, 201-699, 301-699, 401-699, 501-699) were amplified by PCR. The amplified products were subcloned into the EcoRI-BamHI sites of the activation domain and binding domain fusion vectors provided by the Clontech Two-Hybrid Kit. The DNA binding domain vector pGBKT7 contains the *TRP1* gene for yeast selection and has kanamycin resistance for bacterial selection. The DNA activation domain vector pGADT7 contains *LEU2* gene for yeast selection and ampicillin resistance for bacterial selection. The activation and binding domain vectors

carrying the truncated MCM proteins were transformed into AH109 (**Mat a**) and Y187 (Mat α) yeast strains.

At the beginning, all the bait proteins were examined for protein expression by Western blot analysis. Next, they were tested for toxicity effects by comparing their growth rate with that of cells transformed with an empty pGBKT7–BD vector. Inspection for transcriptional activation was performed by mating (mixing) them with an empty pGADT7-AD vector. Leaky *HIS3* expression was tested by plating the transformed baits on selective dropout media lacking histidine (SD/H). All the experiments confirmed that the bait proteins could be used in the two-hybrid system efficiently since they expressed well, did not activate transcription or cause leaky HIS3 expression, and were not toxic.

Protein-protein interactions between the MCM deletion mutants were assessed by mating the two opposite yeast strains containing the activation and binding domain vectors. The two opposite yeast strains were cross-gridded together and plated onto YPD plates, where cells of different mating types can mate and form diploids at the intersection line. The diploid yeast, containing the protein couples, was replica-plated on selective dropout plates lacking leucine and tryptophan (SD/H, W). To select for positive interactions, the diploid yeast colonies were replica-plated on plates lacking adenine, histidine, leucine and tryptophan (SD/A, H, L, W). The positive colonies were further analyzed by the β-galactosidase assay to eliminate false positives. The mating procedures and selections for the positive interactions were repeated as above by switching the yeast strains and vectors. As a positive control, strains containing pGBKT7-53 and pGADT7-T were mated. These encode fusions between the GAL4 DNA-BD and AD and murine p53 and SV40 large T-antigen, respectively (Clontech). As a negative control pGBKT7-Lam, which encodes a fusion of the GAL4 DNA-BD with human lamin was used. This provides a control for an interaction between unrelated protein and the GADT7-T control (Clontech).

At the beginning, full length *Af*MCM was chosen to act as bait. Interactions between the full-length MCM and the MCM truncation library are shown on Fig. 3.1A. MCM interacts positively with itself and the following deletion mutants: residues 1-599,

1-499, 1-399, 101-699 and 201-699. These first results suggested that the oligomerisation domain of *Af*MCM is between residues 201-399.

To confirm the interaction domain, the region of MCM between residue numbers 201-399 was cloned into the BD fusion vector and used as bait in further two-hybrid analyses. Positive interactions were observed with the full length *Af*MCM and residues 201-699, 1-399 and with itself (Fig. 3.1B). All the observed interactions between MVM deletion mutants are shown on Fig. 3.2. As the figure suggests, the region responsible for MCM-MCM interactions might be between residues 301-399 (Fig. 3.2).

The proposed oligomerisation domain for *A.fulgidus* MCM differs from that observed in *M.thermoautotrophicum* MCM (Kasiviswanathan *et al.*, 2004) (Fig. 3.3). MCM is a member of the AAA+ family of ATPases (Neuwald *et al.*, 1999) and as such it contains Walker A and Walker B motifs, responsible for nucleotide binding and hydrolysis. In several cases, these motifs have been found at subunit interface (Mourez *et al.*, 1997). The suggested *A.fulgidus* MCM oligomerisation domain includes the full Walker A, but not Walker B motif.

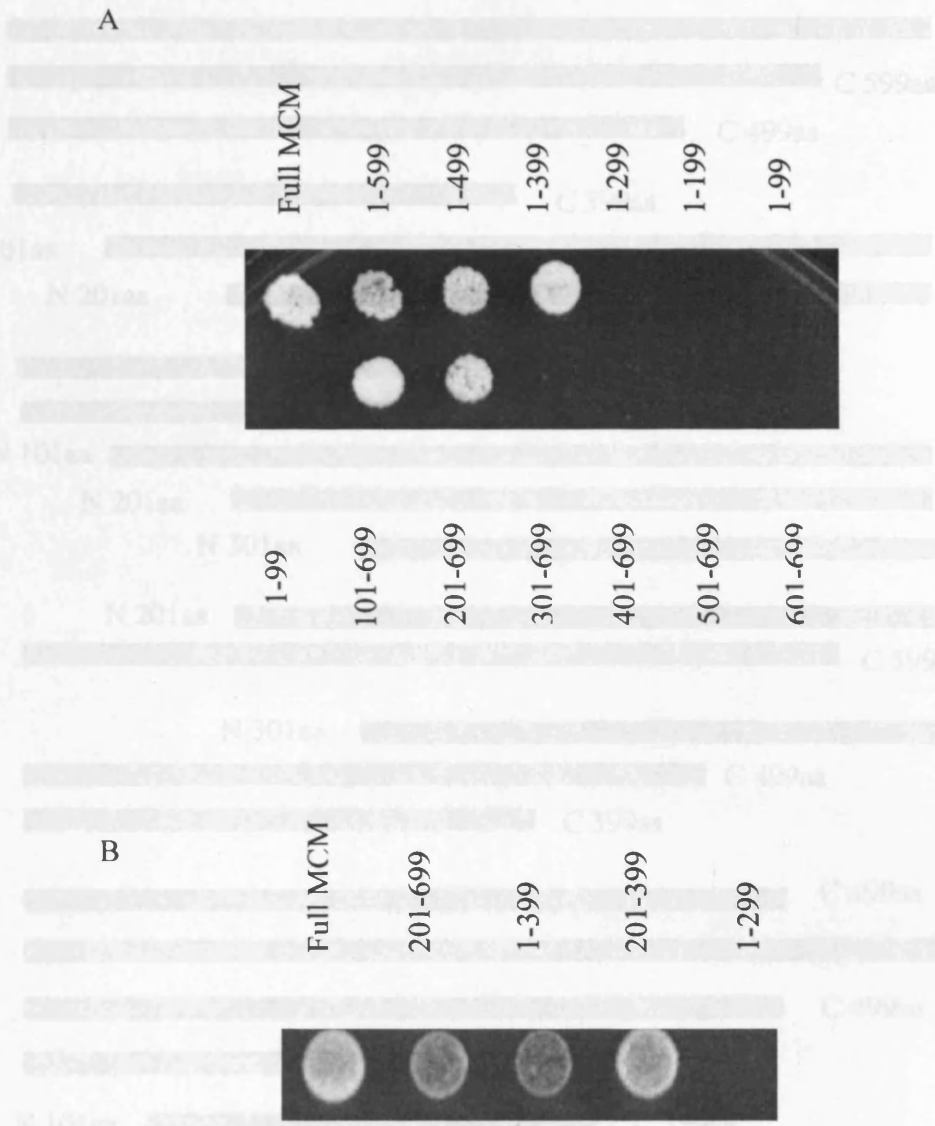


Fig. 3.1 Interaction mating experiments. (A) Interactions of the full-length *Af*MCM protein with the truncated AfMCM library. The oligomerisation domain of *Af*MCM appears to be between residues 201-399, (B) 201-399 *Af*MCM used as bait in the mating experiments. The observed positive self-interaction confirmed the oligomerisation region. Mating experiments were performed on SD/A,H,L,W plates.

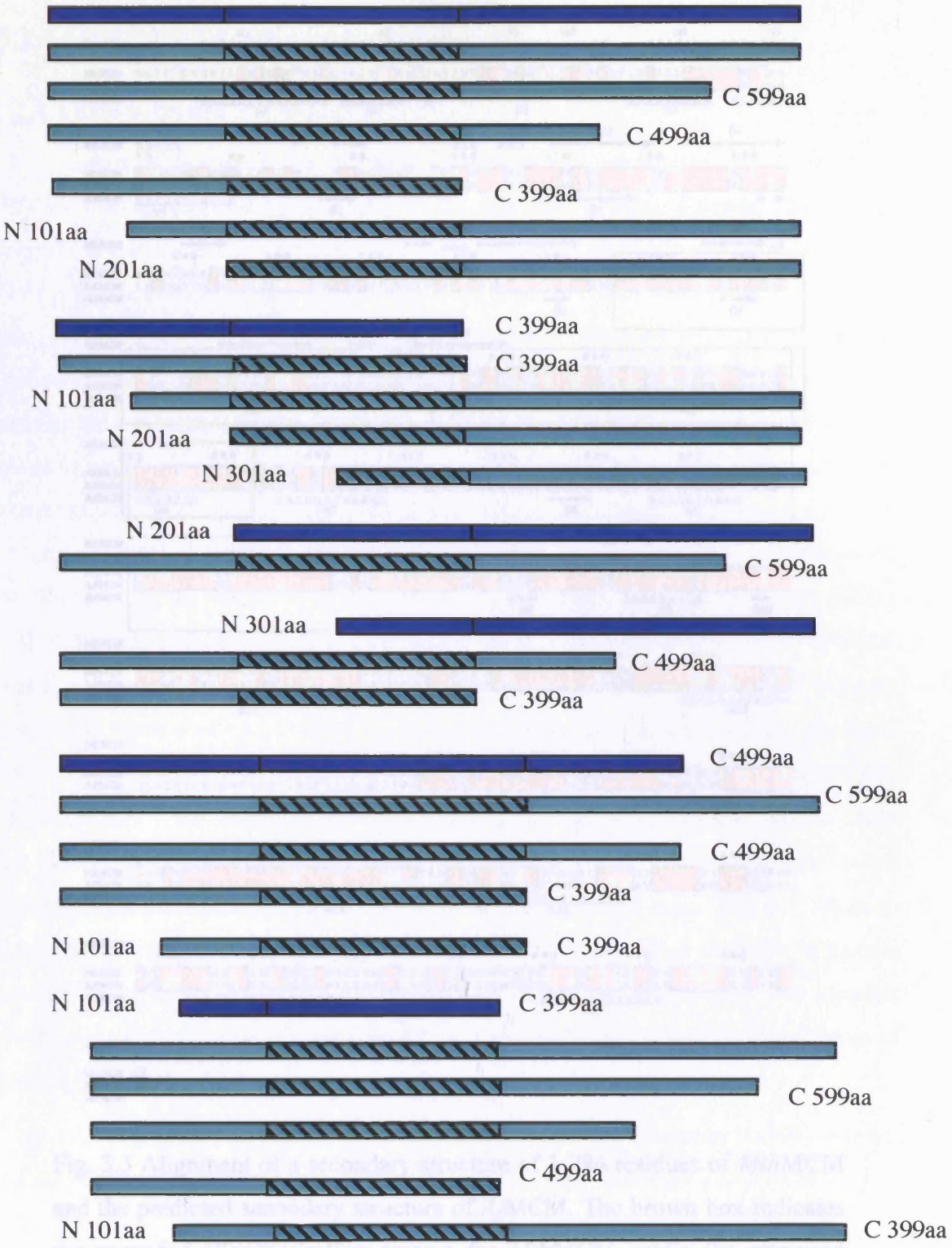


Fig. 3.2 Summary of the observed interactions within the *Af*MCM truncated library. The dark blue bars represent the protein truncations used as baits, the green bars the proteins used as preys. The textured bar indicates the region that appear to be responsible for the MCM-MCM interactions.

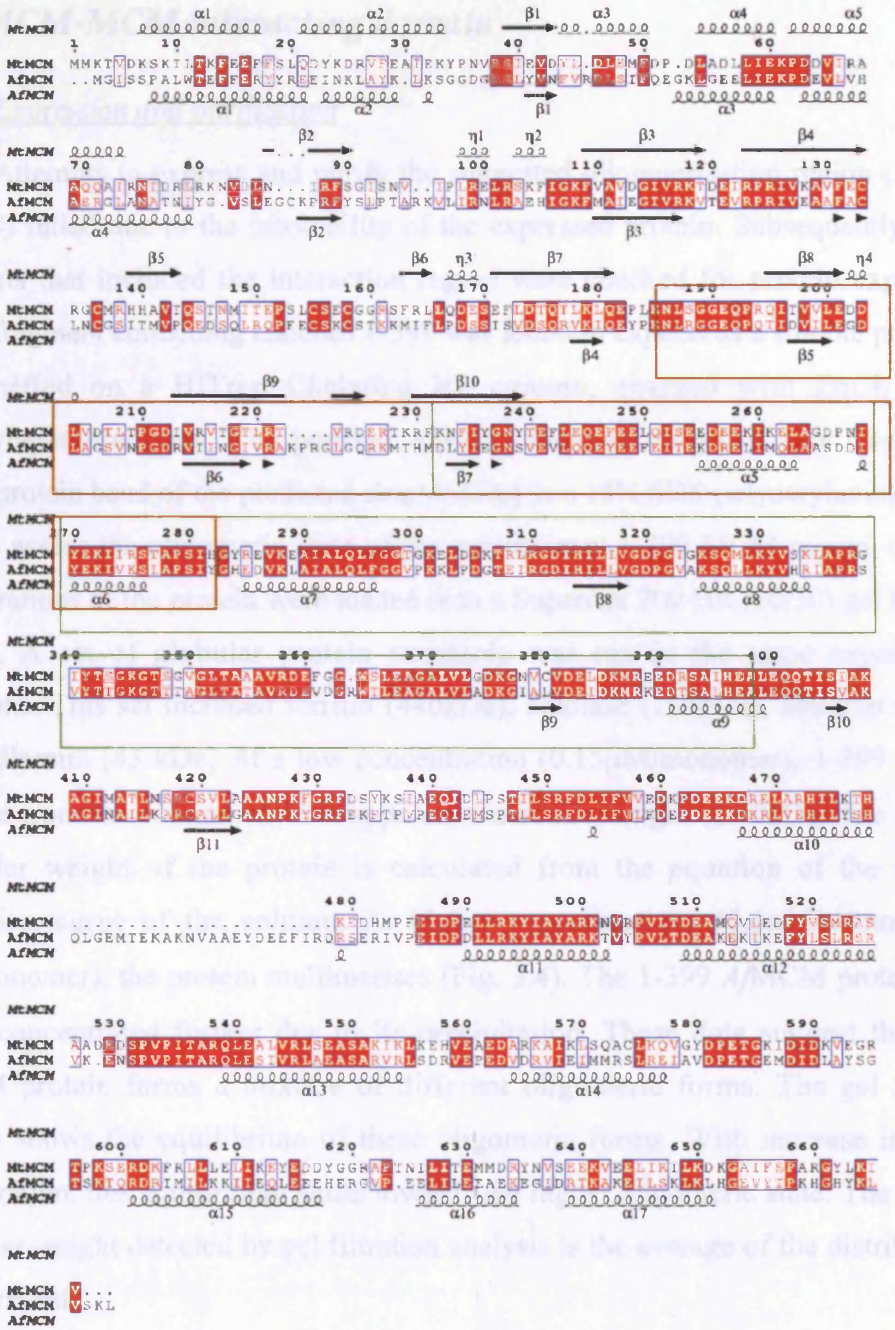


Fig. 3.3 Alignment of a secondary structure of 1-286 residues of *Mth*MCM and the predicted secondary structure of *Af*MCM. The brown box indicates the proposed oligomerisation domain for *Mth*MCM, while the proposed oligomerisation domain for *Af*MCM is shown with the green box.

3.1.4 MCM-MCM interacting domain

3.1.4.1 Expression and purification

Attempts to express and purify the suggested oligomerisation region (residues 301-399) failed due to the insolubility of the expressed protein. Subsequently, larger fragments that included the interaction region were checked for protein expression. Only a fragment containing residues 1-399 was found to express as a soluble protein. It was purified on a HiTrap Chelating HP column, charged with $ZnCl_2$, and a hydroxylapatite column. The purified *Af*MCM fragment migrated as a Coomassie-stained protein band of the predicted size (45kDa) in a 15% SDS-polyacrylamide gel. In order to assess the oligomeric state of the recombinant 1-399 MCM region, different concentrations of the protein were loaded onto a Superdex 200 HR (10/30) gel filtration column. A set of globular protein standards was run in the same experimental conditions. This set included ferritin (440kDa), aldolase (158kDa), albumin (67kDa) and ovalbumin (43 kDa) At a low concentration (0.15μM/monomer), 1-399 *Af*MCM migrated from the column with an apparent molecular weight of 49kDa. The apparent molecular weight of the protein is calculated from the equation of the standard calibration curve of the column. At higher concentrations (0.6μM/monomer and 3μM/monomer), the protein multimerises (Fig. 3.4). The 1-399 *Af*MCM protein could not be concentrated further due to its precipitation. These data suggest that 1-399 *Af*MCM protein forms a mixture of different oligomeric forms. The gel filtration analysis shows the equilibrium of these oligomeric forms. With increase in protein concentration, this equilibrium shifts towards the higher oligomeric state. The apparent molecular weight detected by gel filtration analysis is the average of the distribution of forms present.

The suspected hexamerisation domain of the protein (fragment 1-399) was not suitable for structural studies since it does not form a stable homogenous oligomer.

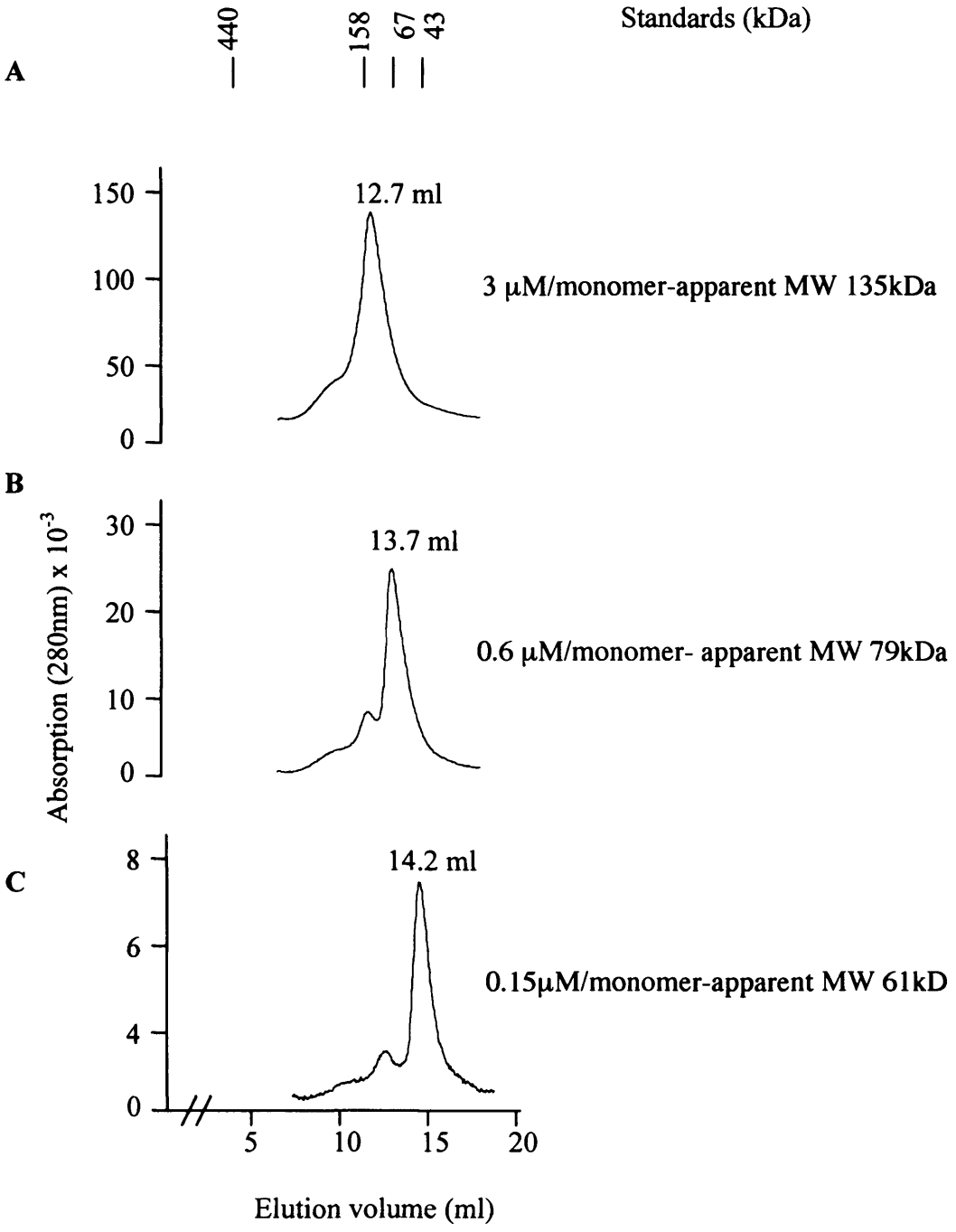


Fig. 3.4 Gel filtration of different concentrations of 1-399 *Af*MCM protein. (A) 3μM/monomer, (B) 0.6μM/monomer and (C) 0.15μM/monomer.

3.1.4.2 DNA binding characteristics of 1-399 AfMCM protein

Studies performed with *Mth*MCM demonstrated that the protein binds to DNA in the presence of magnesium (Kelman *et al.*, 1999; Chong *et al.*, 2000). A bubble substrate (Fig. 3.5) was used to assess the DNA binding ability of the 1-399 *Af*MCM protein. DNA binding ability of the full length *Af*MCM was used as a control. 1-399 *Af*MCM protein binds DNA as a single band as tight as the full-length protein.

In addition, EM studies (carried out by Ian Grainge) of 1-399 *Af*MCM performed in the absence of DNA, showed a mixture of different oligomeric forms. However in the presence of double-stranded (but not single-stranded) DNA, the protein formed hexameric rings with the DNA appearing to pass through the central channel of the ring (Fig. 3.6).

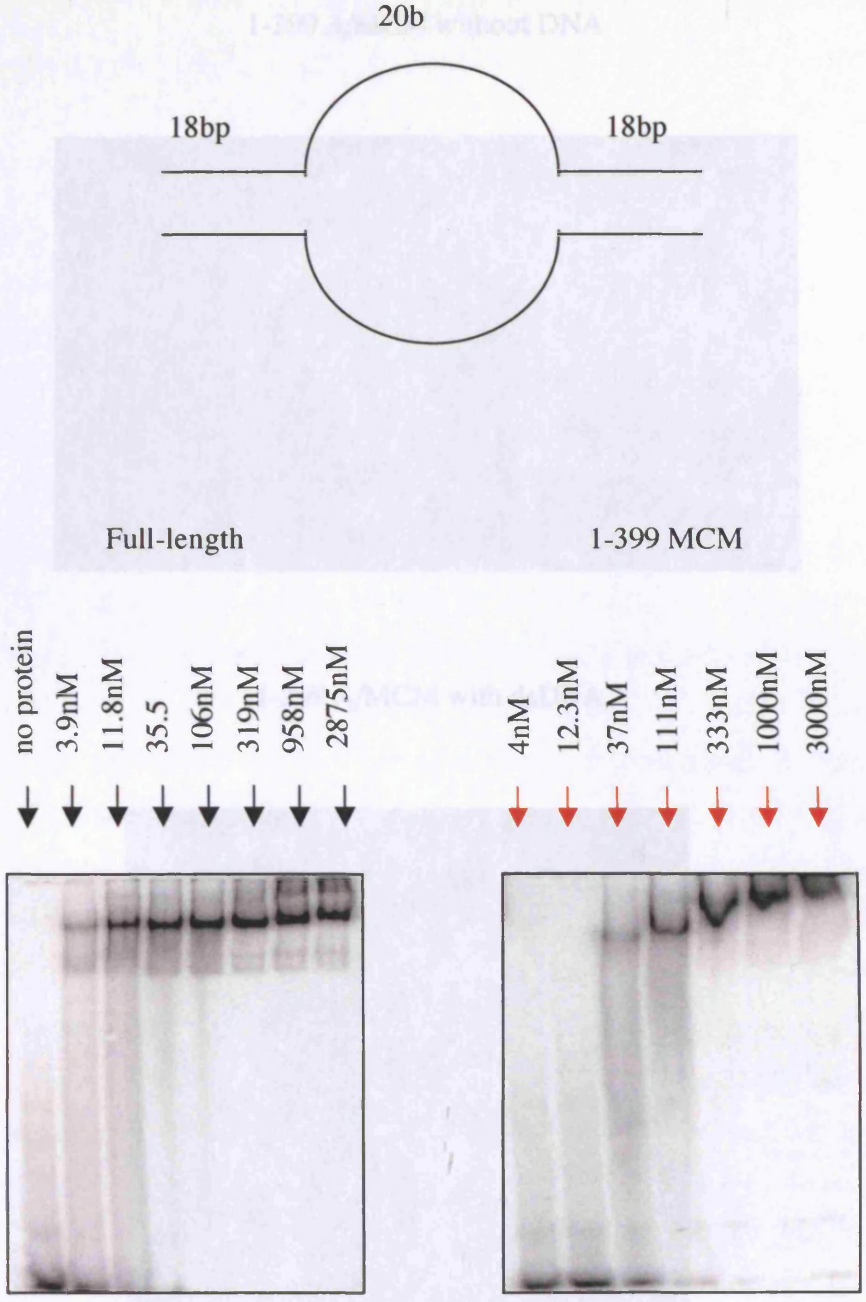


Fig. 3.5 DNA-binding assay for 1-399 *Af*MCM with a bubble substrate. Full-length *Af*MCM was used as a control. 1-399 *Af*MCM protein binds DNA as tight as the full-length protein.

1-399 *Af*MCM without DNA

1-399 *Af*MCM with dsDNA

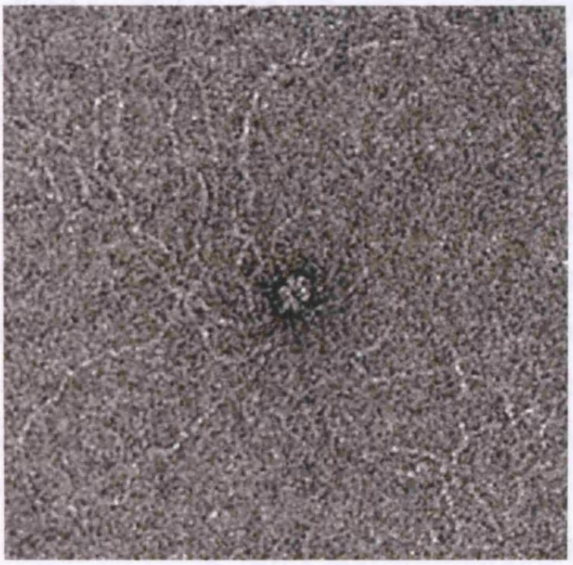


Fig. 3.6 EM studies of 1-399*Af*MCM, performed in the presence and absence of double-stranded DNA (made by Ian Grainge). While 1-399*Af*MCM aggregates in the absence of DNA, the addition of double-stranded (but not single-stranded) DNA stabilises the hexamer formation.

3.2 Protein-protein interactions within the archaeal pre-replication complex

3.2.1 Introduction

The process of DNA replication initiation in archaea is currently unknown. *In silico* analysis suggests that archaeal DNA replication proteins are more similar to those in Eukarya than those found in bacteria. However, the archaeal replication complexes contain fewer proteins than the eukaryotic homologues (Kelman *et al.*, 2003a and Grabowski and Kelman, 2003). Regarding the ORC proteins and CDC6, the situation is again simplified in most archaeal species to the extent that several species encode a single gene with homology both to ORC1 and CDC6 proteins (Grabowski and Kelman, 2003). In many other species there are two genes, one with greater homology to the yeast *Cdc6* gene and the other more like *Orc1*. To date, it has not yet been determined whether the archaeal CDC6/ORC homologues function as ORC, CDC6 or both. Sequence analysis suggests the existence of at least two subfamilies of these proteins. The basis for the division of the subfamilies is a conserved motif that may be a recognition helix (Singleton *et al.*, 2004). The three dimensional structure of the CDC6/ORC homologues from the archaeons *P.aerophylium* (Liu *et al.*, 2000) and *A.pernix* (Singleton *et al.*, 2004) revealed the expected domains found in other members of the AAA+ superfamily of ATPases (Neuwald *et al.*, 1999 and Ogura *et al.*, 2001). In addition to the ATPase domains (domains I and II), the proteins contain a C-terminal winged-helix (WH) domain (domain III). Amino acid substitutions and deletions within the WH domain demonstrated that the domain plays an important role in DNA binding (Grabowski *et al.*, 2001; De Felice *et al.*, 2004; Capaldi and Berger., 2004; Singleton *et al.*, 2004). Homologues of other proteins involved in the pre-RC formation in eukaryotes (CDT1, CDC45, MCM10) have not been identified in Archaea.

The interactions among the archaeal initiation proteins are currently unknown. Two-hybrid and Far Western blot analysis has demonstrated the interactions between MCM and the two ORC/CDC6 homologues identified in the *M.thermoautotrophicus* genome (*Mth*CDC6-1 and *Mth*CDC6-2) (Shin *et al.*, 2003a). Although the ORC/CDC6 homologues are very similar in primary amino acid sequence (Giraldo, 2003) and were

suggested to have similar structure and domain organisation (Liu *et al.*, 2003; Singleton *et al.*, 2004), they appear to employ different regions for MCM binding. While *Mth*CDC6-1 binds MCM predominately via the WH domain (Shin *et al.*, 2003), this domain of *Mth*CDC6-2 does not interact with the MCM helicase. Only the full-length *Mth*CDC6-2 protein has been shown to bind MCM. This is similar to the observations made with a CDC6 homologue from *S. solfataricus* in which an indirect assay suggested that the WH domain of one of the three CDC6 homologues is not required for MCM binding (De Felice *et al.*, 2003). The major contact between *Mth*MCM and *Mth*CDC6 homologues is *via* the N-terminal portion of MCM (Kasiviswanathan *et al.*, 2004). In contrast, the CDC6 homologues from *P.aerophilum* (*Pa*CDC6) did not bind to MCM (Shin *et al.*, 2003).

3.2.2 Aim of the project

GAL4 based yeast two-hybrid and co-immunoprecipitation analyses were used to test the interactions between the known proteins, involved in the pre-replication complex of *A.fulgidus*. Two-hybrid system was used for the screening of an archaeal genomic library as well.

A.fulgidus contains 2,436 predicted coding sequences and encodes one MCM (AF0517) and two ORC/CDC6 proteins (AF0244 and AF0695). Based on the nomenclature proposed by Singleton *et al.*, 2004, henceforth the AF0244 will be referred as ORC-1 and AF0695 as ORC-2. To date, only limited studies have been reported on the protein-protein interactions among the archaeal pre-RC proteins. Moreover, no homologues of the eukaryotic MCM10, CDC45 and CDT1 have been found based on sequence similarities. However, it seems likely that orthologues of some of these proteins might be present in archaea.

3.2.3 *Two-hybrid investigation of the potential interactions*

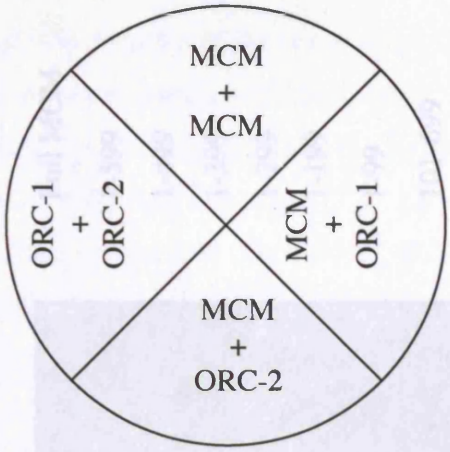
The full-length proteins were amplified by PCR and subcloned into EcoRI-BamHI sites of the activation and binding domain fusion vectors, provided by the Clontech Two-Hybrid Kit. The plasmids containing the bait protein sequences were transformed into AH109 (**Mat a**) yeast cells and examined for protein expression, toxicity, transcriptional activation and leaky *HIS3* expression as described above. The bait proteins were not toxic, expressed well and did not activate transcription or cause leaky HIS3 expression.

Subsequently, *A. fulgidus* MCM, ORC-1 and ORC-2 proteins were fused to the activation domain and transformed into the opposite mating type yeast strain-Y187 (MATα).

Protein-protein interactions were assessed by interaction mating as described above. Positive interactions were not detected among the full-length MCM, ORC-1 and ORC-2 proteins (Fig. 3.7). MCM self-interaction was used as a positive control.

The three dimensional structure of the CDC6/ORC homologues from other archaeal species (Liu *et al.*, 2000 and Singleton *et al.*, 2004) showed that these proteins have three domains. Domain I and II contain the ATPase motifs and C-terminal WH domain III binds DNA. Using these structures as a guide, *Af*ORC-1 and *Af*ORC-2 domains were identified. For ORC-1, domain I comprises residues 1-209, domain II-residues 210-283 and domain III consists of residues 284-409. ORC-2 domain I included residues 1-208, domain II-residues 209-286 and domain III-residues 287-376. These different domains were also cloned into the activation and binding domain vectors used previously, and transformed into AH109 and Y187 yeast strains.

Mating experiments were performed using the MCM truncation library (as described in the previous screen) with different domains of ORC-1 and ORC-2 (Fig. 3.8). Positive interactions were not observed in this experiment.



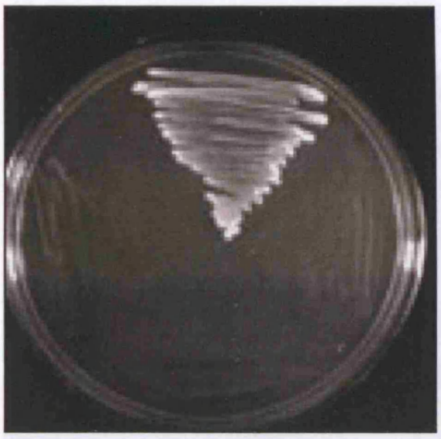


Fig. 3.7 Interaction mating between full-length *Af*MCM, *Af*ORC-1 and *Af*ORC-2 proteins.

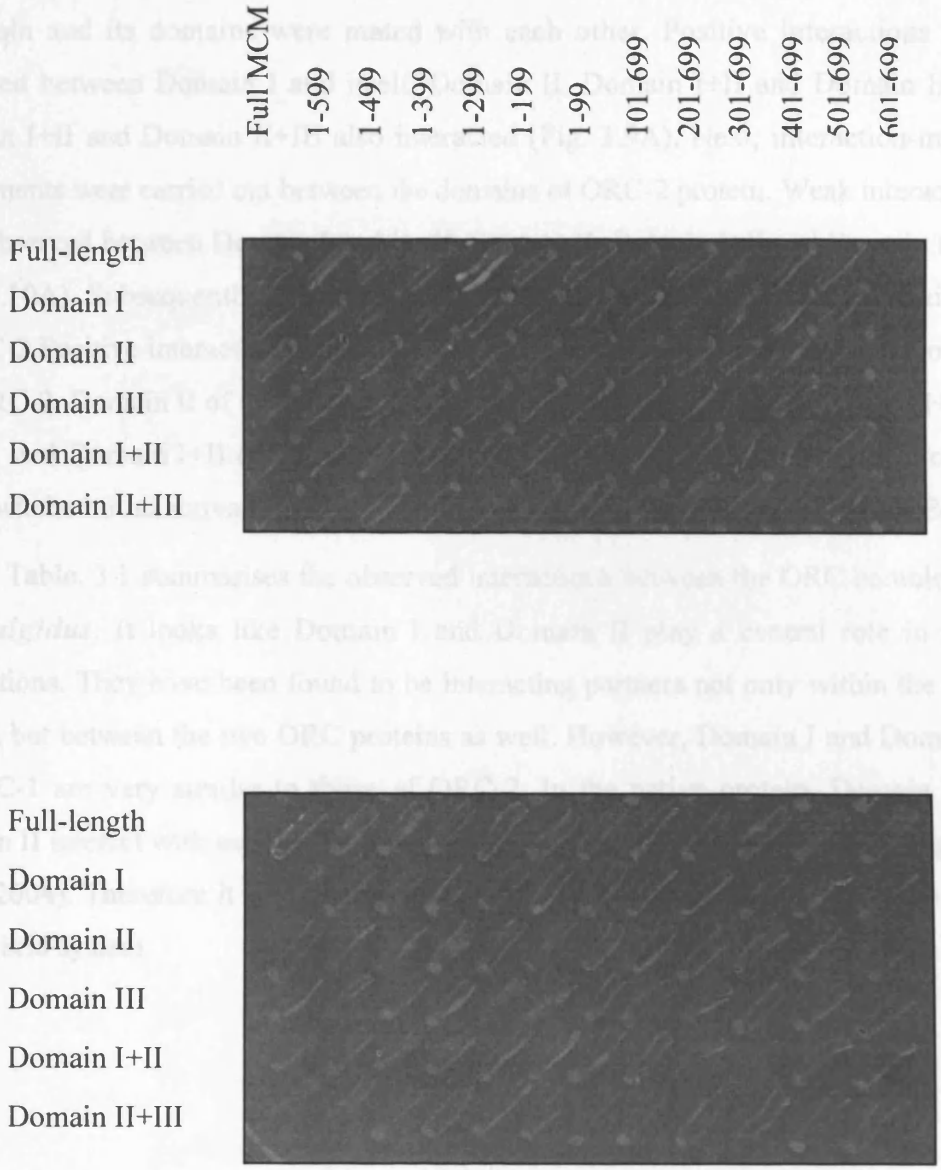


Fig. 3.8 Interaction mating between *Af*MCM truncation library and (A) *Af*ORC-1 and (B) *Af*ORC-2 domains.

Protein-protein interactions were assessed between different domains of the same protein and between the domains of the two distinct ORC proteins. First, *Af*ORC-1 protein and its domains were mated with each other. Positive interactions were observed between Domain I and itself, Domain II, Domain I+II and Domain II+III. Domain I+II and Domain II+III also interacted (Fig. 3.9A). Next, interaction-mating experiments were carried out between the domains of ORC-2 protein. Weak interactions were observed between Domain I and itself, Domain II, Domain I+II and Domain II+III (Fig. 3.10A). Subsequently, the domains of *Af*ORC-1 were mated with the domains of *Af*ORC-2 Positive interactions were detected between Domain I of ORC-1 and Domain I of ORC-2, Domain II of ORC-1 and Domain II of ORC-2 and between Domain I+II of ORC-1 and Domain I+II of ORC-2. These interactions were confirmed when proteins were switched from activation to binding domain vectors (Fig. 3.9B and Fig. 3.10B).

Table. 3.1 summarises the observed interactions between the ORC homologous of *A.fulgidus*. It looks like Domain I and Domain II play a central role in these interactions. They have been found to be interacting partners not only within the same protein but between the two ORC proteins as well. However, Domain I and Domain II of ORC-1 are very similar to those of ORC-2. In the native protein, Domain I and Domain II interact with each other as they form together the ATPase domain (Singleton *et al.*, 2004). Therefore it is not surprising that those two domains interact in the yeast two-hybrid system.

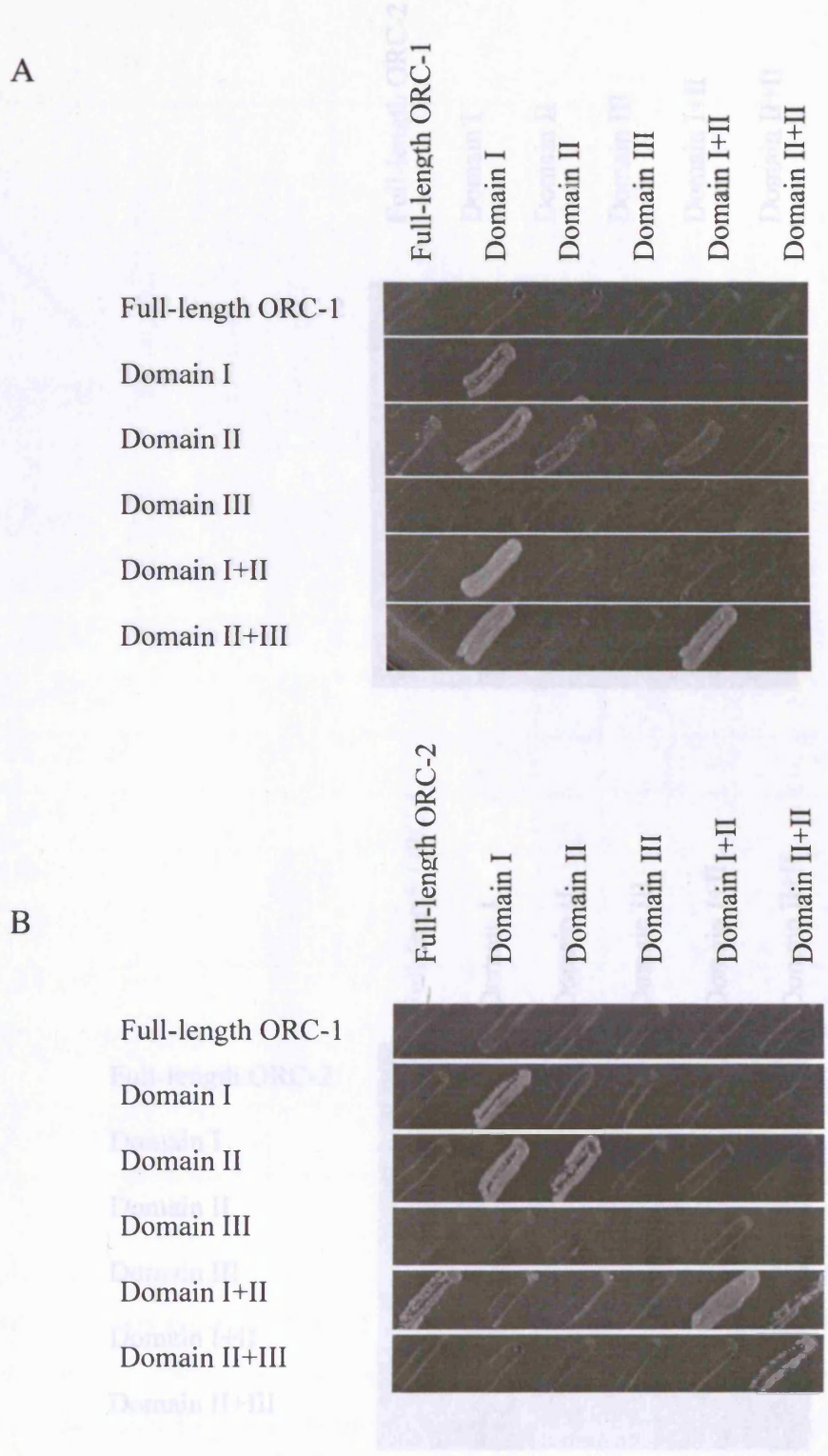


Fig. 3.9 (A) Interaction mating between the *Af*ORC-1 truncated mutants; (B) Interactions between *Af*ORC-1 and *Af*ORC-2 truncation libraries, *Af*ORC-1 mutants were used as bait proteins.

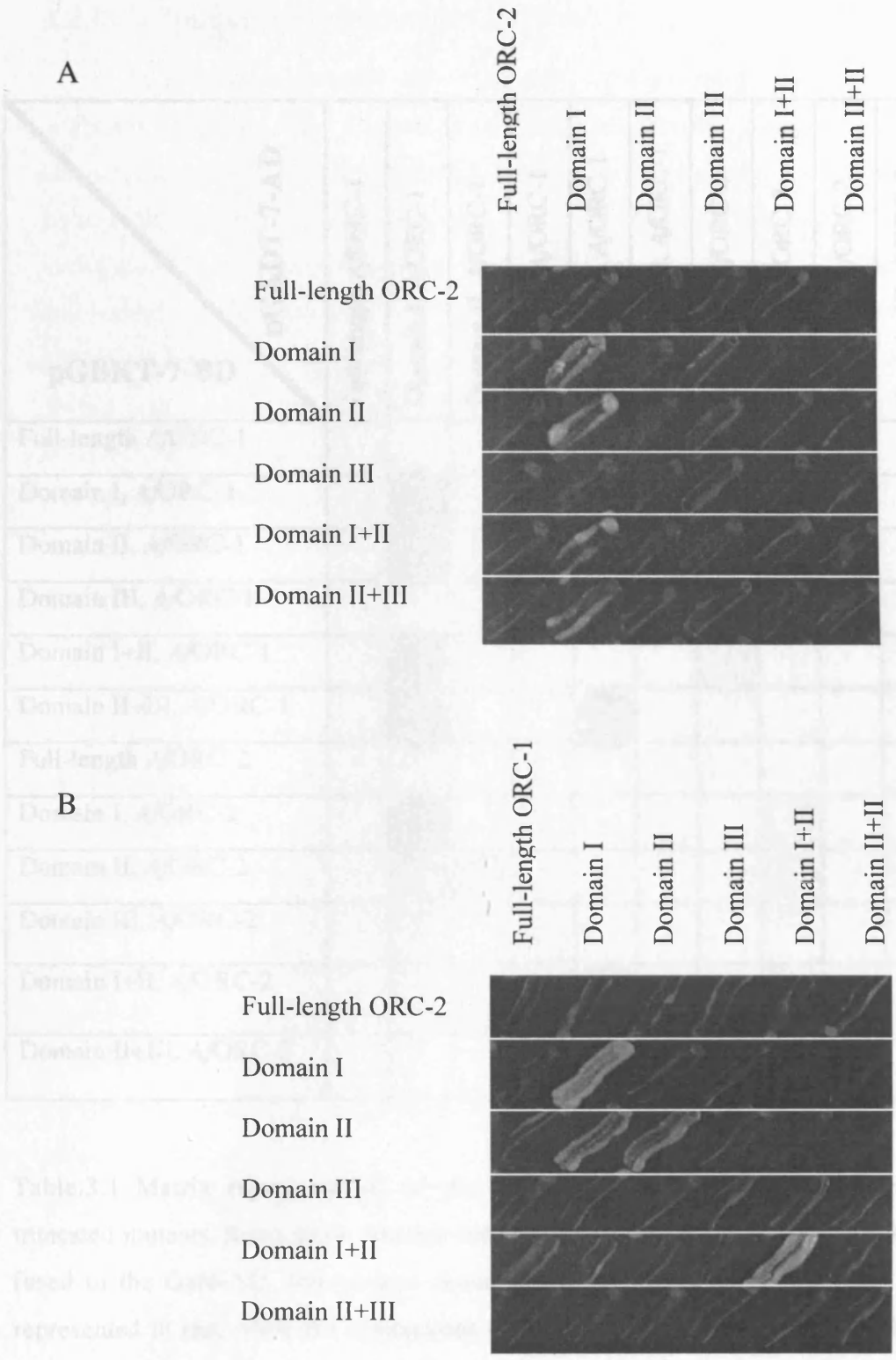


Fig. 3.10 (A) Interaction mating between the *Af*ORC-2 truncated mutants, (B) Interactions between *Af*ORC-2 and *Af*ORC-1 truncation libraries, *Af*ORC-2 mutants were used as bait proteins.

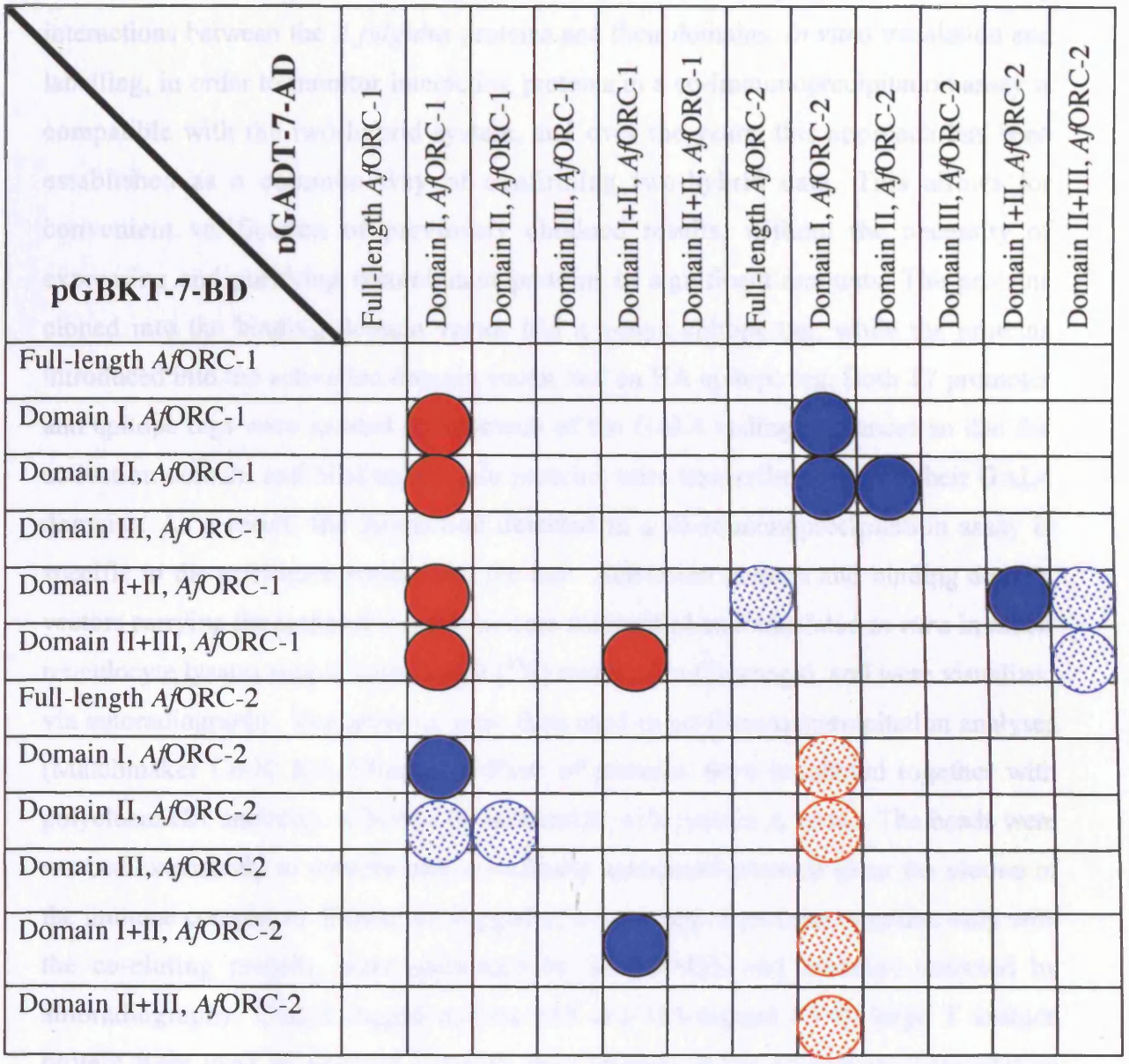


Table.3.1 Matrix representation of the interactions between *Af*ORC-1 and *Af*ORC-2 truncated mutants. Rows show proteins fused to the Gal4-DNA-BD, columns show proteins fused to the Gal4-AD. Interactions found within the domains of the same proteins are represented in red, while the interactions observed between the domains of the two ORC proteins are embodied in blue. The filled pattern symbolises strong interactions, while the textured pattern-the weak ones.

3.2.4 Co-immunoprecipitation approach

An *in vitro* co-immunoprecipitation analysis was employed to detect and verify interactions between the *A.fulgidus* proteins and their domains. *In vitro* translation and labelling, in order to monitor interacting proteins in a co-immunoprecipitation assay is compatible with the two-hybrid system, and over the years, this approach has been established as a common way of confirming two-hybrid data. This allows for convenient verification of previously obtained results, without the necessity of expressing and purifying recombinant proteins in significant amounts. The proteins cloned into the binding domain vector had a c-myc epitope tag, while the proteins introduced into the activation domain vector had an HA epitope tag. Both T7 promoter and epitope tags were located downstream of the GAL4 coding sequences so that the activation domain and binding domain proteins were transcribed without their GAL4 domains. As a result, the interaction detected in a co-immunoprecipitation assay is specific to the activation protein and the bait. Activation domain and binding domain vectors carrying the archaeal sequences were transcribed and translated *in vitro* in rabbit reticulocyte lysates supplemented with [^{35}S] methionine (Promega), and were visualized via autoradiography. The proteins were then used in co-immunoprecipitation analyses (Matchmaker Co-IP Kit, Clontech). Pairs of proteins were incubated together with polyclonal HA antibody, followed by incubation with protein A beads. The beads were washed extensively to remove non-specifically associated proteins prior the elution of the immune complexes. Eluted HA-tagged or c-myc-tagged proteins, together only with the co-eluting protein, were separated by SDS-PAGE and proteins detected by autoradiography. C-myc-tagged murine p53 and HA-tagged SV40 large T antigen protein were used as positive controls; they interact in the co-immunoprecipitation assay *in vitro* and in the yeast two-hybrid screen *in vivo* (Clontech). These experiments did not confirm the results seen with the yeast two-hybrid system (Fig. 3.11, Fig. 3.12 and Fig. 3.13).

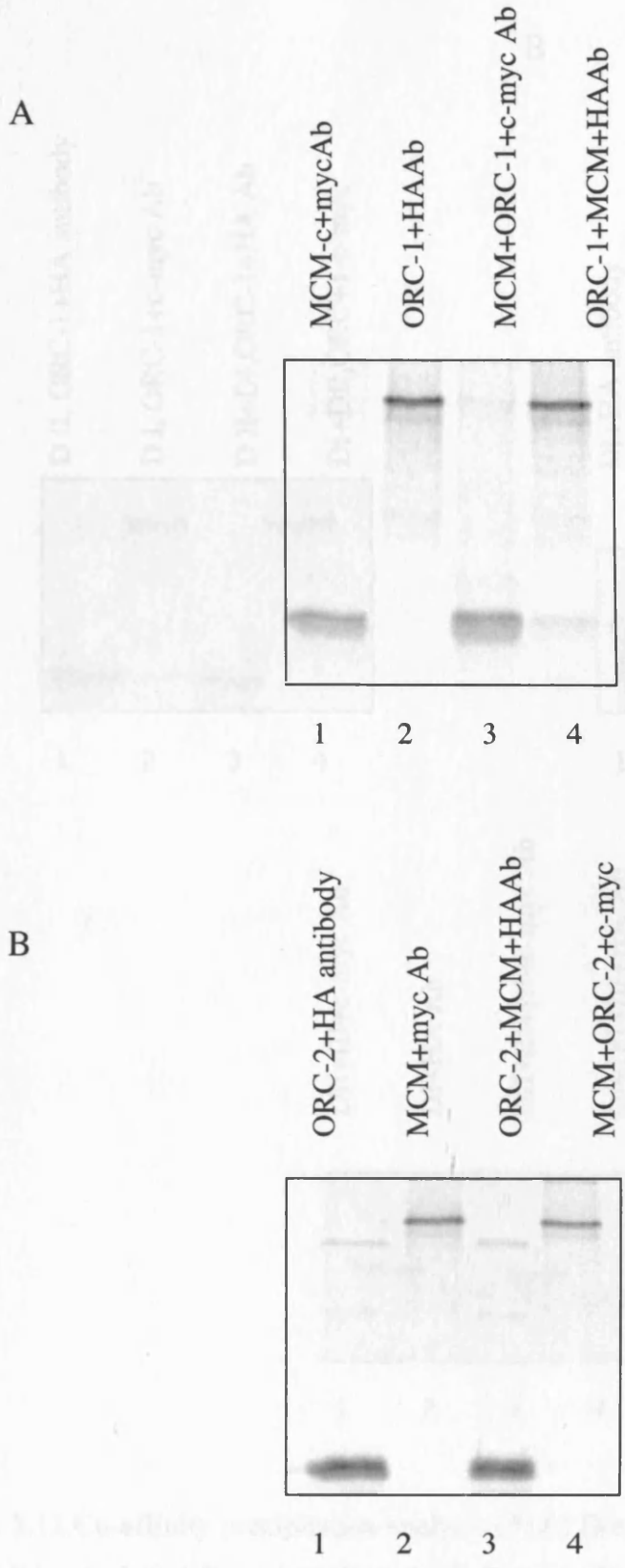


Fig. 3.11 Co-affinity precipitation analysis of the *Af*MCM and (A) *Af*ORC-1 and (B) *Af*ORC-2. *Af*MCM was detected with c-myc antibody, while *Af*ORC-1 and *Af*ORC-2 were detected with HA antibodies. Positive interactions were not observed using this method (lane 3 and lane 4)

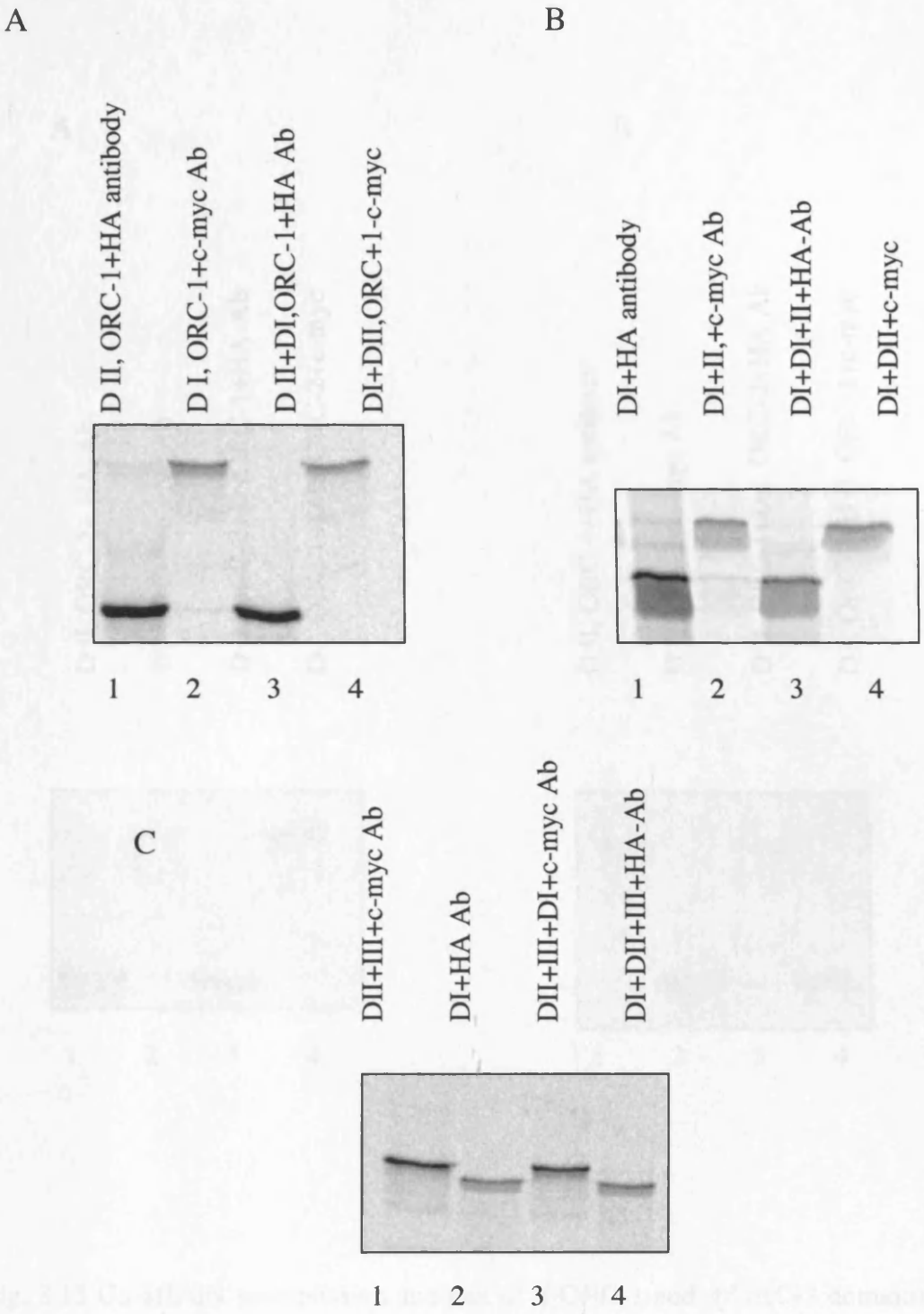


Fig. 3.12 Co-affinity precipitation analysis of (A) Domain I and Domain II, (B) Domain I and Domain I+II and (C) Domain I and Domain II+III of *Af*ORC-1. Domain I of *Af*ORC-1 was cloned into pGBKT7-BD and was detected with c-myc antibody, while Domain II, Domain I+II and Domain II+III of *Af*ORC-1 were cloned into pGADT7-AD and detected with HA antibodies. These experiments did not confirm the results observed with the two-hybrid system (lanes 3 and 4).

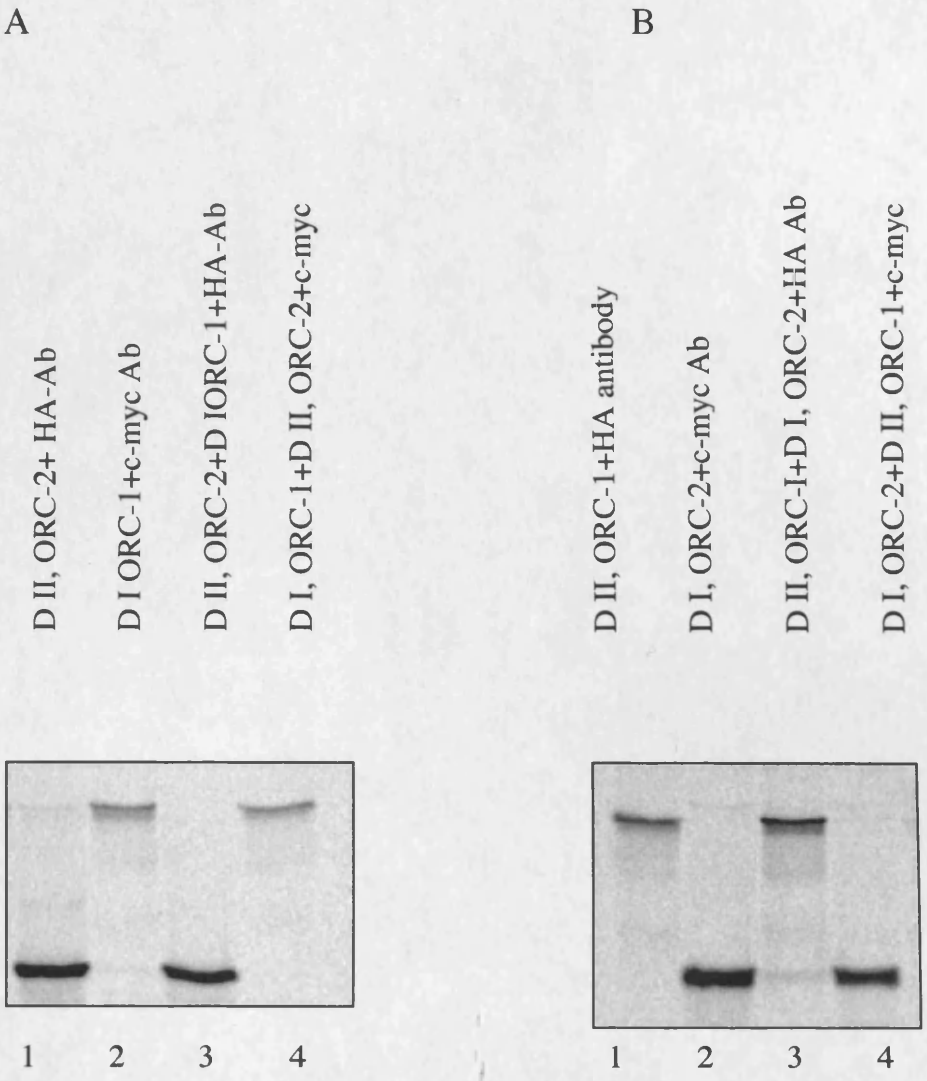


Fig. 3.13 Co-affinity precipitation analysis of *Af*ORC-1 and *Af*ORC-2 domains. (A) Domain I, *Af*ORC-1 and Domain II, *Af*ORC-2. Domain I of *Af*ORC-1 was cloned into pGBKT7-BD and was detected with c-myc antibody, while Domain II of *Af*ORC-2 was cloned into pGADT7-AD and detected with HA antibodies (B) Domain II, *Af*ORC-1 and Domain I, *Af*ORC-2. Domain I of *Af*ORC-2 was detected with c-myc antibody, while Domain II, *Af*ORC-1 with HA antibodies. These experiments did not confirm the results observed with the two-hybrid system (lanes 3 and 4).

3.2.5 *Yeast two-hybrid screen of A. fulgidus genomic DNA library*

*Af*MCM and the two *Af*ORC proteins were as baits and transformed separately into the AH109 strain. *A.fulgidus* genomic library (prepared by N. Cook) was cloned into the activation domain vector and transformed in Y187. The library contained at least 2.5 x 10^6 clones.

The pre-transformed genomic library and a bait culture were mated, and the diploid yeast mix was plated onto appropriate selection media containing X-α-gal to directly assay for α-galactosidase. In the primary screens a total of 1 x 10^7colonies were screened and 73 putative positive clones for *Af*MCM, 124 for *Af*ORC-1 and 117 for *Af*ORC-2 were observed. All clones were then re-tested for β-galactosidase activity (Materials and Methods, section 2) to eliminate some of the false positives. This resulted in 52 positive clones for *Af*MCM, 74 for *Af*ORC-1 and 86 for *Af*ORC-2. These positive clones were re-tested for their nutritional and reporter phenotypes. Sometimes library clones can contain more than one AD/library plasmid, but in the process of re-plating colonies on selective media, these plasmids are lost, while maintaining selective pressure on DNA-BD and AD vectors. Two such clones were discarded here; leaving 72 clones as confirmed positives for *Af*ORC/CDC6-1. The DNA inserts of all positive clones were amplified by colony PCR using AD vector-specific primers and the products were analysed on an agarose gel. Some products were clearly represented more than once, but all clones were analysed by DNA sequencing.

The nucleotide sequences obtained from the putative positives were used in standard BLAST searches of the nucleotide/protein database at the National Centre for Biotechnology Information (NCBI). The identified positive interactors of *Af*MCM, *Af*ORC-1 and *Af*ORC-2 proteins are presented in Table 3.2, 3.3 and 3.4. Three of the open reading frames encoded conserved hypothetical proteins: AF0995, AF1019 and AF0491. These three clones appeared often and showed positive interactions with all three bait proteins. Therefore their interaction needed to be confirmed by re-testing their binding specificity in the yeast system.

Protein	*Function*	*Number of appearances*
AF0232	Glutamine ABC transporter, permease protein	3
AF0380	Transmembrane ologosaccharyl transferase, putative	2
AF0491	**Conserved hypothetical protein**	**11**
AF0506	Iron-sulphur binding reductase	5
AF0770	Signal-transducing histidine kinase	2
AF0733	Thiamine monophosphate kinase (thiL)	4
AF0882	Asparaginase	2
AF0995	**Conserved hypothetical protein**	**8**
AF1019	**Conserved hypothetical protein**	**9**
AF1770	Dipeptide ABC transporter, ATP-binding protein	4
AF1968	Transcriptional regulatory protein, Rok family	2

Table 3.2 *Af*MCM-interacting proteins found in yeast two-hybrid screen against the genomic library. The three conserved hypothetical proteins used for further analysis are indicated in bold.

Protein	*Function*	Number of appearances
AF0034	3-ketoacyl-CoA thiolase	2
AF0087	Nitrate ABC transporter, ATP-binding protein (nrtC-1)	3
AF0491	**Conserved hypothetical protein**	**9**
AF0889	Ribose ABC transporter, permease protein	3
AF0950	Carbon monoxide dehydrogenase, iron sulfur subunit	4
AF0995	**Conserved hypothetical protein**	12
AF1019	**Conserved hypothetical protein**	13
AF1101	Acetyl-CoA decarbonylase/synthase, subunit epsilon	4
AF1222	Proline permease	3
AF1531	Conserved hypothetical protein	2
AF1763	Lipase, putative	2
AF1769	Dipeptide ABC transporter, permease protein	2
AF2129	Aspartate aminotransferase	5
AF2254	ATP-dependent RNA helicase, DEAD-family	4
AF2409	Iron-sulfur cluster binding protein	7

Table 3.3 *Af*ORC-1-interacting proteins found in yeast two-hybrid screen against the genomic library. The three conserved hypothetical proteins used for further analysis are indicated in bold.

Protein	*Function*	*Number of appearances*
AF0273	Sarcosine oxidase, subunit alpha	3
AF0491	**Conserved hypothetical protein**	10
AF0506	Iron-sulfur binding reductase	4
AF0745	Dihydroorotase dehydrogenase (pyrD)	2
AF0814	Conserved hypothetical protein	3
AF0995	**Conserved hypothetical protein**	14
AF1019	**Conserved hypothetical protein**	16
AF1101	Acetyl-CoA decarbonylase/synthase, subunit epsilon	4
AF1252	Oxaloacetate decarboxylase, subunit alpha	3
AF1368	Hydrogenase expression/formation protein	2
AF1490	LSU ribosomal protein L1P	5
AF1769	Dipeptide ABC transporter, permease protein	4
AF1813	TBP-interacting protein TIP49	5
AF1829	F420H2: quinone oxidoreductase, putative	3
AF2037	Translation initiation factor eIF-2B, subunit delta	3
AF2409	Iron-sulfur cluster binding protein	5

Table 3.4 *Af*ORC-2-interacting proteins found in yeast two-hybrid screen against the genomic library. The three conserved hypothetical proteins used for further analysis are indicated in bold.

Firstly, the three positive clones were mated with unrelated bait proteins (human lamin C, murine p53, and yeast elongation factors ELP 1-6) and with an "empty" DNA-BD plasmid. The three archaeal proteins interacted with all irrelevant baits (Fig. 3.14A). The interaction ability of the full-length positive clones was tested as well (Fig. 3.14B). As a result of the pairwise mating, all three clones proved to be false positives since their interactions with the archaeal bait proteins of interest were not specific. The three conserved hypothetical proteins isolated from the screen interacted with all unrelated bait proteins and when their full-length sequences were used in the screen, their interactions with the pre-RC proteins were not confirmed.

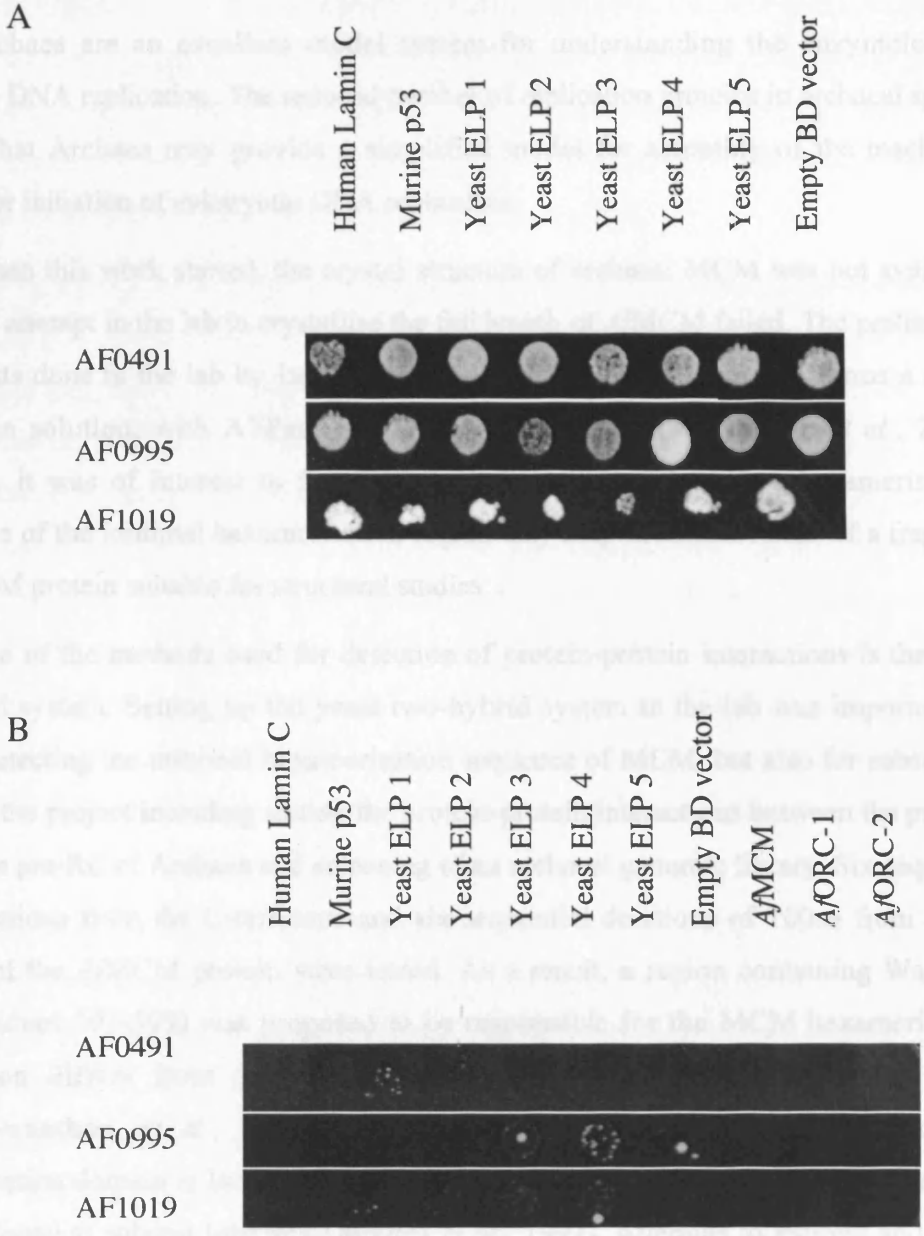


Fig. 3.14 Test for interaction specificity of AF0491, AF0995 and AF1019 conserved hypothetical proteins (A) Mating of the isolated clones with unrelated bait proteins, (B) Mating of the full-length hypothetical proteins with the *Af*MCM, *Af*ORC-1, *Af*ORC-2 and irrelevant bait proteins.

3.3 Conclusions and future work

Archaea are an excellent model system for understanding the enzymology of eukaryotic DNA replication. The reduced number of replication proteins in archaeal species suggests that Archaea may provide a simplified model for assembly of the machinery required for initiation of eukaryotic DNA replication.

When this work started, the crystal structure of archaeal MCM was not available. The initial attempt in the lab to crystallise the full length of *Af*MCM failed. The preliminary experiments done in the lab by Ian Grainge showed that *Af*MCM protein forms a single hexamer in solution, with ATPase and DNA helicase activity (Grainge *et al.*, 2003). Therefore, it was of interest to find the region(s) involved in MCM hexamerisation. Knowledge of the minimal hexamerisation region may help in identification of a fragment of the MCM protein suitable for structural studies.

One of the methods used for detection of protein-protein interactions is the yeast two-hybrid system. Setting up the yeast two-hybrid system in the lab was important not only for detecting the minimal hexamerisation sequence of MCM, but also for subsequent aspects of the project including testing the protein-protein interactions between the proteins involved in pre-RC of Archaea and screening of an archaeal genomic library. Six sequential 100aa deletions from the C-terminus and six sequential deletions of 100aa from the N-terminus of the *Af*MCM protein were tested. As a result, a region containing Walker A motif (residues 301-399) was proposed to be responsible for the MCM hexamerisation. This region differs from the one observed in *M.thermoautotrophicum* MCM (Kasiviswanathan *et al.*, 2004). However it is not surprising that the proposed oligomerisation domain is located around the Walker A motif. In several cases, this motif has been found at subunit interface (Mourez *et al.*, 1997). Attempts to express and purify the suggested oligomerisation region (residues 301-399) failed due to the insolubility of the expressed protein. Subsequently, larger fragments that included the interaction region were checked for protein expression. Only a fragment containing residues 1-399 was found to express as a soluble protein. This fragment was purified and its oligomeric state assessed by a Superdex 200 HR (10/30) gel filtration column. The gel filtration analysis showed that 1-

399*Af*MCM protein forms a mixture of different oligomeric forms (not a stable homogenous oligomer) and therefore was not suitable for structural studies. However, EM studies (carried out by Ian Grainge) of 1-399 *Af*MCM performed in the presence of double-stranded DNA showed that the protein forms hexameric rings with the DNA passing through the central channel of the ring. Future work in the lab will include crystallisation trials of the 1-399 *Af*MCM in the presence of double-stranded DNA.

The *AfMCM* truncation library was used also to test the possible interactions with ORC and CDC6. Although interactions between these proteins were detected in other archaeal species (Shin *et al.*, 2003; De Felice *et al.*, 2003; Kasiviswanathan *et al.*, 2004; Kasiviswanathan *et al.*, 2005), they were not identified in *A.fulgidus* using the two-hybrid approach and co-immunoprecipitation analysis. Positive interactions were observed only between some domains of ORC-1 and ORC-2. Domain I and Domain II appeared to play a central role in these interactions. They have been found to be interacting partners not only within the same protein but between the two ORC proteins as well. However, these interactions were not confirmed with the co-immunoprecipitation analysis. Future work in the lab could involve isothermal titration calorimetry and analytical ultracentrifuge analyses for a further characterisation of the protein-protein interactions between these archaeal proteins.

In Archaea, no homologues of MCM10, CDC45 and CDT1 have been found based on sequence similarities. However, it seems likely that orthologues of some of these proteins might be present in Archaea. Therefore, two-hybrid screening of archaeal genomic library was utilised to search for new partners of the archaeal pre-RC proteins. The library screening did not produce any significant positive interactors. The reasons for that could be many. Not all of the reported protein-protein interactions can be detected in the yeast two-hybrid system and are therefore considered as false negatives. False negatives can be cause by different characteristics of the yeast two-hybrid system. First, the activation domain and binding domain proteins may fail to localise to the yeast nucleus. Second, two proteins may be unable to function within the context of a binding domain and activation domain fusion. Third, the interaction between two proteins may depend on post-translational modifications

that are absent in yeast cells. Finally, it has been reported that a number of protein-protein interactions can only be detected in the two-hybrid system when either one or the other protein is truncated. This suggests that large-scale two-hybrid analyses are useful to obtain partial coverage of protein-protein interactions within a proteome of interest. However, in order to achieve complete protein interaction network, alternative large-scale approaches have to be applied. Several proteomics technologies have been developed and adapted to chart protein-protein interactions on a proteome-wide scale. Protein microarrays allow the detection of binary protein-protein interactions *in vitro*. In this method, the proteins are covalently attached to a solid support and screened with fluorescently labelled probes. Although such protein microarrays are in principal suitable to screen in parallel for proteome-wide interactions, the techniques has only been applied to identify domain-domain interactions and antibodies and antigens (Espejo *et al.*, 2002). Mass spectrometry analysis is another powerful and sensitive technique, which allow the detection of peptides from purified protein complexes. Traditional affinity methods using specific antibodies or affinity ligands can be combined with mass spectrometry to identify protein complexes from cells or tissues (Husi *et al.*, 2000). Obviously these approaches rely on the availability of specific antibodies or other capturing ligands. Future work in the lab could involve affinity purification of protein complexes coupled to mass spectrometry detection. Compared with the yeast two-hybrid, this approach is more physiological, because actual molecular assemblies made up by all combinations of direct and cooperative interactions are analysed *in vivo*, rather than re-constituted bimolecular interactions *in vitro*. Whereas in the yeast two-hybrid system both interactors are expressed as fusion proteins, in the complex purification approach only one component of the complex is expressed as a fusion protein, minimising possible steric interference.

CHAPTER 4

BIOCHEMICAL CHARACTERISATION OF ORC-1 AND ORC-2 PROTEINS FROM *A.pernix*

4.1 Introduction

4.1.1 CDC6/ORC1 homologues in Archaea

The process of initiating DNA replication involves a series of proteins that specifically assemble on DNA origins in a precisely determined manner. Formation of the pre-replication complex in both bacteria and eukarya is regulated by ATP. Most of the proteins known to be required for pre-RC formation include consensus nucleotide binding motifs and are members of the AAA+ family of proteins (Neuwald *et al.*, 1999). Consistent with ATP regulating these proteins, mutations expected or known to eliminate ATP binding or hydrolysis result in inhibition of pre-RC formation (Sekimizu *et al.*, 1987; Perkins and Diffley, 1998; Weinreich *et al.*, 1999a; Lee *et al.*, 2000; Lee and Bell, 2000; Klemm and Bell, 2001; Chesnokov *et al.*, 2001; Bell and Dutta, 2002).

In bacteria, DnaA protein is the universal initiator of chromosomal DNA replication DnaA can exist in a complex with ATP or ADP (Sekimizu *et al.*, 1988a; Bramhill and Kornberg, 1988; Hwang and Kornberg, 1992; Speck and Messer, 2001). DnaA protein binds co-operatively to the five DnaA boxes (R1-5) in the *oriC* of the chromosome in an ATP bound form (Bramhill and Kornberg, 1988). It is thought that this binding induces a strain on the hydrogen bonding between sister DNA strands, which causes local distortion of an AT-rich region, located near the left boundary of *oriC* to form an open complex in an ATP-dependent manner (Sekimizu *et al.*, 1987; Bramhill and Kornberg, 1988; Skarstad *et al.*, 1990; Hwang and Kornberg, 1992). DnaA-ADP complex is inactive in terms of origin binding and unwinding (Speck *et al.*, 1999). DnaB, the replicative helicase, is transferred to the replication origin by the helicase-loading factor DnaC (Konieczny and Helinski, 1997). DnaC is a dual ATP/ADP switch protein (Davey *et al.*, 2002). ATP is not required for DnaC to load DnaB onto ssDNA (Davey *et al.*, 2002). However, DnaC-ATP inhibits DnaB helicase activity, while the DnaC-ADP does not (Davey *et al.*, 2002). Therefore, ATP hydrolysis by DnaC is triggered by DnaB and ssDNA (Davey *et al.*, 2002).

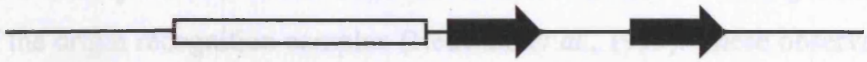
In contrast to replication initiation in bacteria, which starts at a single chromosomal origin and is controlled by a single protein DnaA (Kaguni, 1997; Speck *et al.*, 1997; Weigel

et al., 1997; Messer *et al.*, 2001), replication initiation in eukaryotes starts at multiple origins and is controlled by a six subunit ORC complex (Rowley *et al.*, 1995; Bielinsky *et al.*, 2001; DePamphilis, 2003). ORC is believed to serve as a platform on which the pre-replication complex (pre-RC) is assembled (Kornberg and Baker, 1992). Although few ORC subunits share certain structural features with DnaA, most notably a series of ATP-binding motifs, extensive homology does not exist between the two initiators. As a result, it has not been clear to what extent the mechanisms of replication initiation in eukaryotes are related to those of bacteria. ATP binding and hydrolysis are important for ORC function. ATP binding but not hydrolysis by ORC is required for DNA binding (Bell and Stillman, 1992; Klemm *et al.*, 1997; Austin *et al.*, 1999; Chesnokov *et al.*, 2001). Double-stranded origin DNA inhibits ATP hydrolysis by ORC, while ssDNA stimulates ATP hydrolysis in a length-dependent manner (Lee *et al.*, 2000). These findings suggest that, once bound to the origin, ORC is retained in an ATP-bound state and that DNA unwinding at the origin stimulates ATP hydrolysis by ORC. Targeting of ORC to specific chromosomal locations could be accomplished through its interactions with another pre-RC component CDC6. CDC6 modulates ORC binding by increasing the stability of ORC on chromatin (Harvey and Newport, 2003) and inhibits ORC binding to non-specific DNA (Mizushima *et al.*, 2000). It is suggested that the ATP-bound conformation of ORC is needed for interaction with CDC6 (Klemm and Bell, 2001). The ATP-bound form of CDC6 is required for its association with ORC (Perkins and Diffley, 1998), enhancing the specificity and affinity of ORC for origins (Mizushima *et al.*, 2000). Once localised at replication origins, CDC6 helps to recruit and load MCM2-7 helicase onto DNA in an ATP-dependent manner (Coleman *et al.*, 1996; Romanowksi *et al.*, 1996; Donovan *et al.*, 1997; Tanaka *et al.*, 1997, Perkins and Diffley, 1998; Weinreich *et al.*, 1999). CDC6 protein shows significant similarity to the clamp loaders of DNA polymerases (Perkins and Diffley, 1998), which assemble a ring-shaped sliding clamp that associates with the replicative polymerase on primed DNA and serves as its processivity factor (Jeruzalmi *et al.*, 2004). The similarities between CDC6 and the clamp loaders of DNA polymerase, together with the genetic and *in vivo* observations, suggest that CDC6 may act as a loading factor for MCM (Perkins and Diffley, 1998; Weinreich *et al.*, 1999). Following MCM loading and the initiation of S

phase, CDC6 is phosphorylated by CDKs, transported to the cytoplasm, and eventually degraded in an ubiquitin-dependent manner (Saha *et al.*, 1998).

In archaea, DNA replication proteins resemble those of eukarya but assume somewhat simpler forms. Within the repertoire of archaeal homologues of eukaryotic replication proteins, the most likely candidates for initiator proteins that mediate origin recognition are the archaeal ORC/CDC6 homologues (Edgell and Doolittle, 1997; Baker and Bell, 1998; Leipe *et al.*, 1999). Most archaeal species that have been sequenced contain at least one such homologue (Myllykallio and Forterre, 2000). The archaeal ORC/CDC6 proteins are related to both the eukaryotic ORC1 subunit of the origin recognition complex and CDC6. This observation raises the question as to whether the archaeal ORC/CDC6 may play a dual role in both origin recognition and MCM recruitment. As a result, different research labs adopted different nomenclatures for the names of the archaeal homologues. Insight into the mechanism of ORC/CDC6 function has come from the determination of the crystal structure of ORC/CDC6 from the crenarchaeons *P.aerophilum* and *A.pernix* (Liu *et al.*, 2000; Singleton *et al.*, 2004). The three dimensional structure of the archaeal ORC/CDC6 homologue revealed the expected conserved domains found in other members of the AAA+ superfamily of ATPases (Ogura and Wilkinson, 2001). Together with the nucleotide binding domains, the structure disclosed the presence of a winged helix-turn-helix (WH) domain that has been proposed to interact with DNA (Liu *et al.*, 2000; Singleton *et al.*, 2004; Robinson *et al.*, 2004). It has been proposed that interaction between WH proteins and their DNA targets is *via* the most positively charged surfaces that interact with the phosphodiester backbone (Gajiwala *et al.*, 2000). Based on that and the sequence comparison of ORC/CDC6 homologues of variety of archaeal species, two classes of WH proteins were defined (Fig. 4.1) (Singleton *et al.*, 2004). Class I proteins have a conserved positively charged recognition helix, while in class II proteins the number of positively charged residues is greater in the wing region.

CLASS I



```
CONSENSUS   LTxRRISDLIxELDMLGILxxxVVSRGR-YGRTKEIxL

ApeORC-1    LEHVTLRRVSGIISELDMLGIVKSRVVSRGR-YGKTREVSLDADR
SsORC-1     VEAVTQRRVSDIINELDMVGILTAKVVNRGR-YGKTKEIGLAVDK
PfORC       LEPLTQRRISDLINELDMLGIINAKVVSKGR-YGRTKEIRLNVTS
AfORC-1     VDPLTQRRISDLISELDMLGILNSVVISKGK-YGRTREMKVEADE
MthORC-1    GDSVSYRRIFDFINEIEMLGLISTNTVSRGRGKGRTNIIDLQCET
```

CLASS II

```
ApeORC-2    ASLTMYNVKPRGYTQYHIYLKHLTSLGLVDAKPS--GRGMRGRTTLFR
SsORC-2     -LSREFNEEPRRHTQVYEYLRKLKVIGIINTRQS--GKGMRGRTTLVS
PaeORC-2    IVCEEYGERPRVHSQLWSYLNDLREKGIVETRQNKRGEGVEGRTTLIS
AfORC-2      ------EIKMSYRKYYNILEKLERLKLIEISFG--EKG-RGKTRYVQ
MthORC-2     ------STGSSYSSFNRIIEKLEFLRLIDTRLT--GKGVRGNSRILI

CONSENSUS           Xxxxxxx LxxLxxLGIIxxxxx--GxGxRGR
```

Fig. 4.1 Sequence alignment of archaeal ORC/Cdc6 proteins showing the division between class I and class II proteins (Singleton *et al.*, 2004). The secondary structure of the winged helix region is shown at the top (helix as open rectangle, beta-strands as arrows). Conserved identical residues are highlighted in green and those conserved in the majority of proteins are shown in yellow. Archaeal species: *Ape*, *Aeropyrum pernix; Ss*, *Sulfolobus solfataricus*; *Pf*, *Pyrococcus furiosus*; *Pae*, *Pyrobaculum aerophilum*; *Af*, *Archaeoglobus fulgidus*; *Mth*, *Methanobacterium thermoautotrophicum*.

4.1.2 Aim of the project

Pair-wise global sequence alignments the *Ape*ORC proteins revealed that they share about 33% similarity with the *S.cerevisiae* CDC6 factor and 20% similarity with the ORC1 subunit of the origin recognition complex (Neuwald *et al.*, 1999). These observations raise the question as to whether the archaeal ORC/CDC6 may play a dual role in both origin recognition and MCM recruitment. To address this question, recombinant ORC/CDC6 homologues from *A.pernix* were expressed and purified and their affinity for single-stranded and double-stranded DNA examined. ATP regulates formation of the pre-replication complex in both, bacteria and eukarya. Therefore, the ATPase activities of both *A.pernix* ORC/CDC6 homologues were assessed.

4.2 Sequence homology and analysis

The analysis of the *A.pernix* genome sequence revealed the presence of two ORFs coding for putative homologues of the eukaryotic ORC1/CDC6 proteins (Kawarabayasi *et al.*, 1999). Henceforth, the *APE*0152 will be referred as *Ape*ORC-1 and *APE*0475 as *Ape*ORC-2 (Singleton *et al.*, 2004). The Walker A and Walker B motifs, typically found in the AAA+ proteins, are present in *Ape*ORC-1 and *Ape*ORC-2 (Neuwald *et al.*, 1999) (Fig.4.2).

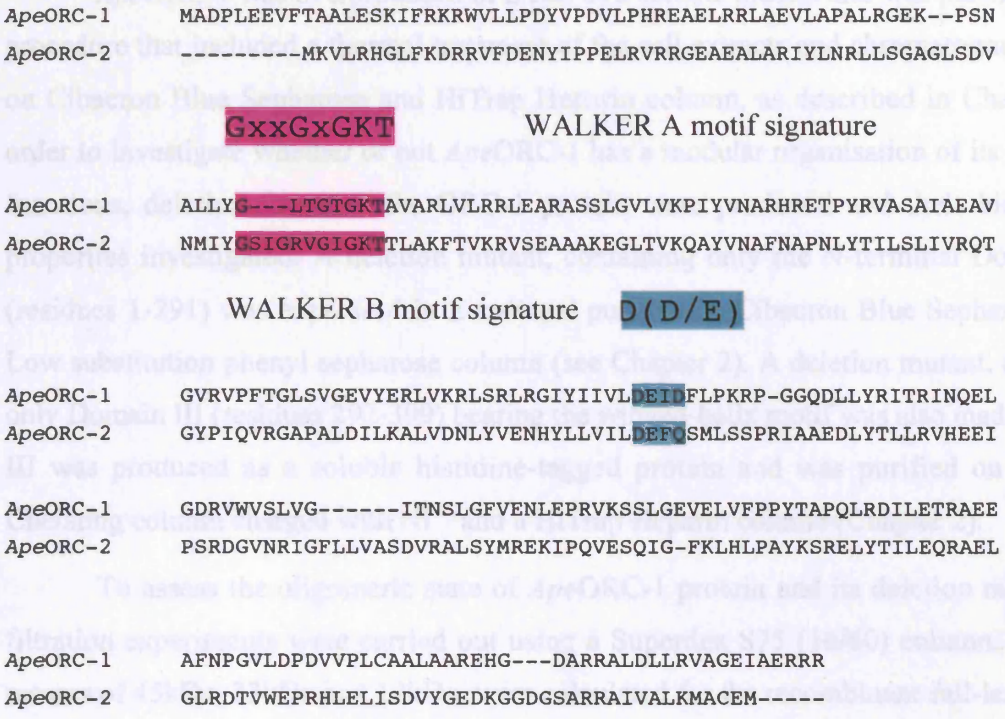


Fig. 4.2 Sequence alignment of *A.pernix* ORC-1 and ORC-2 proteins showing the location of Walker A (in pink) and Walker B (in blue) motifs (Walker *et al.*, 1982).

4.3 *A.pernix* ORC-1 protein

4.3.1 Identification and purification

*Ape*ORC-1 was overproduced in *E.coli* as a soluble protein and was purified using a procedure that included a thermal treatment of the cell extracts and chromatographic steps on Cibacron Blue Sepharose and HiTrap Heparin column, as described in Chapter 2. In order to investigate whether or not *Ape*ORC-1 has a modular organisation of its biological functions, deletion forms of the ORC-1 protein were produced and their biochemical properties investigated. A deletion mutant, containing only the N-terminal Domain I+II (residues 1-291) was expressed in *E.coli* and purified on Cibacron Blue Sepharose and a Low substitution phenyl sepharose column (see Chapter 2). A deletion mutant, containing only Domain III (residues 292-399) bearing the winged-helix motif was also made. Domain III was produced as a soluble histidine-tagged protein and was purified on a HiTrap Chelating column charged with Ni^{2+} and a HiTrap Heparin column (Chapter 2).

To assess the oligomeric state of *Ape*ORC-1 protein and its deletion mutants, gel filtration experiments were carried out using a Superdex S75 (16/60) column. Molecular masses of 45kDa, 33kDa and 12kDa, were calculated for the recombinant full-length ORC-1, DI+II and DIII, respectively (Fig. 4.3). Therefore, initially, it was concluded that *Ape*ORC-1 is a monomer in solution. However, more recent analytical ultracentrifugation experiments suggested that the full-length ORC-1 protein can form dimers at high concentrations (M.Gaudier and S.Westcott, unpublished data). The dimerisation interface is believed to be within Domain I+II since Domain III alone did not dimerise (M.Gaudier and S.Westcott, unpublished data). Dimerisation of archaeal ORC/CDC6 homologue was also observed by differential scanning calorimetry experiments with *M.thermoautotrophicum* ORC-2 (Shin *et al.*, 2003a). *In vivo* and *in vitro* studies of the eukaryotic CDC6 homologue from *S.cerevisiae* have also suggested oligomerisation (Saha *et al.*, 1998; Herbig *et al.*, 1999).

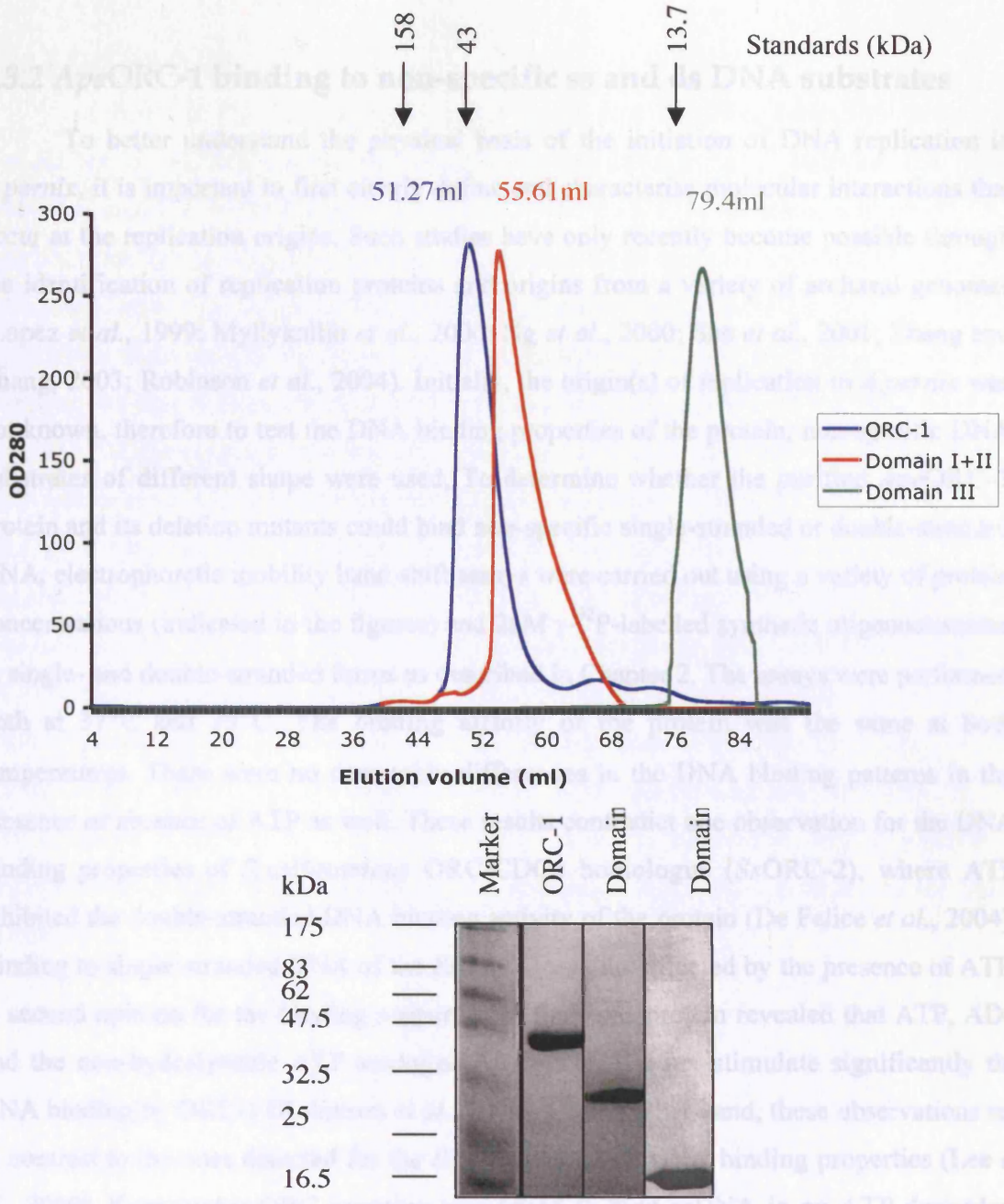


Fig. 4.3 Gel filtration analysis of *Ape*ORC-1 protein. (A) Elution profile of the purified proteins from Superdex S75 (16/60) gel filtration column. The arrows indicate the position of the protein markers used to calibrate the column (Aldolase, Ovalbumin and RNaseA). (B) Coomassie blue-stained SDS-PAGE of the full-length, Domain I+II and Domain III proteins after GF purification procedure.

4.3.2 *Ape*ORC-1 binding to non-specific ss and ds DNA substrates

To better understand the physical basis of the initiation of DNA replication in *A.pernix*, it is important to first clearly define and characterise molecular interactions that occur at the replication origins. Such studies have only recently become possible through the identification of replication proteins and origins from a variety of archaeal genomes (Lopez *et al.*, 1999; Myllykallio *et al.*, 2000; Ng *et al.*, 2000; She *et al.*, 2001; Zhang and Zhang, 2003; Robinson *et al.*, 2004). Initially, the origin(s) of replication in *A.pernix* was not known, therefore to test the DNA binding properties of the protein, non-specific DNA substrates of different shape were used. To determine whether the purified *Ape*ORC-1 protein and its deletion mutants could bind non-specific single-stranded or double-stranded DNA, electrophoretic mobility band shift assays were carried out using a variety of protein concentrations (indicated in the figures) and 2nM γ-^{32}P-labelled synthetic oligonucleotides in single- and double-stranded forms as described in Chapter 2. The assays were performed both at 37°C and 75°C. The binding affinity of the protein was the same at both temperatures. There were no detectable differences in the DNA binding patterns in the presence or absence of ATP as well. These results contradict one observation for the DNA binding properties of *S.solfataricus* ORC/CDC6 homologue (*Ss*ORC-2), where ATP inhibited the double-stranded DNA binding activity of the protein (De Felice *et al.*, 2004). Binding to single-stranded DNA of the *Ss*ORC-2 was not affected by the presence of ATP. A second opinion for the binding properties of the same protein revealed that ATP, ADP and the non-hydrolysable ATP analogue AMP-PNP, did not stimulate significantly the DNA binding by ORC-1 (Robinson *et al.*, 2004). On the other hand, these observations are in contrast to the ones detected for the *S.cerevisiae* ORC DNA binding properties (Lee *et al.*, 2000). *S.cerevisiae* ORC complex was found to bind ssDNA in an ATP dependent manner. Although ATP is not required for the *Sc*ORC to bind ds or ssDNA, the presence of ATP increases the affinity of the protein for ds origin DNA, but weakens the affinity for ssDNA (Lee *et al.*, 2000).

4.3.2.1 Full-length ORC-1

As shown in Fig. 4.5 and Fig. 4.6, full-length *A.pernix* ORC-1 was able to bind ss and dsDNA of different forms, expressing a preference for longer T bubbles. The DNA substrates used for the binding assays are shown on Fig. 4.4. To quantitatively analyse the data for these interactions, the binding curves were fitted to an independent binding site equation (Chapter 2). This equation describes the association of a protein with a ligand at equilibrium when each binding event is entirely independent and in such cases can be used to derive the dissociation constant for the interaction. Recent work on the DNA binding activity of *M.thermoautotrophicum* ORC homologues show that both proteins bind longer DNA substrates in a cooperative manner (Capaldi and Berger, 2004). Indeed, when binding data were fitted to the Hill equation, sigmoidal binding curves were observed, which suggest cooperative binding (Fig. 4.5 and Fig. 4.6). Binding curves were repeated in triplicate and the plots show the average and standard deviations from the three data sets. In the Hill analysis, the Hill coefficient is an index of the degree of cooperativity, and for a positively cooperative system has a value greater than unity with a maximum equal to the number of sites per oligomeric protein molecule. When the binding data were fitted to the Hill equation (Chapter 2), Hill coefficients were >1, consistent with a positively cooperative binding mechanism between the full-length *Ape*ORC-1 protein and the DNA substrates (Table 4.1).

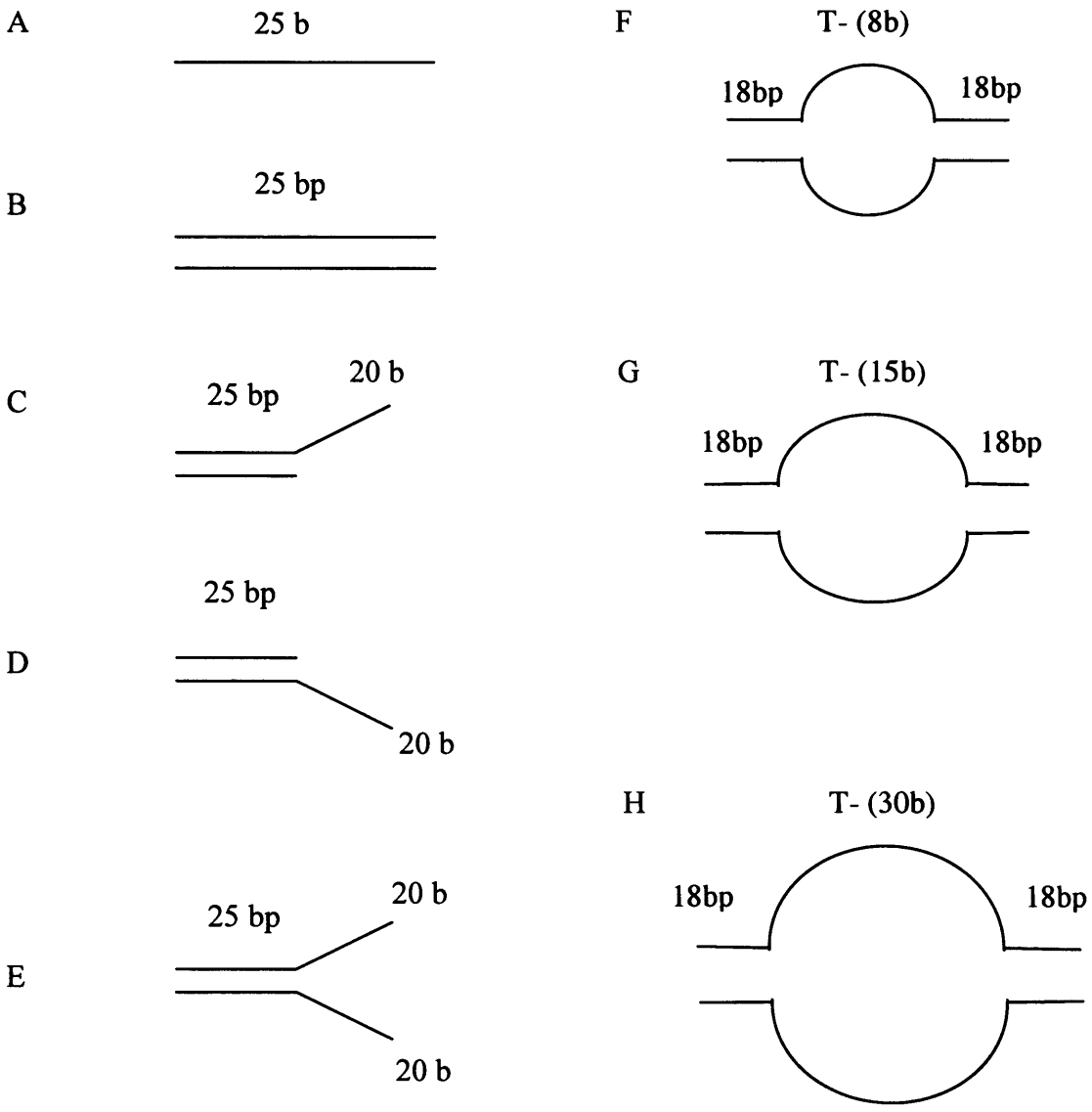


Fig. 4.4 Diagrammatic representation of the nucleotide-based substrates used in Chapter 4 and Chapter 5 (sequences of the oligonucleotides can be found in Materials and Methods). (A) Single-stranded DNA; (B) Blunt duplex; (C) 3'-tailed duplex; (D) 5'-tailed duplex; (E) Flayed duplex; (F) T(8) bubble substrate; (G) T(15) bubble substrate; (H) T(30) bubble substrate.

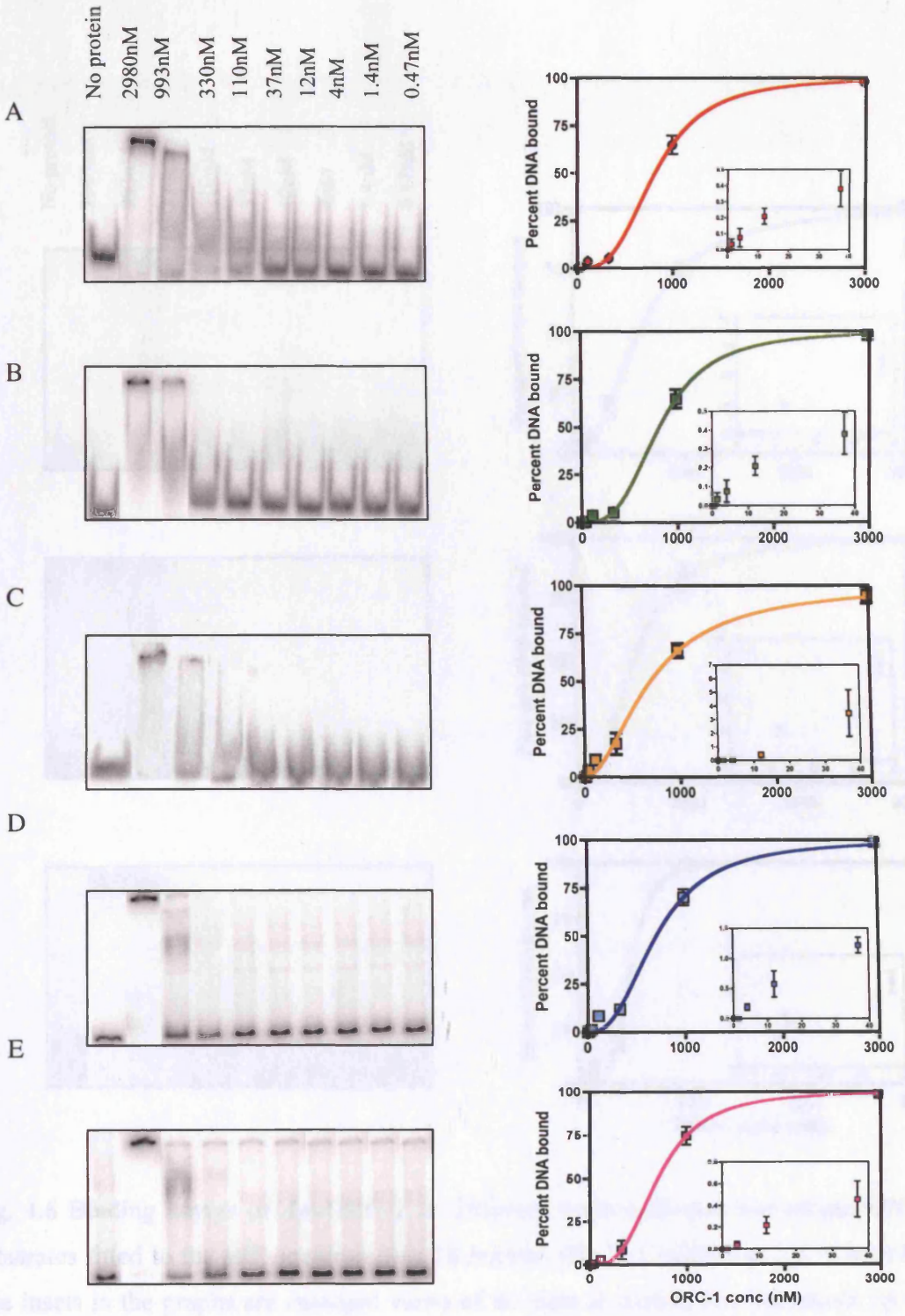


Fig. 4.5 Binding assays of *Ape*ORC-1 to different non-origin DNA substrates fitted to the Hill equation. (A) Single-stranded; (B) Blunt duplex; (C) 3'-tailed; (D) 5'-tailed; (E) Flayed duplex. The inset of the graph represents the binding affinity of *Ape*ORC-1 at concentrations up to 40nM.

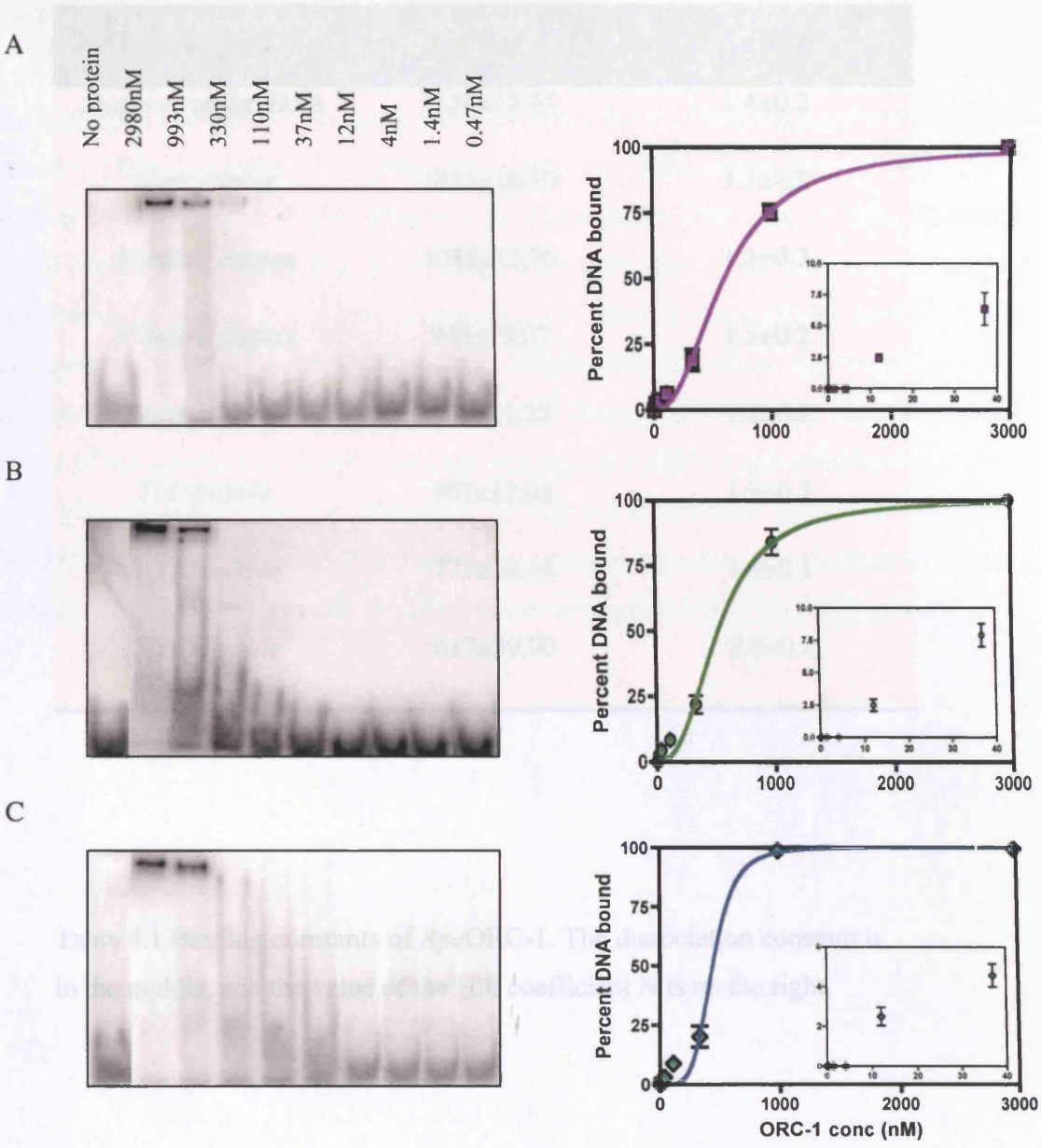


Fig. 4.6 Binding assays of *Ape*ORC-1 to different bubble-shaped non-origin DNA substrates fitted to the Hill equation. (A) T8 bubble; (B) T15 bubble; (C) T30 bubble. The insets in the graphs are enlarged views of the data at protein concentrations up to 40nM. The points on the plot are the averaged values from three independent binding experiments with the standard deviation indicated by the error bars.

DNA substrate	K_D *(nM)*	N
Single-stranded DNA	1154±12.44	1.4±0.2
Blunt duplex	1331±16.50	1.3±0.2
3'-tailed duplex	1085±13.70	1.3±0.2
5'-tailed duplex	945±19.07	1.5±0.2
Flayed duplex	900±11.28	1.2±0.2
T(8) bubble	867±12.05	1.9±0.2
T(15) bubble	771±18.44	1.9±0.1
T(30) bubble	617±39.90	2.0±0.2

Table 4.1 Binding constants of *Ape*ORC-1. The dissociation constant is in the middle, and the value of the Hill coefficient N is on the right.

4.3.2.2 Domain I+II and Domain III

It has been previously shown that ORC proteins comprise three domains (Liu *et al.*, 2000; Singleton *et al.*, 2004). Domains I and II contain the ATP-binding site, while domain III binds to DNA (De Felice *et al.*, 2003; Singleton *et al.*, 2004). As expected, Domain I+II of *Ape*ORC-1 did not bind single-stranded or double-stranded DNA (Fig. 4.7). As observed for ORC homologues from *A.fulgidus* (Grainge *et al.*, 2003) and for all three homologues in *S.solfataricus* (De Felice *et al.*, 2003; Robinson *et al.*, 2004), Domain III was able to form a stable complex with DNA, albeit with lower affinity than that of the full-length protein. This observation is consistent with the proposal that Domain III is the DNA-binding site on the basis of the structural similarity between this region of ORC-1 and other known DNA-binding proteins that contain the WH motif. As Fig. 4.8 shows there was no cooperativity of binding when only Domain III is used. This probably explains the reduction in overall affinity of the isolated domain for the DNA targets compared to the full-length protein.

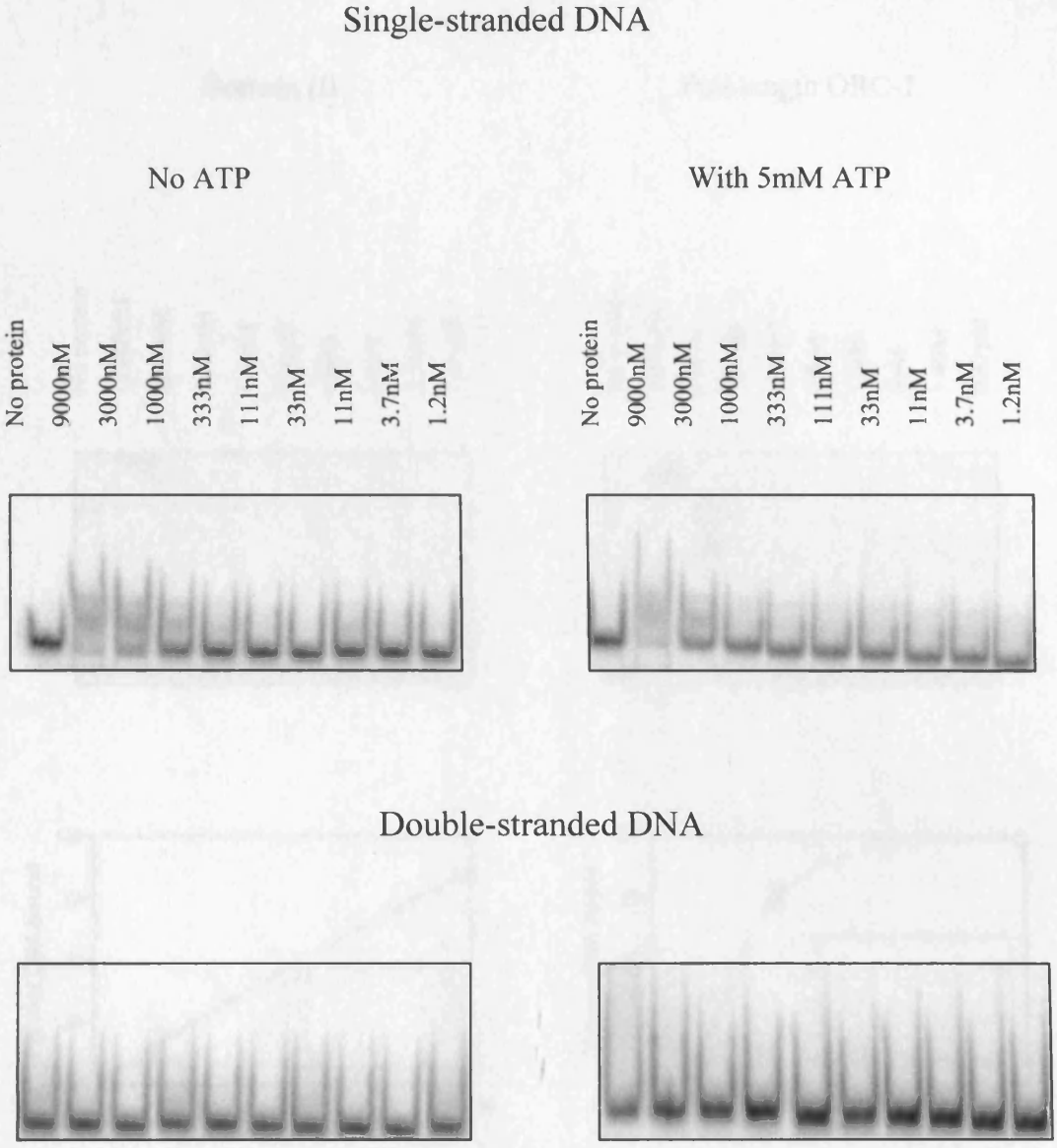


Fig. 4.7 Domain I+II of *Ape*ORC-1 binding to non–specific DNA. The binding reactions were performed in the presence or absence of 5mM ATP per reaction. As expected Domain I+II did not bind ss or ds DNA.

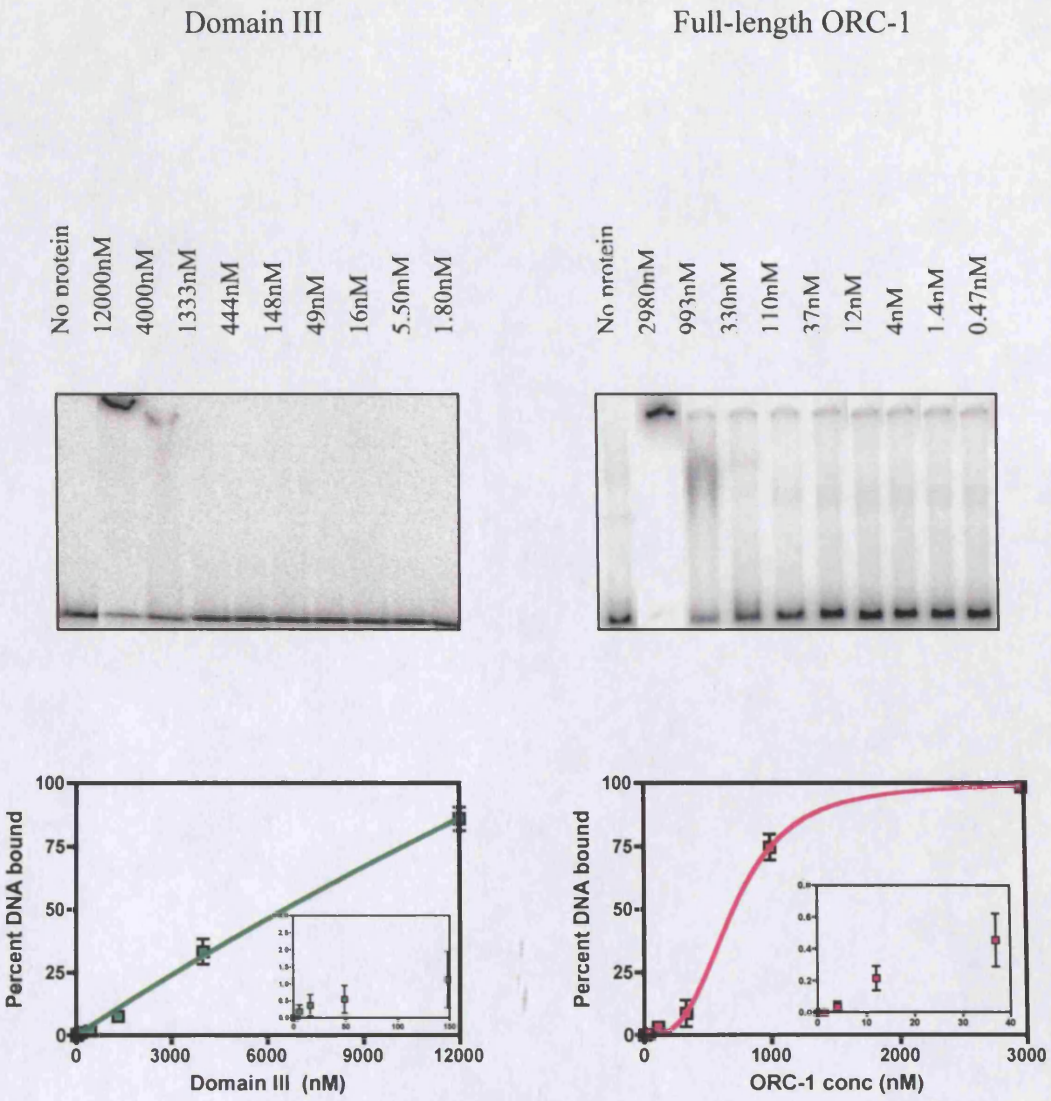


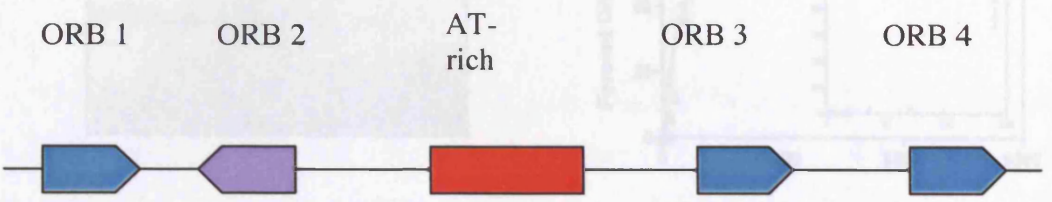
Fig. 4.8 Comparison between the DNA binding affinity of Domain III and the full-length *Ape*ORC-1 to a flayed duplex DNA. The insets of the graphs represent the binding affinity of Domain III or full-length *Ape*ORC-1 at concentrations up to 150 nM or 40nM, respectively.

4.3.3 *Ape*ORC-1 binding to origin DNA

Consensus sequences, referred to as origin recognition boxes (ORB), were identified in a number of archaeal species (Myllykallio *et al.*, 2000; Robinson *et al.*, 2004; Capaldi and Berger, 2004). Such boxes were originally found in *S.solfataricus* replication origins (Robinson *et al.*, 2004). These sites can be in similar orientations but frequently one or more of these is reversed relative to the others. On the basis of these ORB consensus sequences, the location of a replication origin in *A.pernix* was predicted (Robinson *et al.*, 2004). Although archaeal replication origins are frequently closely associated with ORC genes, the predicted *A.pernix* origin is not close to an ORC gene (Lundgren *et al.*, 2004).

In order to investigate whether the proposed site might indeed be a replication origin in *A.pernix*, the DNA binding properties of ORC-1 in this region were analysed. ORC-1 binds to a series of four repeats at the proposed origin hereafter referred to as *Ori1*. Each of the repeats contains an ORB sequence, with three sites in a similar orientation, but the forth being inverted relative to the others. An AT-rich sequence is situated at the centre of *Ori1* between the ORB2 and ORB3 boxes (Fig. 4.9) (I. Grainge). It has been shown that DnaA binds repeating elements within the replication origin cooperatively as a monomer and oligomerises upon binding (Doran *et al.*, 1999; Messer *et al.*, 2001; Majka *et al.*, 2001). Therefore, it is possible that *Ape*ORC-1 binds to multiple sites of the replication origin in a manner similar to DnaA. The basis for this cooperativity of binding between and within the repeat sequences of *Ori1* from *A.pernix* was investigated.

Initially, the DNA binding affinity of ORC-1 protein to each of the 26 base pairs-long ORB boxes was analysed on 4% SDS-PAGE gels (Fig. 4.10). Binding to each of the four single boxes show hyperbolic binding properties and fit well to the Hill equation (Chapter 2). ORB sequences are poor palindromes and it is unlikely that ORC proteins bind to different sequences with the same affinity. Indeed, ORC-1 protein binds more tightly to ORB1 and ORB4, which have the same sequence, than to ORB2 and ORB3 that are different.



ORB 1

GCTCCACAGGAAACGGAGGGGTCCTT

ORB 2

TGATGCTCCAGAGGAACCTGGGGGGT

ORB 3

AGAGGATCCAGCGGAAAACAGGGGGT

ORB 4

GCCGTCTCCACAGGAAACGGAGGGGT

Fig. 4.9 *A.pernix Ori1* structure and origin recognition box sequences.

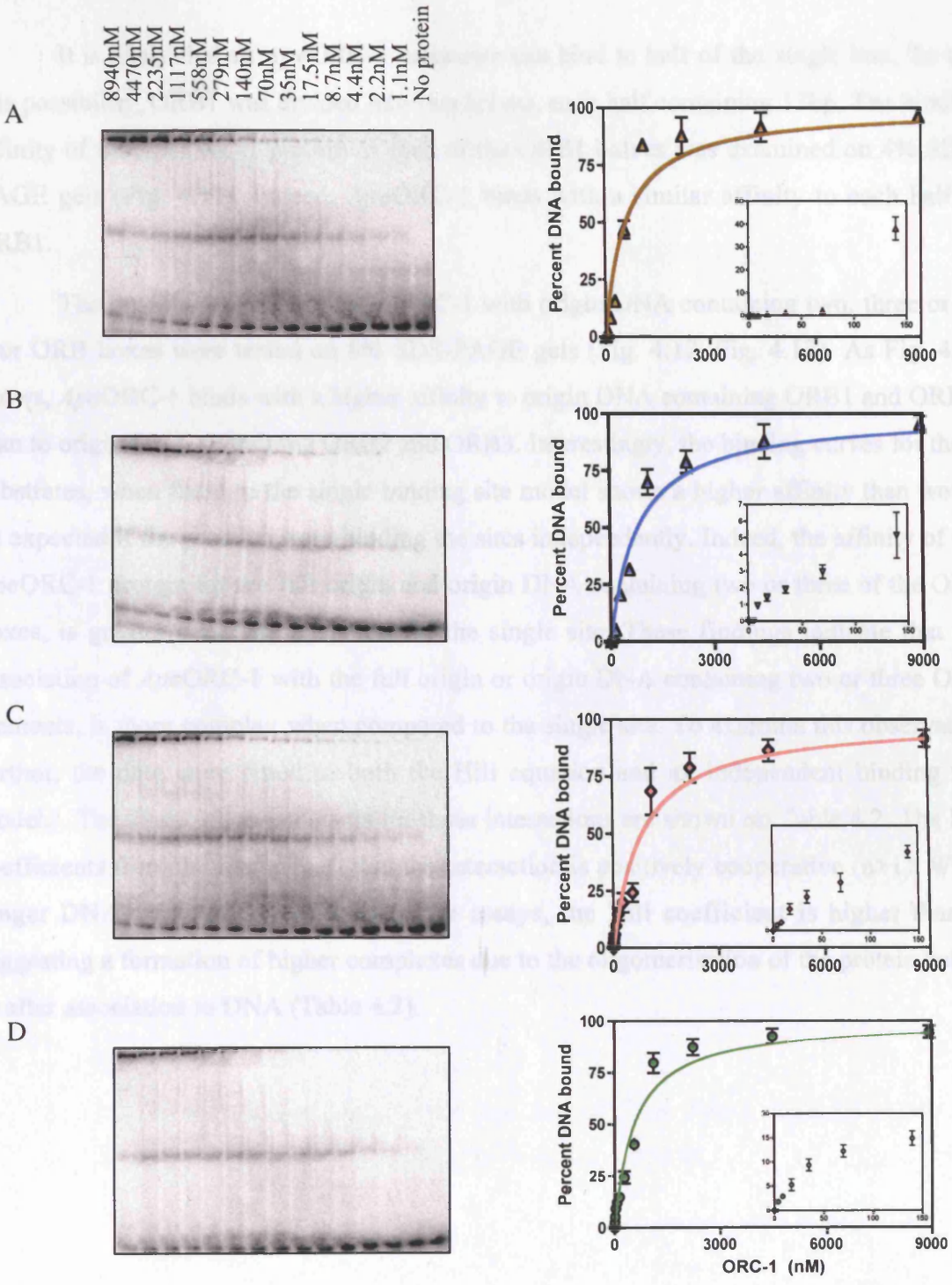


Fig. 4.10 *Ape*ORC-1 binding to single origin recognition boxes. (A) ORB1; (B) ORB2; (C) ORB3; (D) ORB4. The insets of the graphs represent the binding affinity of the protein at concentrations up to 150nM. The data is fitted to the independent binding site model.

It is likely that an *Ape*ORC-1 monomer can bind to half of the single box. To test this possibility, ORB1 was divided into two halves, each half containing 13bp. The binding affinity of the *Ape*ORC-1 protein to each of the ORB1 halves was examined on 4% SDS-PAGE gels (Fig. 4.11). Indeed, *Ape*ORC-1 binds with a similar affinity to each half of ORB1.

The binding properties of *Ape*ORC-1 with origin DNA containing two, three or all four ORB boxes were tested on 8% SDS-PAGE gels (Fig. 4.12, Fig. 4.13). As Fig. 4.14 shows, *Ape*ORC-1 binds with a higher affinity to origin DNA containing ORB1 and ORB4, than to origin DNA containing ORB2 and ORB3. Interestingly, the binding curves for these substrates, when fitted to the single binding site model shows a higher affinity than would be expected if the proteins were binding the sites independently. Indeed, the affinity of the *Ape*ORC-1 protein for the full origin and origin DNA containing two or three of the ORB boxes, is greater than that observed for the single site. These findings indicate that the association of *Ape*ORC-1 with the full origin or origin DNA containing two or three ORB elements, is more complex when compared to the single site. To examine this observation further, the data were fitted to both the Hill equation and an independent binding site model.). The dissociation constants for these interactions are shown on Table 4.2. The Hill coefficients from the fits suggest that the interaction is positively cooperative ($n>1$). When longer DNA substrates were used in the assays, the Hill coefficient is higher than 2, suggesting a formation of higher complexes due to the oligomerisation of the protein before or after association to DNA (Table 4.2).

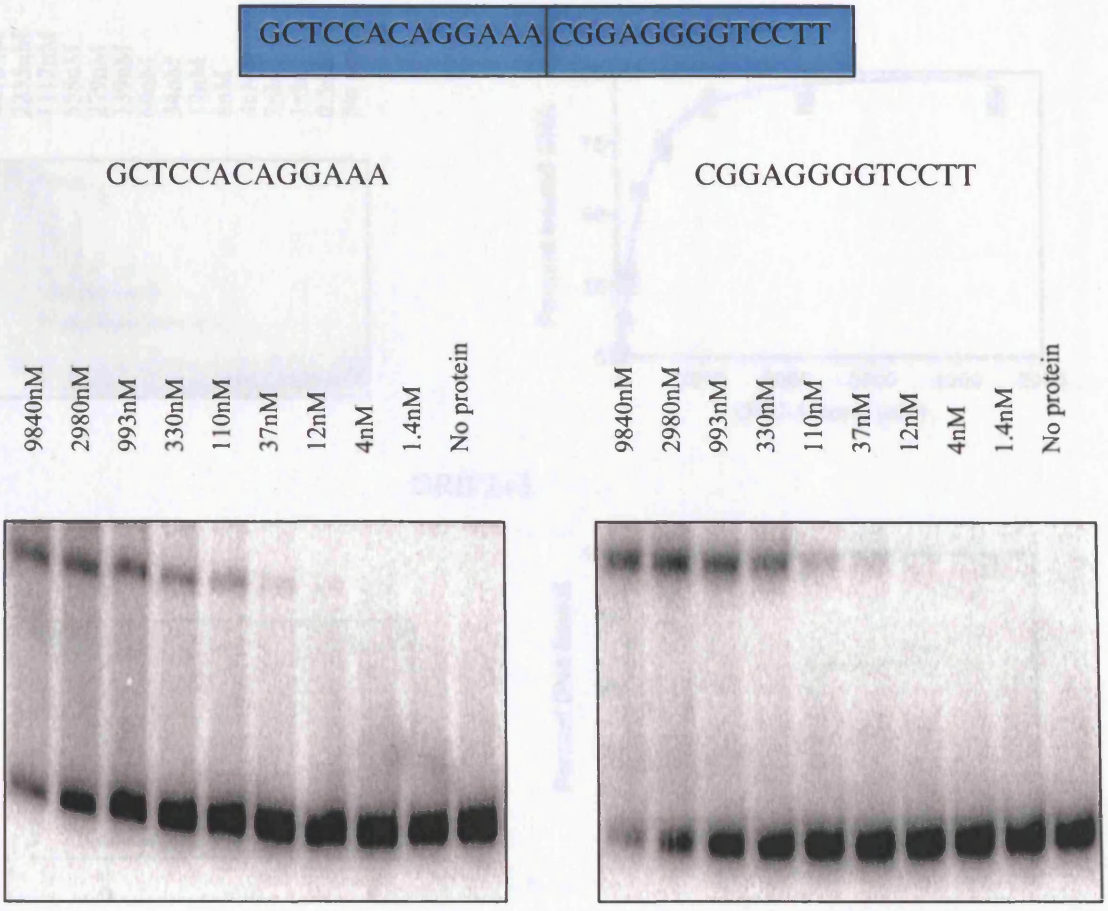
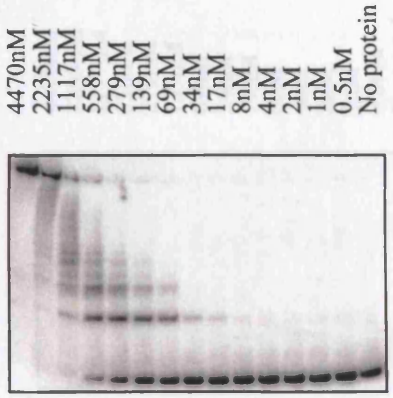
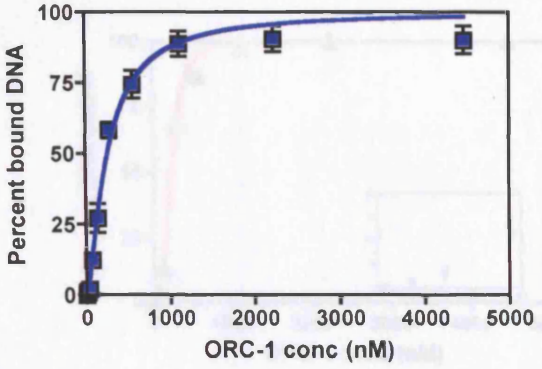


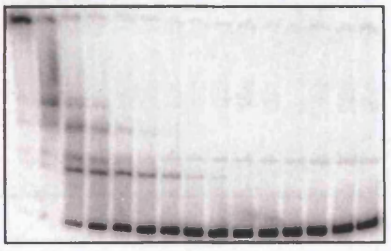
Fig. 4.11 *Ape*ORC-1 binding to the two halves of an origin recognition box (ORB 1). *Ape*ORC-1 protein can bind both halves of ORB1.

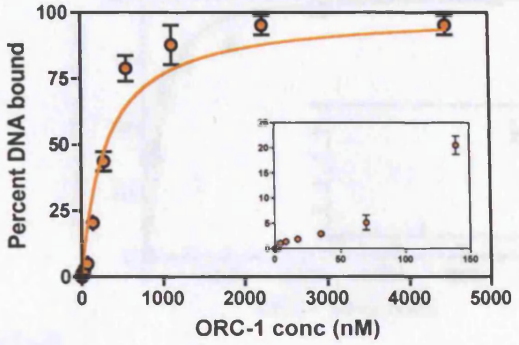
ORB 1+2



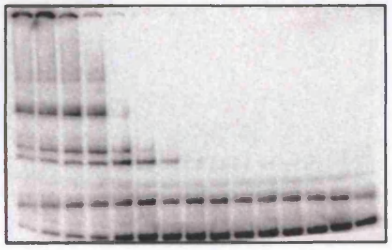


ORB 2+3





ORB 3+4



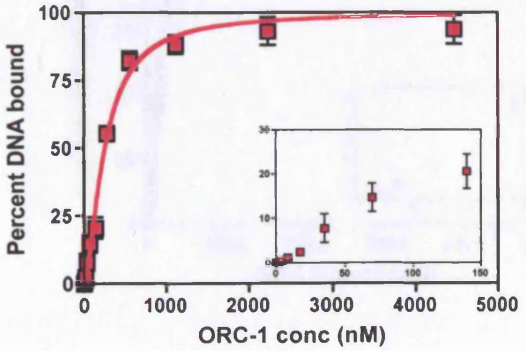
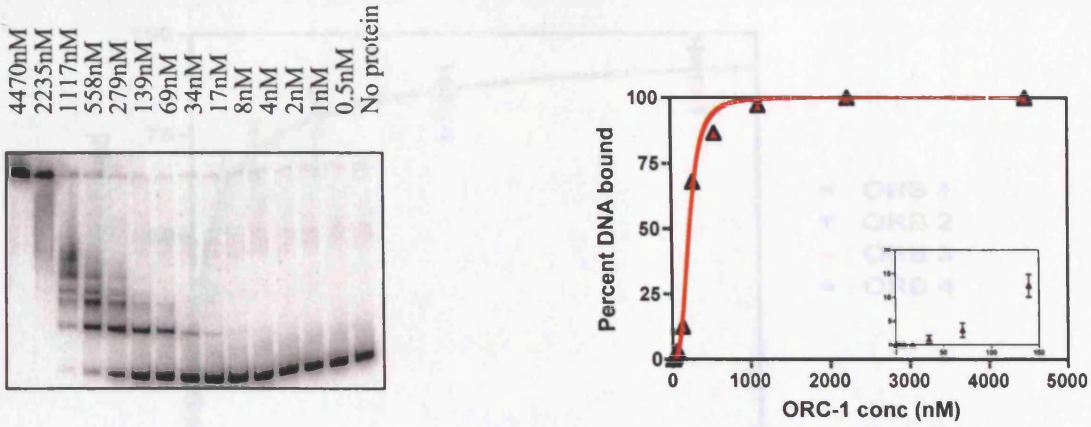
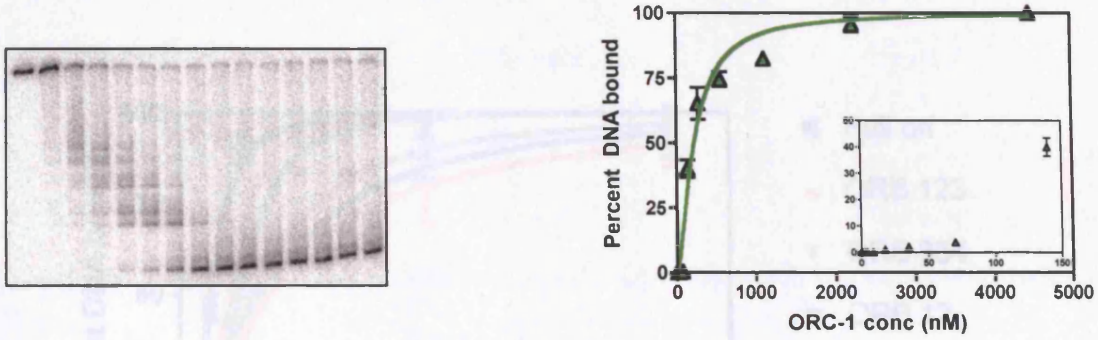


Fig. 4.12 *Ape*ORC-1 binding to different origin recognition boxes. The insets in the graphs are enlarged views of the data at protein concentration <150nM. `The points on the plot are the averaged values from three independent binding experiments with the standard deviation indicated by the error bars.

ORB 1+2+3



ORB 2+3+4



ORB 1+2+3+4

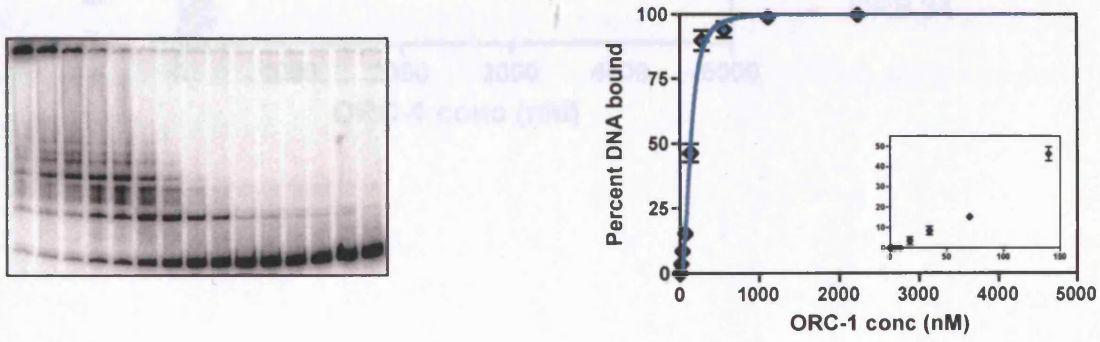
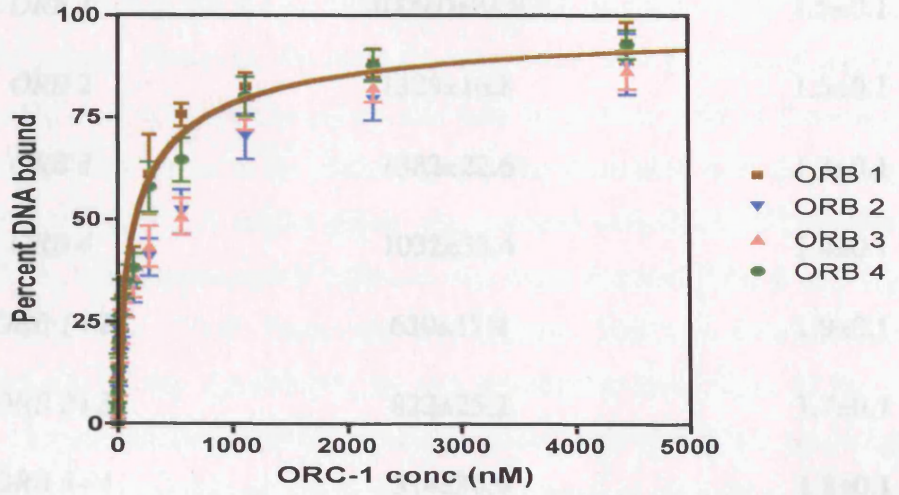


Fig. 4.13 *Ape*ORC-1 binding to modified origin sequences fitted to the Hill equation. The insets in the graphs are enlarged views of the data at protein concentration <150nM. The points on the plot are the averaged values from three independent binding experiments with the standard deviation indicated by the error bars.

A



B

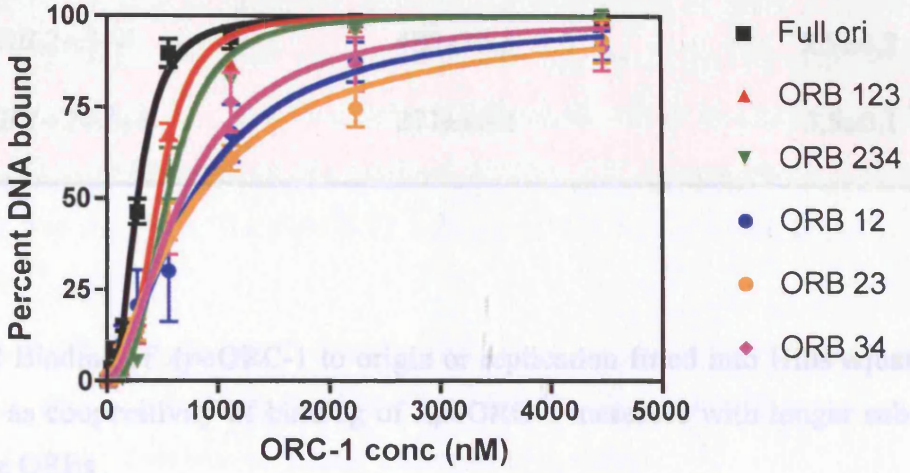


Fig. 4.14 Graphic representation of *Ape*ORC-1 binding to the origin of replication. (A) Binding to single origin recognition boxes; (B) Binding to modified origin sequences.

DNA substrate	K_D *(nM)*	*N*
ORB 1	1009.0±21.9	1.5±0.1
ORB 2	1329±16.8	1.5±0.1
ORB 3	1382±22.6	1.3±0.1
ORB 4	1032±33.4	1.4±0.1
ORB 1+2	630±17.4	1.9±0.1
ORB 2+3	822±25.2	1.7±0.1
ORB 3+4	514±36.6	1.8±0.1
ORB 1+2+3	493±34.1	2.2±0.2
ORB 2+3+4	488±27.3	2.9±0.2
ORB 1+2+3+4	271±16.2	3.5±0.1

Table 4.2 Binding of *Ape*ORC-1 to origin or replication fitted into Hills equation. It appeared as cooperativity of binding of *Ape*ORC-1 increases with longer substrates with more ORBs .

4.4.4 ATPase activity of ApeORC-1

As a member of the AAA+ family of ATPases, *Ape*ORC-1 was expected to possess an ATPase activity. However, the recombinant protein, purified from *E.coli* was found to be deficient of such activity. The crystal structure of *A.pernix* ORC-2 showed the presence of an ADP molecule bound to the protein, despite the absence of ADP during crystallisation (Singleton *et al.*, 2004). A similar result was obtained with ORC/CDC6 homologue from *P.aerophilum*, which presumably indicates that these proteins form a very tight complex with ADP (Liu *et al.*, 2000; Grabowski and Kelman, 2001). In order to investigate the ATPase activity of the *Ape*ORC-1, an apo (nucleotide-free) form of the protein was prepared. The purified recombinant protein was denatured by dialysis against guanidinium chloride containing activated charcoal, and renatured as described in Chapter 2. The ATPase activity of the apo protein was tested. The reaction mixture was incubated at 75°C. To determine the role of thermally induced autohydrolysis of ATP, a reaction mixture without the protein was used as a control. The release of [γ-^{32}P] orthophosphate was measured by a thin layer chromatography (TLC), as described in Chapter 2. 6μM of ATP and 3μM of protein were used per reaction. As Fig. 4.15 shows, single turnover of ATP hydrolysis was observed. The slow ATP hydrolysis may be due to the formation of protein complexes.

The ATPase activity of Domain I+II was also examined. The apo protein was prepared as above and as described in Chapter 2. The Domain I+II protein was found to retain the same ATPase activity as the wild type. (Fig.4.16)

Next, the effect of ss or ds non-specific DNA on the ATPase activity of *Ape*ORC-1 was examined. As Fig. 4.17 shows, the presence of 9μM of single-stranded DNA (25b oligonucleotide) did not affect the ATPase activity of the protein, while the presence of 9μM of double-stranded DNA of the same length inhibited the ATPase activity of the protein. This inhibition was even greater when ds origin DNA was used (26bp long ORB1 box). The presence of single-stranded origin DNA of the same size did not have any effect on the ATPase activity of the protein (Fig.4.18).

These results suggest similarity between the *Ape*ORC-1 and the eukaryotic ORC activity, where double-stranded origin DNA also inhibited the ATPase activity of the protein (Lee *et al.*, 2000). However, ssDNA did not affect the ATPase activity of the archaeal protein, but was found to stimulate the ATPase activity of the eukaryal ORC complex in a length-dependent manner (Lee *et al.*, 2000).

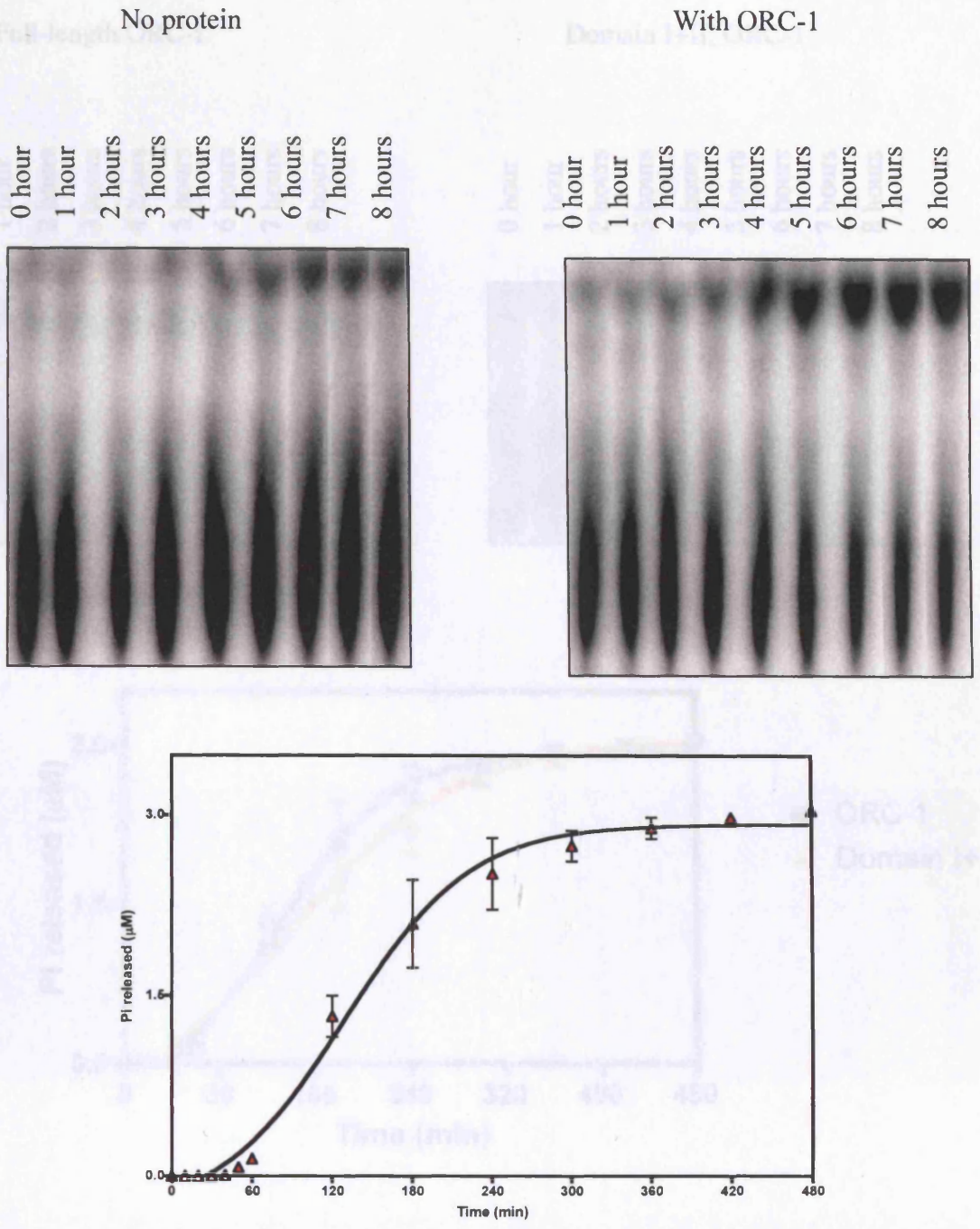


Fig. 4.15 ATPase activity of *Ape*ORC-1 protein. Reactions were performed in the presence of 6μM of ATP and 3μM of protein. The *Ape*ORC-1 protein undergoes a single turnover of ATP. Reactions without the protein were used as controls to delimit the role of thermally induced autohydrolysis of ATP, which was then substracted from the rate in the presence of enzyme. The points on the graph are the averaged values from three independent ATPase experiments with the standard deviation indicated by the error bars.

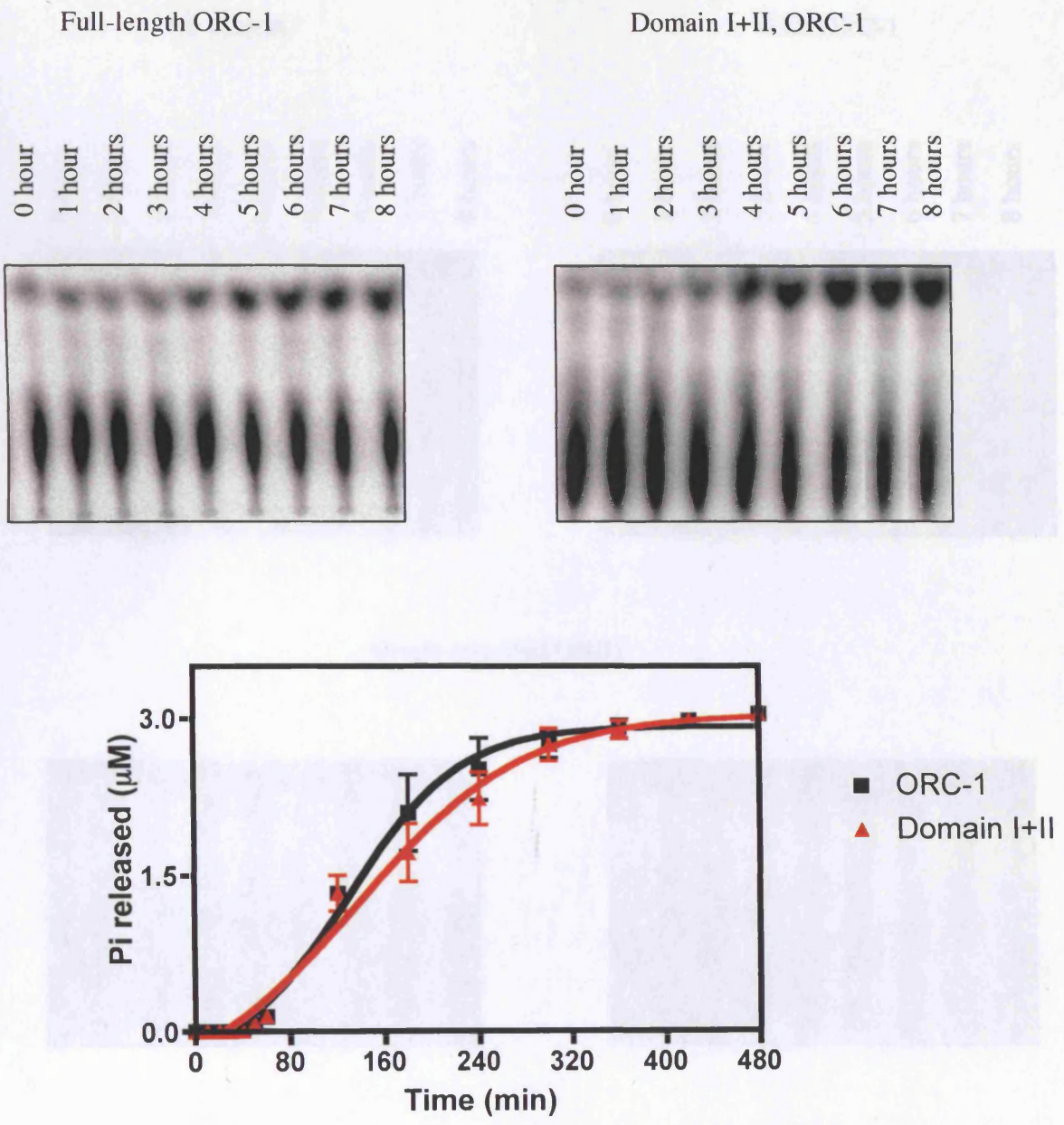


Fig. 4.16 Comparison between the ATPase activity of the full-length ORC-1 and Domain I+II of ORC-1. The assays were performed at 75°C for 8 hours. Time points were collected every hour. Domain I+II and the full-length ORC-1 protein have similar ATPase activities.

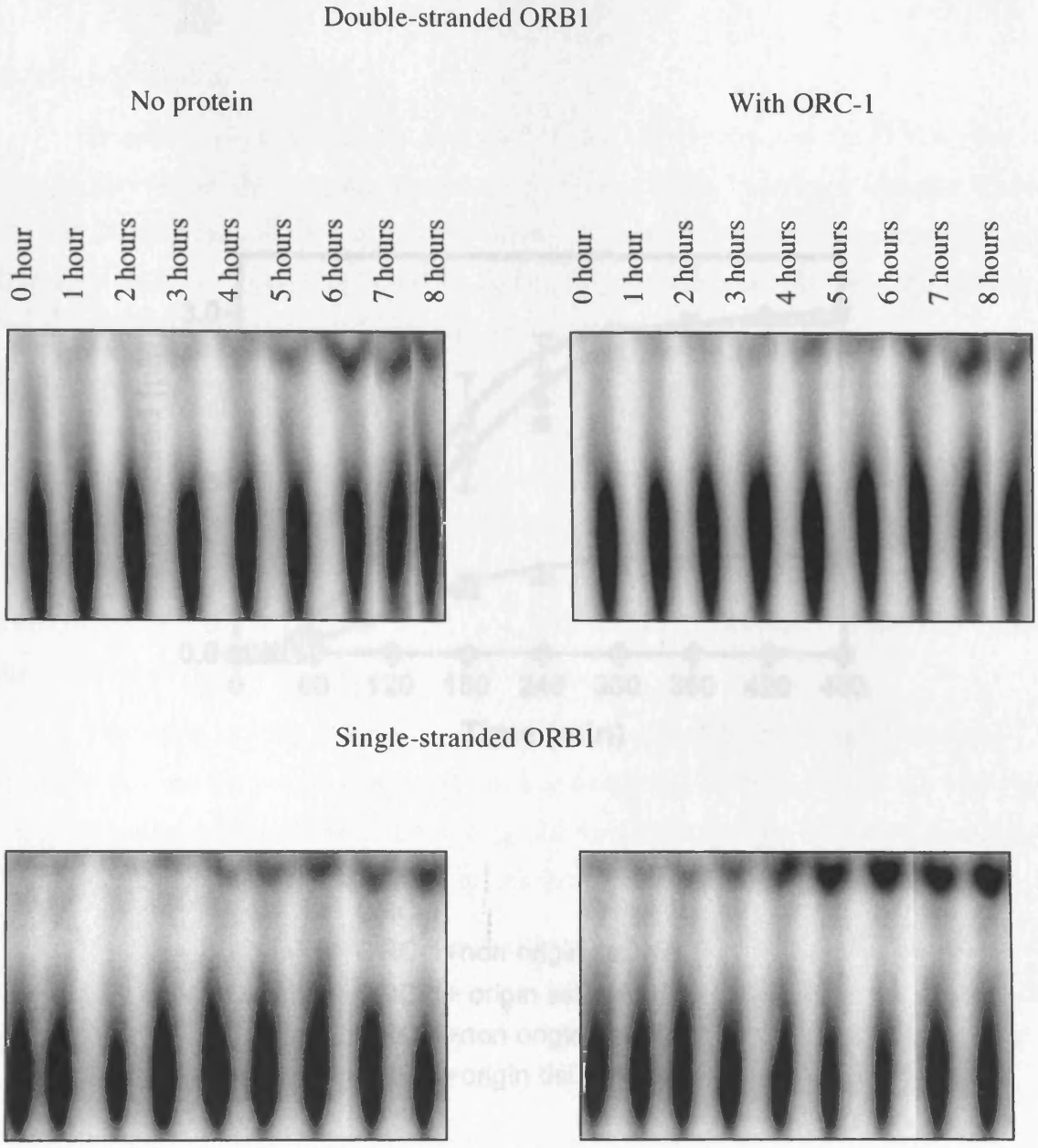


Fig. 4.17 ATPase activity of *Ape*ORC-1 in the presence of double-stranded origin DNA (top) and single-stranded origin DNA (bottom). Reaction mixtures, containing 6μM ATP, 3μM protein and 9μM DNA, were incubated at 75°C for 8 hours. Reactions without protein were used as a control and subtracted from rates with protein.

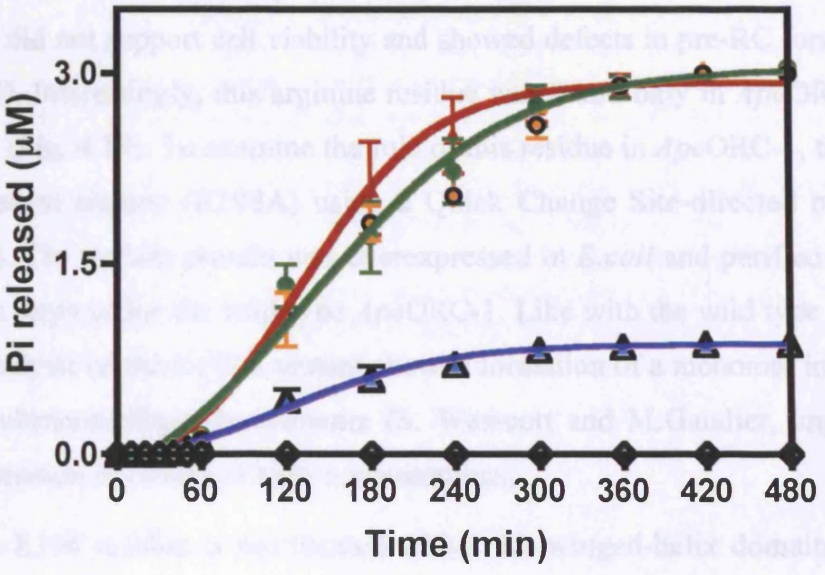


Fig. 4.18 Regulation of the ATPase activity by *Ape*ORC-1. The presence of a double-stranded origin DNA completely abolished the ATPase activity of the protein, while the presence of single-stranded DNA did not have any effect.

4.4.5 *Arginine mutant*

Sequence alignment of the archaeal ORC-1 homologues and the ORC4 proteins from *S.cerevisiae* and humans suggested the presence of a conserved arginine residue (R198). Mutations of this arginine residue in *S.cerevisiae* ORC4 eliminated ATP hydrolysis, did not support cell viability and showed defects in pre-RC formation (Bowers *et al.*, 2004). Interestingly, this arginine residue was found only in *Ape*ORC-1, but not in *Ape*ORC-2 (Fig. 4.19). To examine the role of this residue in *Ape*ORC-1, the arginine 198 was mutated to alanine (R198A) using a Quick Change Site-directed mutagenesis Kit (Chapter 2). The mutant protein was overexpressed in *E.coli* and purified using the same purification steps as for the wild-type *Ape*ORC-1. Like with the wild type *Ape*ORC-1, gel filtration analysis of the R198A mutant showed formation of a monomer in solution, while analytical ultracentrifuge experiments (S. Westcott and M.Gaudier, unpublished data) showed formation of dimers at high concentrations.

The R198 residue is not located within the winged-helix domain, therefore it is expected that the mutant protein would bind and footprint DNA as well as the wild type. Fig. 4.20 shows a DNaseI footprinting assay for the R198A and the wild type *Ape*ORC-1. This experiment provided evidence for the correct folding of the mutant protein.

Next, the R198A mutant was examined for ATPase activity by TLC. The mutant apo protein was obtained in exactly the same way as the wild type. As previously shown for the arginine mutant of *S.cerevisiae* ORC4 (Bowers *et al.*, 2004), the R198A mutant of *Ape*ORC-1 was impaired in ATP hydrolysis activity (Fig. 4.21). In *S.cerevisiae*, ATP crosslinking to the ORC4 subunit was found to be dependent on origin DNA and ATP binding by ORC1 (Klemm *et al.*, 1997), suggesting that ORC1 and ORC4 interact with the same ATP molecule. Structural analysis of other protein complexes containing multiple AAA+ proteins has demonstrated that ATP hydrolysis by one subunit is often triggered by an essential arginine residue in an adjacent subunit (Ogura *et al.*, 2004). These reactions are thought to couple conformational changes to ATP hydrolysis.

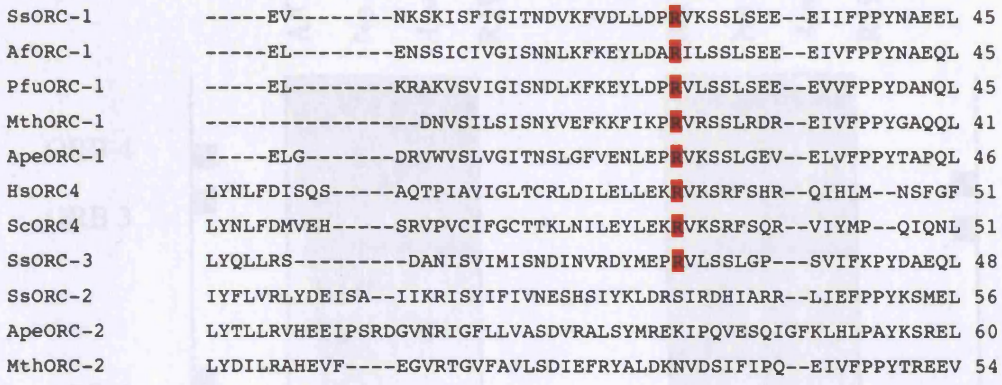


Fig. 4.19 Sequence alignment of archaeal ORC-1 and ORC-2 and eukaryotic ORC4 proteins. A conserved arginine residue (in red), which eliminates the ATP hydrolysis in *S.cerevisiae* ORC4 (Bowers *et al.*, 2004), was found within the archaeal ORC-1, but not ORC-2. Archaeal species: *Ape*, *Aeropyrum pernix; Ss*, *Solfolobus solfataricus*; *Pf*, *Pyrococcus furiosus*; *Af*, *Archaeoglobus fulgidus*; *Mth*, *Methanobacterium thermoautotrophicum.* Eukaryotic species: *Hs, Homo sapiens; Sc, Saccharomyces cervisiae.*

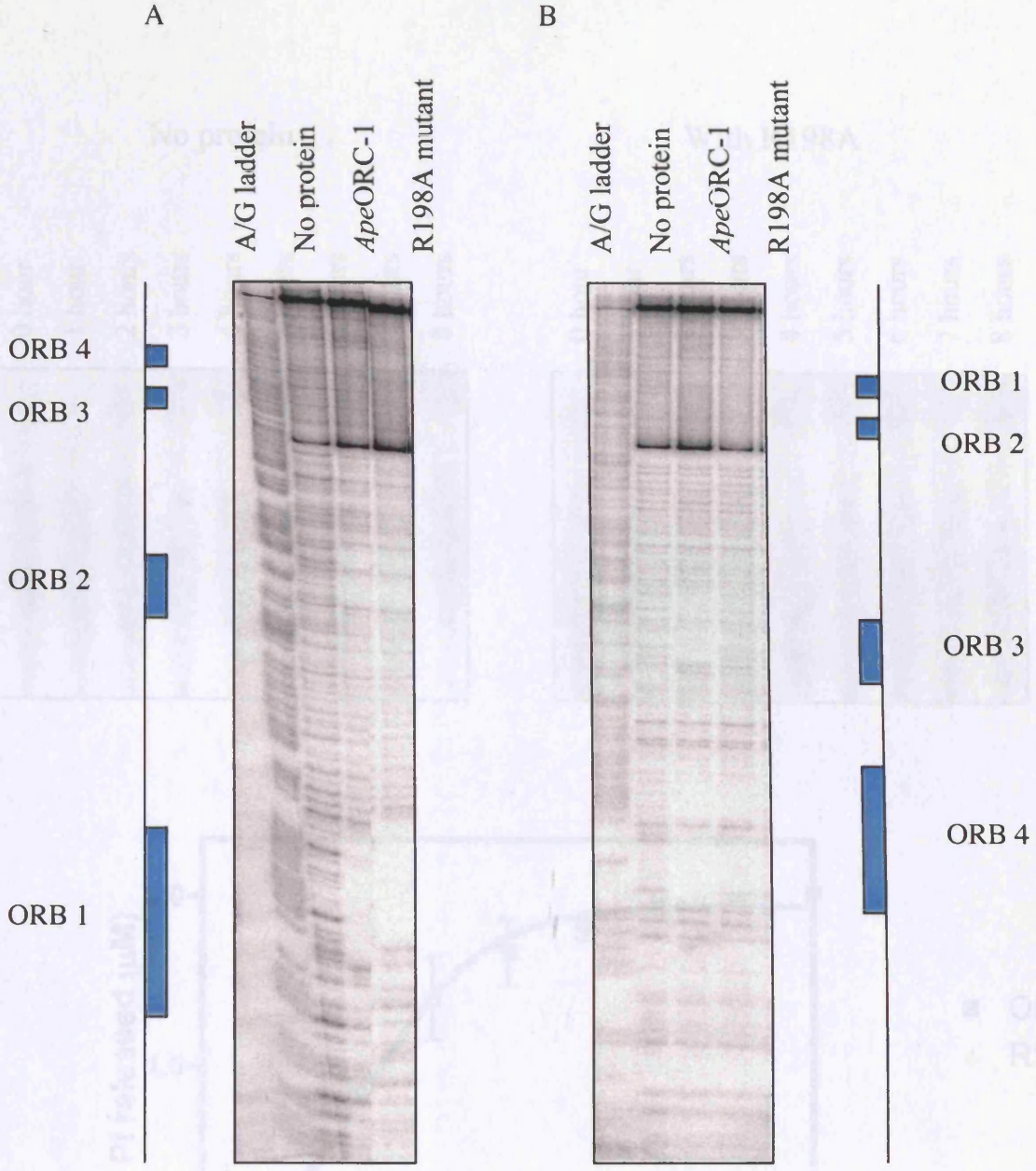


Fig. 4.20 DNase I footprinting reaction of *Ape*ORC-1 and the R198A mutant on the *Ori1* region. Each strand of the DNA was labelled separately and the protection patterns for each strand are shown. The strand with ORB1 at its 5' end is denoted as the top strand. (A) Shows footprinting on the top strand; (B) Footprint on the bottom strand. The four conserved Origin Recognition boxes (ORB) are shown as blue boxes alongside the radiograph. 4.5μM of wild type or mutant protein were used per reaction.

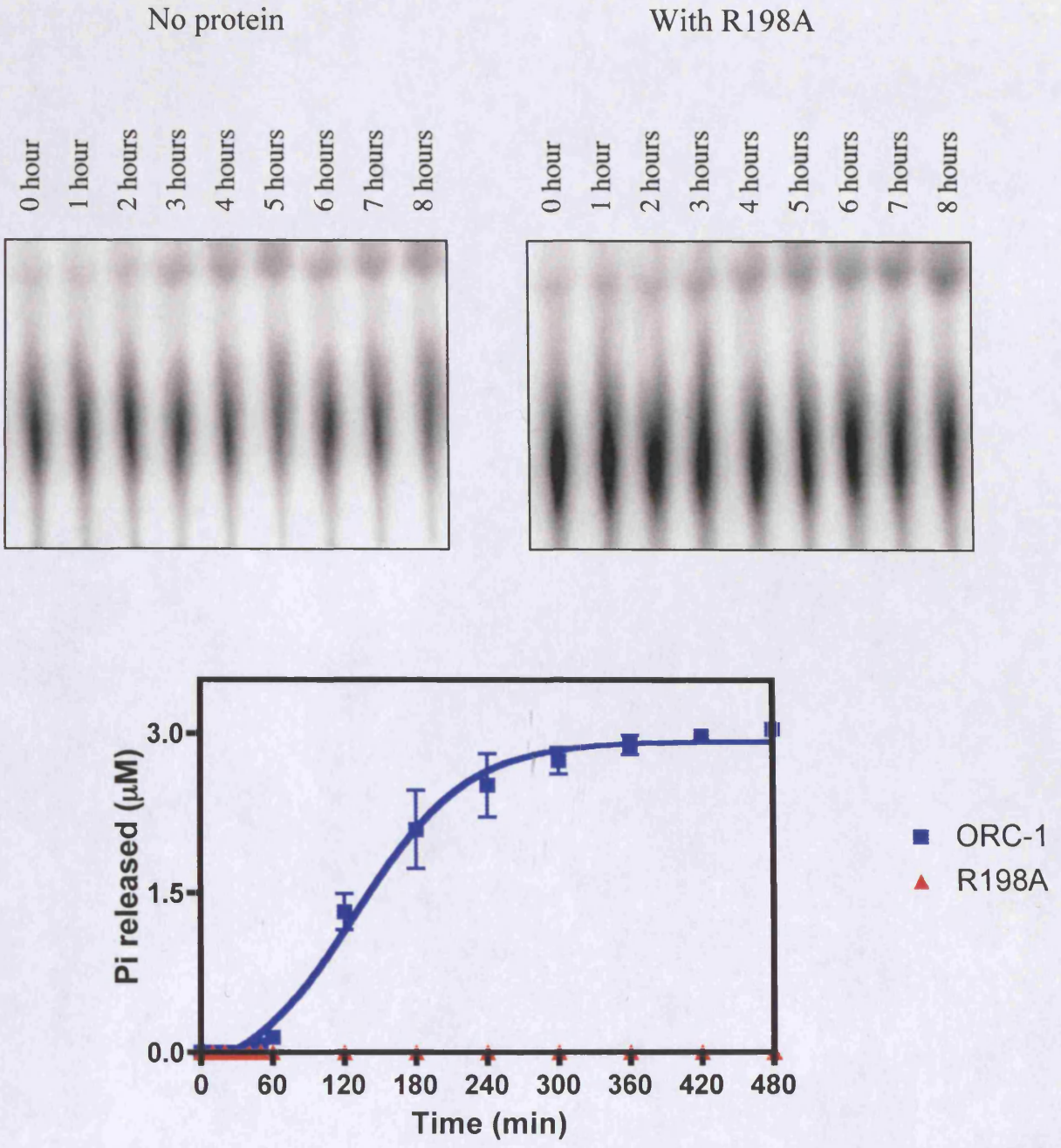


Fig. 4.21 ATPase assay for R198A mutant of *Ape*ORC-1. Reactions were performed in a presence of 3μM protein and 6μM ATP. No ATPase activity was detected for the mutant protein.

4.5 *Ape*ORC-2 protein

4.5.1 Identification and purification

*Ape*ORC-2 protein was overproduced in *E.coli* as a soluble protein and was purified using a procedure that included a thermal treatment of the cell extracts and chromatographic steps on blue and heparin columns, as described in Chapter 2. According to its crystal structure (Singleton *et al.*, 2004), the protein consists of three structural domains. Domain I comprises residues 23-213, domain II is in two parts (residues 1-22 and residues 214-299), and domain III consists of residues 300-409. Deletion mutants consisting Domain I +II and Domain III were expressed in *E.coli* and purified, using the same chromatographic steps as for the deletion mutants of *Ape*ORC-1 described above.

To assess the oligomeric state of *Ape*ORC-2 protein and its deletion mutants, gel filtration analyses were carried out using a S75 (16/60) gel filtration column. Molecular masses of 46kDa, 30kDa and 16kDa, were calculated for the recombinant full-length, Domain I+II and Domain III, respectively. Therefore, it was concluded that *Ape*ORC-2 is a monomer in solution (Fig. 4.22). Unlike the *Ape*ORC-1, analytical ultracentrifugation experiments confirmed the monomeric state of the protein (M.Gaudier and S. Westcott, unpublished data).

Sequence alignment of the archaeal ORC-2 proteins and the ORC-4 proteins from *S.cerevisiae* and humans (Fig. 4.19) show that the conserved arginine residue found in *Ape*ORC-1 is not present in the archaeal ORC-2 proteins.

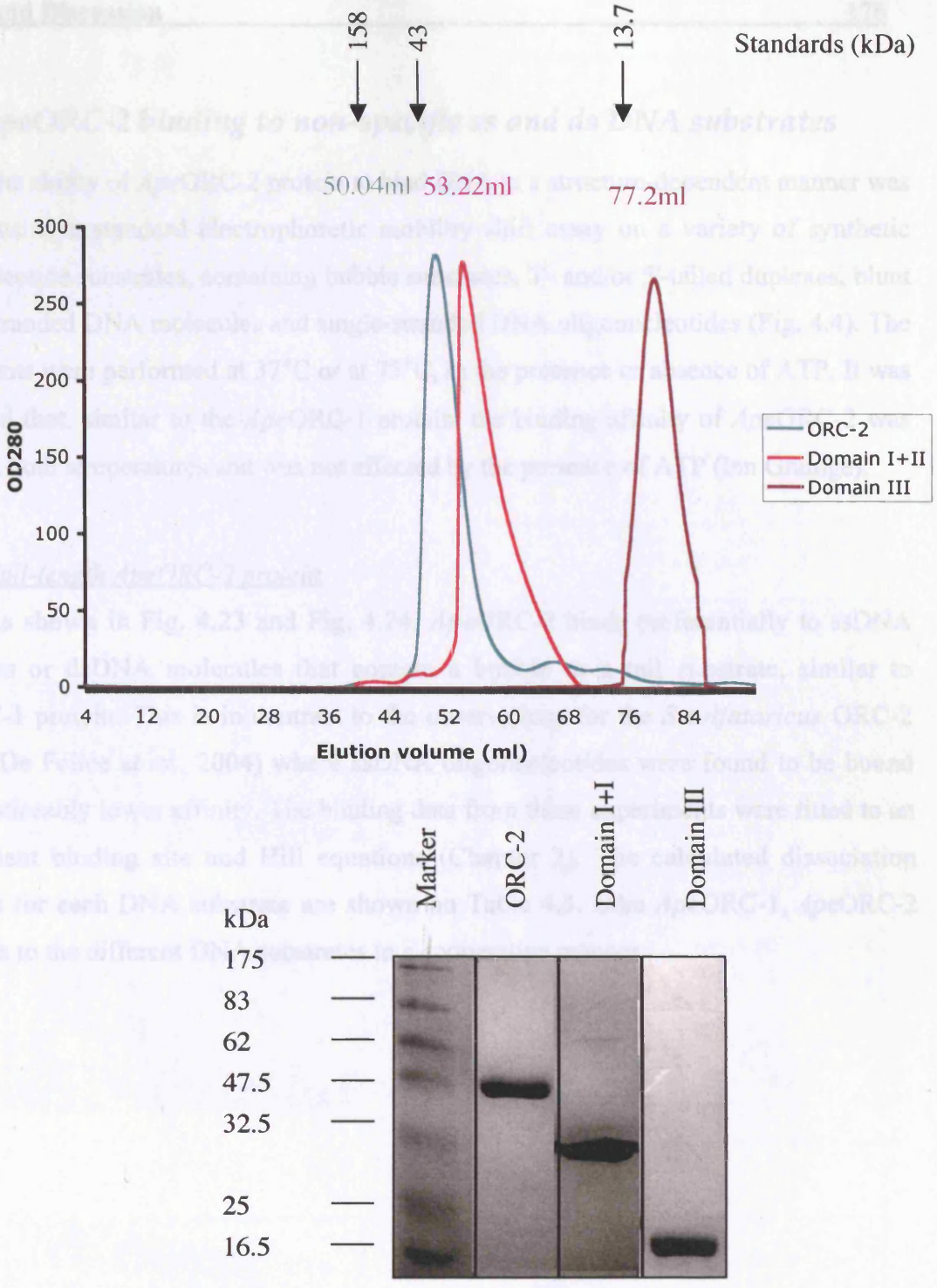


Fig.4.22 Gel filtration analysis of *Ape*ORC-2 protein. (A) Elution profile of the purified proteins from a Superdex S75 (16/60) gel filtration column. The arrows indicate the position of the protein markers used to calibrate the column (Aldolase, Ovalbumin and RNaseA). (B) Coomassie blue-stained SDS-PAGE of the full-length, Domain I+II and Domain III proteins after the GF purification procedure.

4.5.2. *ApeORC-2 binding to non-specific ss and ds DNA substrates*

The ability of *Ape*ORC-2 protein to bind DNA in a structure-dependent manner was assayed using a standard electrophoretic mobility shift assay on a variety of synthetic oligonucleotide substrates, containing bubble substrates, 3'- and/or 5'-tailed duplexes, blunt double-stranded DNA molecules and single-stranded DNA oligonucleotides (Fig. 4.4). The experiments were performed at 37°C or at 75°C, in the presence or absence of ATP. It was concluded that, similar to the *Ape*ORC-1 protein, the binding affinity of *Ape*ORC-2 was similar at both temperatures and was not affected by the presence of ATP (Ian Grainge).

4.5.2.1 Full-length ApeORC-2 protein

As shown in Fig. 4.23 and Fig. 4.24, *Ape*ORC-2 binds preferentially to ssDNA molecules or dsDNA molecules that contain a bubble or a tail substrate, similar to *Ape*ORC-1 protein. This is in contrast to the observations for the *S.solfataricus* ORC-2 protein (De Felice *et al.*, 2004) where ssDNA oligonucleotides were found to be bound with a noticeably lower affinity. The binding data from these experiments were fitted to an independent binding site and Hill equations (Chapter 2). The calculated dissociation constants for each DNA substrate are shown on Table 4.3. Like *Ape*ORC-1, *Ape*ORC-2 also binds to the different DNA substrates in a cooperative manner.

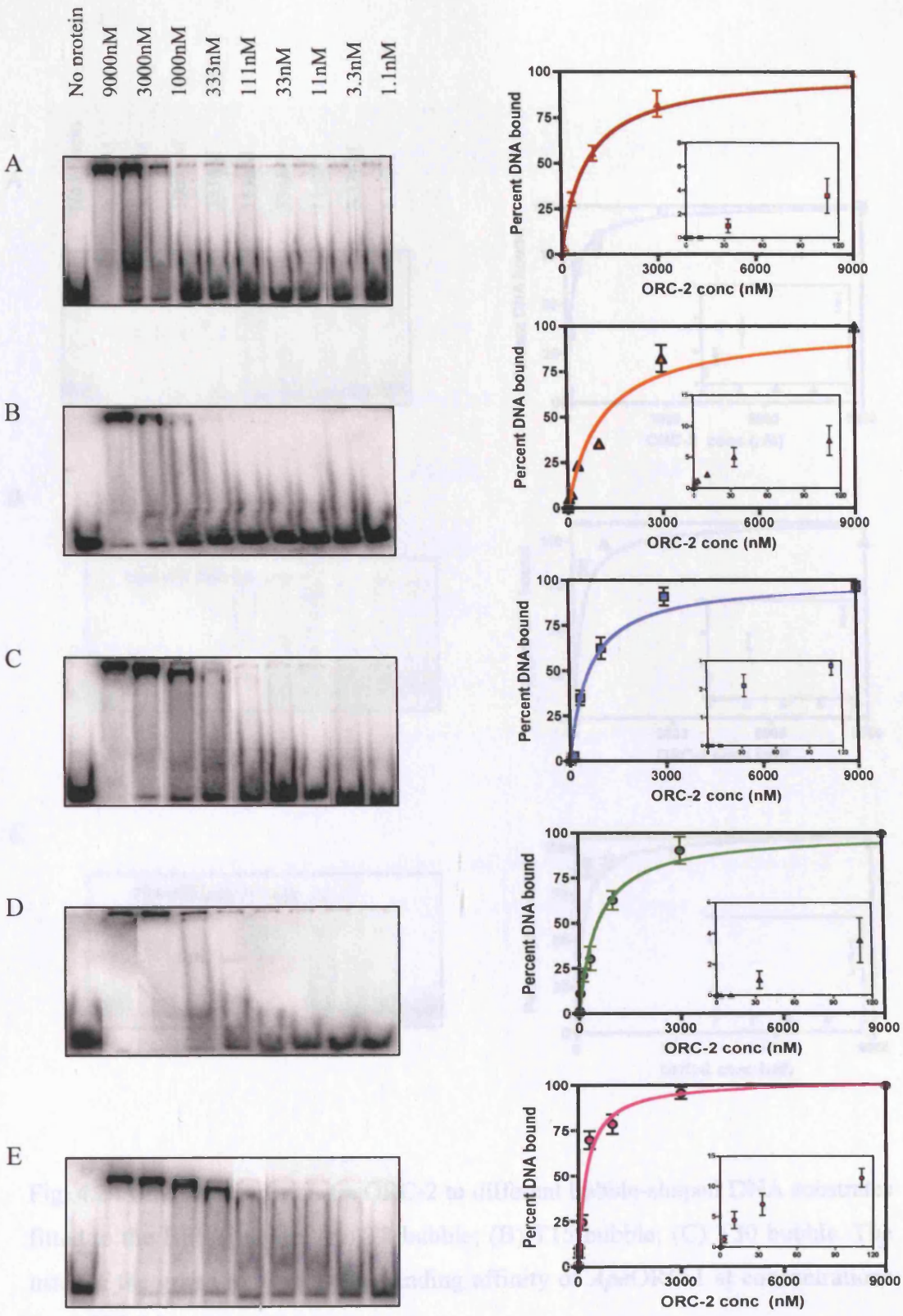


Fig. 4.23 Binding assays of *Ape*ORC-2 to different non-origin DNA substrates. (A) Single-stranded; (B) Blunt duplex; (C) 3'-tailed; (D) 5'-tailed; (E) Flayed duplex. The inset of the graph represents the binding affinity of *Ape*ORC-1 at concentrations up to 120nM.

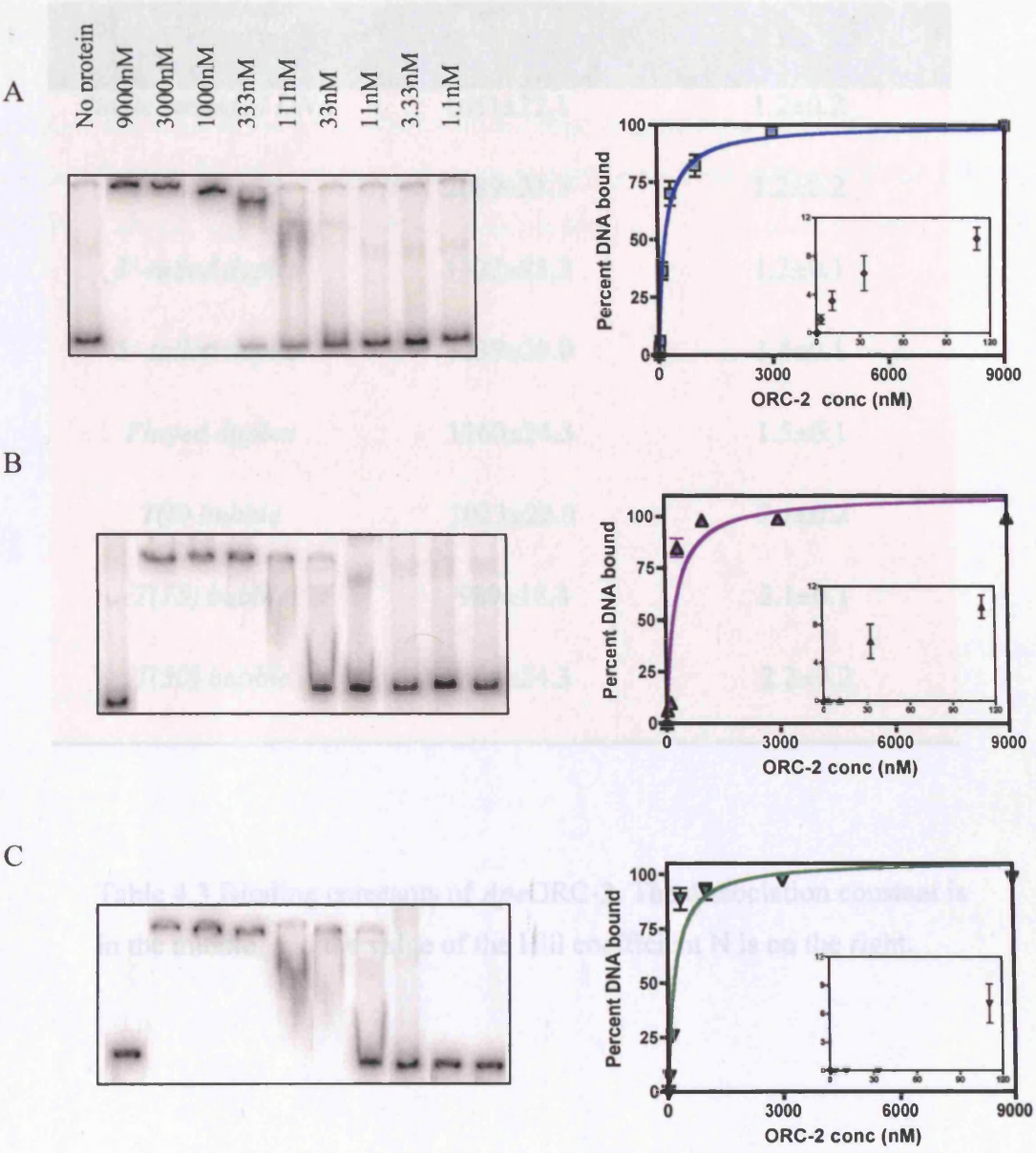


Fig. 4.24 Binding assays of *Ape*ORC-2 to different bubble-shaped DNA substrates fitted to the Hill equation. (A) T8 bubble; (B) T15 bubble; (C) T30 bubble. The inset of the graph represents the binding affinity of *Ape*ORC-1 at concentrations up to 120nM.

DNA substrate	K_D *(nM)*	N
Single-stranded DNA	1651±22.1	1.2±0.2
Blunt duplex	2089±31.7	1.2±0.2
3'-tailed duplex	1322±23.2	1.2±0.1
5'-tailed duplex	1239±23.0	1.4±0.1
Flayed duplex	1160±24.3	1.5±0.1
T(8) bubble	1023±22.0	2.1±0.2
T(15) bubble	989±18.4	2.1±0.1
T(30) bubble	864±24.3	2.2±0.2

Table 4.3 Binding constants of *Ape*ORC-2. The dissociation constant is in the middle, and the value of the Hill coefficient N is on the right..

4.5.2.2 Domain I+II

As observed for the Domain I+II of the *Ape*ORC-1 protein, Domain I+II from *Ape*ORC-2 did not bind ss or ds DNA (Ian Grainge, unpublished data). The winged helix containing domain III of *Ape*ORC-2 alone was able to bind the different DNA substrates with a lower affinity (Ian Grainge, unpublished data)

4.5.3 ATPase activity of *Ape*ORC-2 protein

Like the *Ape*ORC-1 protein, the *Ape*ORC-2 protein is a member of the AAA+ family of ATPases. The protein was purified from *E.coli* in an ADP-bound form. To assess the ATPase activity of the protein, an apo form was made as described for the *Ape*ORC-1 protein. The ATPase reactions were performed at 75°C, a reaction without a protein was used as a control. The release of [γ-^{32}P] orthophosphate was measured by thin layer chromatography (TLC), as described in Chapter 2. 6μM of ATP and 3μM of *Ape*ORC-2 protein were used per reaction. The ATPase activity of the protein was measured every 30 seconds (Fig. 4.25). Like for the *Ape*ORC-1, the *Ape*ORC-2 was found to undergo a single turnover of ATP hydrolysis and then suffered from extreme product inhibition due to the tight binding of ADP. Although the reason why ORC-2 binds ADP so tightly compared with other AAA+ family proteins is not clear, it is likely that there is some biological significance to this observation. One implication of this observation is that ATP turnover controls the activity of the ORC-2 protein and that some additional factors are required to facilitate the release of ADP from the protein, presumably as a part of the processes controlling origin firing in archaea.

The effect of ss and ds non-specific DNA on the ATPase activity of the protein was also examined. As Fig. 4.26 shows, the presence of ds DNA inhibited the ATPase activity of the protein, while the ssDNA did not have any effect.

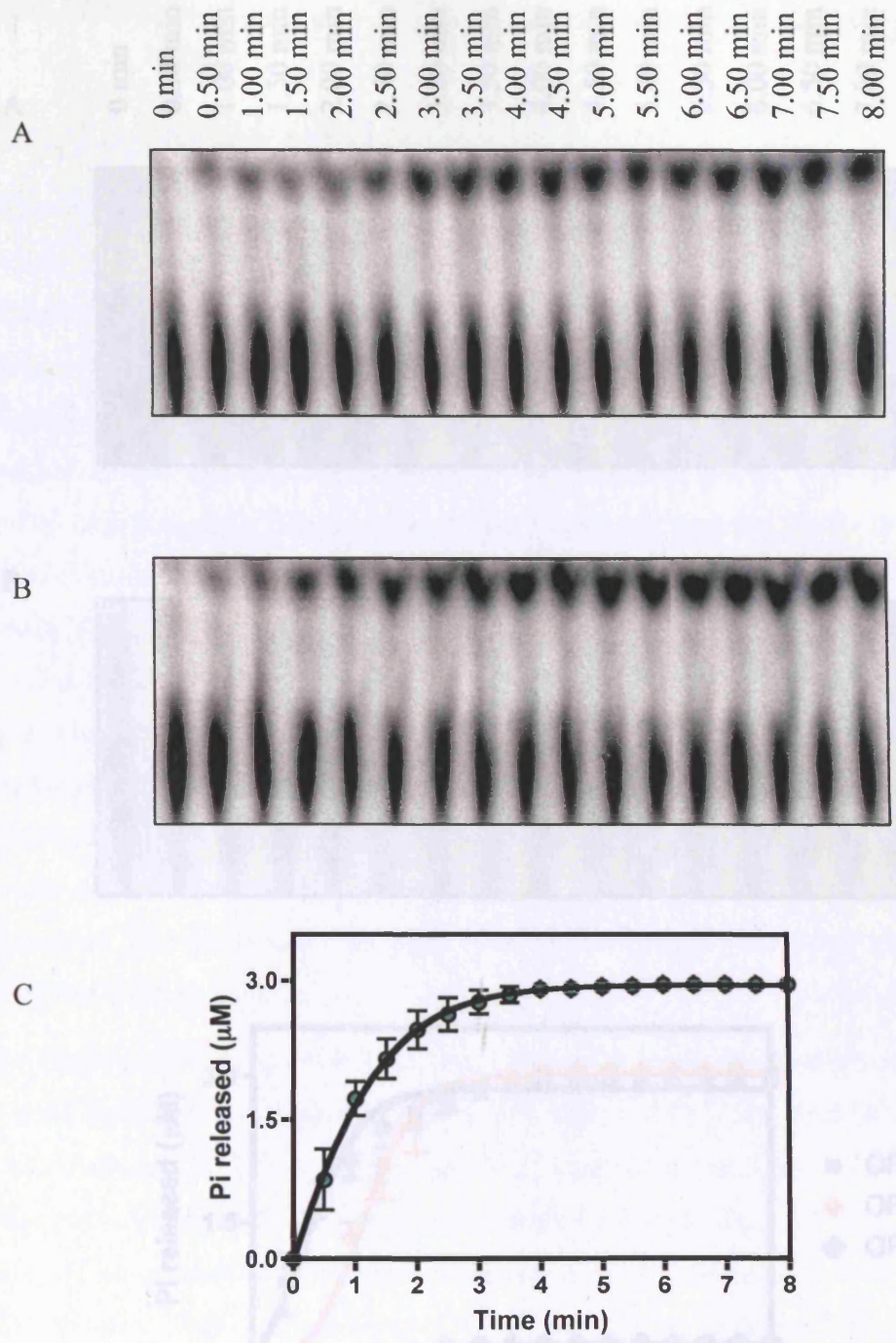


Fig. 4.25 ATPase assay of *Ape*ORC-2. (A) Full-length protein, (B) Domain I+II of ApeORC-2, (C) The ApeORC-1 protein undergoes a single turnover of ATP. 6μM ATP and 3μM protein were used in each reaction. The time points were collected every 30 seconds up to 8 minutes.

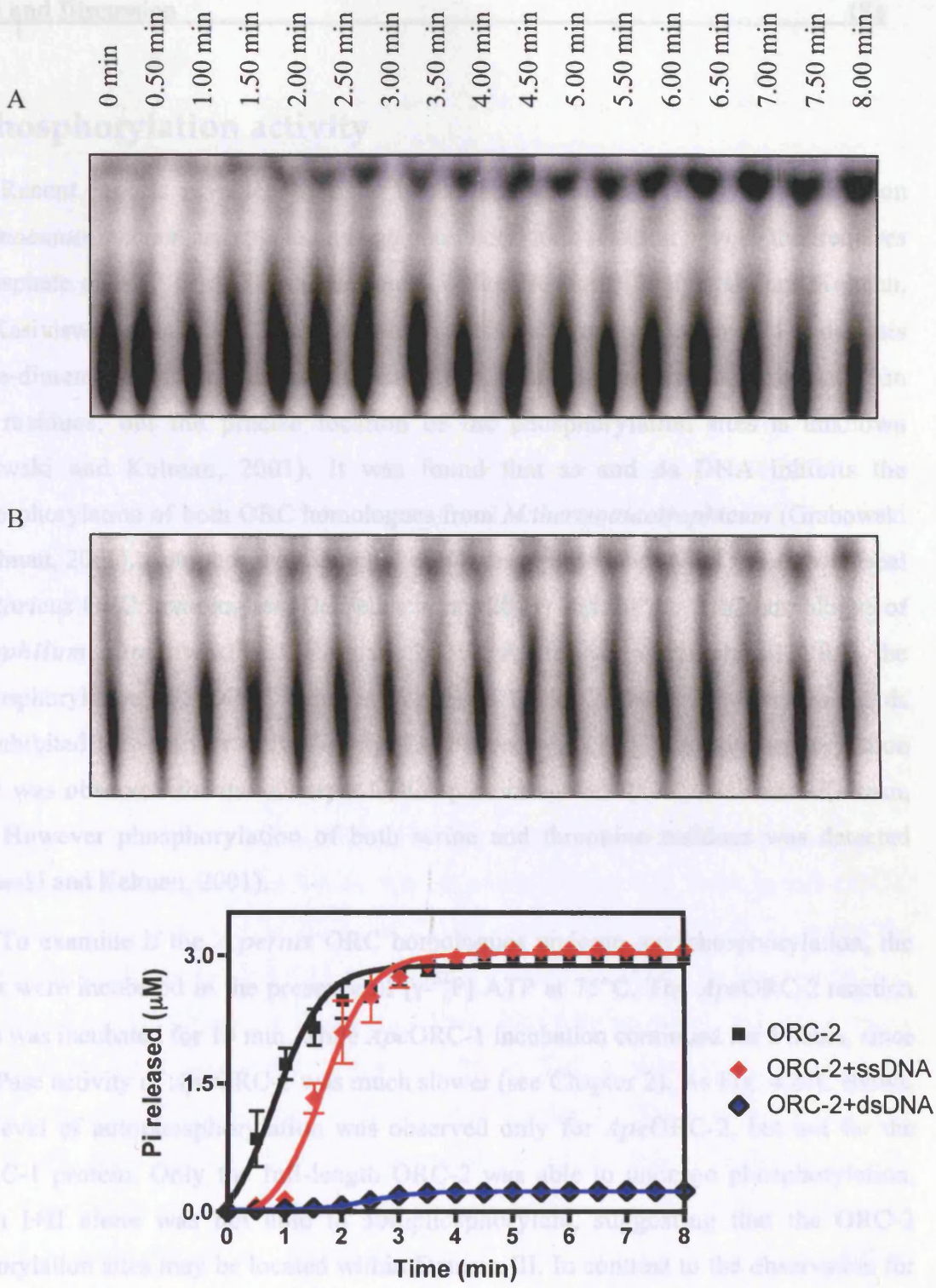


Fig. 4.26 ATPase assay for *Ape*ORC-2. (A) In the presence of ssDNA, (B) In the presence of dsDNA. (C) Regulation of ATPase activity of *Ape*ORC-2 by DNA. 6μM ATP, 3μM protein and 9μM DNA were used in each reaction.

4.6 Phosphorylation activity

Recent reports revealed that the two ORC homologues from euryarchaeaon *M.thermoautotrophicum* are able to undergo phosphorylation reaction *in vitro* that requires a γ-phosphate of ATP or dATP and an intact Walker A motif (Grabowski and Kelman, 2001; Kasiviswanathan *et al.*, 2005). A phosphoamino acid analysis with acid hydrolysis and one-dimensional electrophoresis revealed that both proteins are phosphorylated on serine residues, but the precise location of the phosphorylation sites is unknown (Grabowski and Kelman, 2001). It was found that ss and ds DNA inhibits the autophosphorylation of both ORC homologues from *M.thermoautotrophicum* (Grabowski and Kelman, 2001). Autophosphorylation of ORC was reported for one of the crenarchaeal *S.solfataricus* ORC homologues (De Felice *et al.*, 2003) and for the ORC homologue of *P.aerophilum* (Grabowski and Kelman, 2001). As for the euryarchaeal ORC, the autophosphorylation of SsoORC requires an integral Walker A motif. However, ss and ds DNA inhibited this activity only slightly (De Felice *et al.*, 2003). Autophosphorylation activity was observed for the eukaryal CDC6 protein as well (Grabowski and Kelman, 2001). However phosphorylation of both serine and threonine residues was detected (Grabowski and Kelman, 2001).

To examine if the *A.pernix* ORC homologues undergo autophosphorylation, the proteins were incubated in the presence of [γ-^{32}P] ATP at 75°C. The *Ape*ORC-2 reaction mixture was incubated for 10 min, while *Ape*ORC-1 incubation continued for 5 hours, since the ATPase activity of *Ape*ORC-1 was much slower (see Chapter 2). As Fig. 4.27C shows, a low level of autophosphorylation was observed only for *Ape*ORC-2, but not for the *Ape*ORC-1 protein. Only the full-length ORC-2 was able to undergo phosphorylation, Domain I+II alone was not able to autophosphorylate, suggesting that the ORC-2 phosphorylation sites may be located within Domain III. In contrast to the observation for *M.thermoautotrophicum* and *S.solfataricus* ORC homologues, ss and ds DNA did not inhibit the autophosphorylation ability of *Ape*ORC-2 (Fig. 4.27C). Similar results were observed for the eukaryal CDC18 protein from *S.pombe*, where ss or ds DNA did not have any effect (Grabowski and Kelman, 2001). In eukaryotes, ORC and CDC6 were found to

be phosphorylated by cyclin dependent kinases (Elsasser *et al.*, 1996; Brown *et al.*, 1997; Findeisen *et al.*, 1999). Phosphorylation of CDC6 is associated with its degradation (in yeast) or nuclear export (in mammals) (Muzi-Falconi *et al.*, 1996 a, b; Lopez-Girona *et al.*, 1998; Fujita *et al.*, 1999). Since CDC6 is essential for MCM to load onto chromatin, this has the effect of restricting MCM loading to G1/S, when CDC6 is present and active. Therefore, it was proposed that autophosphorylation in archaea might play a regulatory role during initiation as well (Grabowski and Kelman, 2001; Kelman and Kelman, 2003; Grabowki and Kelman, 2003). It was also hypothesized that ORC-MCM interaction may regulate the phosphorylation activity during assembly of the helicase around DNA at the origin (Grabowski and Kelman, 2001; Kelman and Kelman, 2003). Thus, the effect of MCM on the ORC autophosphorylation was analysed. In contrast to the observations for the *M.thermoautotrophocum* ORC homologues, where MCM stimulated the autophosphorylation of ORC-1, but inhibited the autophosphorylation of ORC-2 (Kasiviswanathan *et al.*, 2005), MCM from *A.pernix* did not affect the level of autophosphorylation of *Ape*ORC-2 (Fig. 4.27). Interestingly, it was found that ORC-2 phosphorylates full-length MCM and MCM in which the N-terminal 251 amino acids were deleted (Fig. 4.27A). The MCM phosphorylation appears to be specific for the *A.pernix* MCM only, since the *A.fulgidus* MCM was not phosphorylated by the *A.pernix* ORC-2 (Fig. 4.27B). The low level of phosphorylation made it very difficult to locate the phosphorylation sites of either MCM or ORC-2. To detect possible phosphorylation sites within *Ape*MCM, 16 amino acids long peptides (which overlap every two amino acids) from *Ape*MCM were spotted on a filter membrane. This membrane was used in the radioactive phosphorylation reaction as described in Chapter 2. After several washings and treatment with λ-phosphatase to delimit nonspecific phosphorylation, the membrane was dried and exposed overnight in a phosphoimager cassette. As Fig. 4.28 shows, a few regions of the membrane showed phosphorylation. Future work involving mutation of these residues will confirm that observation.

The observed phosphorylation of *Ape*MCM is not surprising. In eukaryotes, different subunits of the MCM2-7 complex was found to be phosphorylated by the cyclin dependent kinases. It is believed that MCM chromatin binding and helicase activity are

regulated by the phosphorylation of the MCM subunits (Musahl *et al.*, 1995; Hendrickson *et al.*, 1996; Coue *et al.*, 1996; Young *et al.*, 1997; Fujita *et al.*, 1999; Ishimi, 2000; Ishimi *et al.*, 2001). In archaea, the mechanism that control the origin firing and preventing of replication re-initiation is not known. Homologues of CDK or DDK were not found within the archaeal genomes to date, which suggest that, despite of the sequence similarity between archaeal and eukaryal replication proteins, an alternative mechanism for replication regulation may be used. It is possible that the archaeal ORC homologues play different roles during the initiation of replication. Apart from origin recognition and/or helicase loading, they may regulate the replication by phosphorylating and inactivating the archaeal MCM.

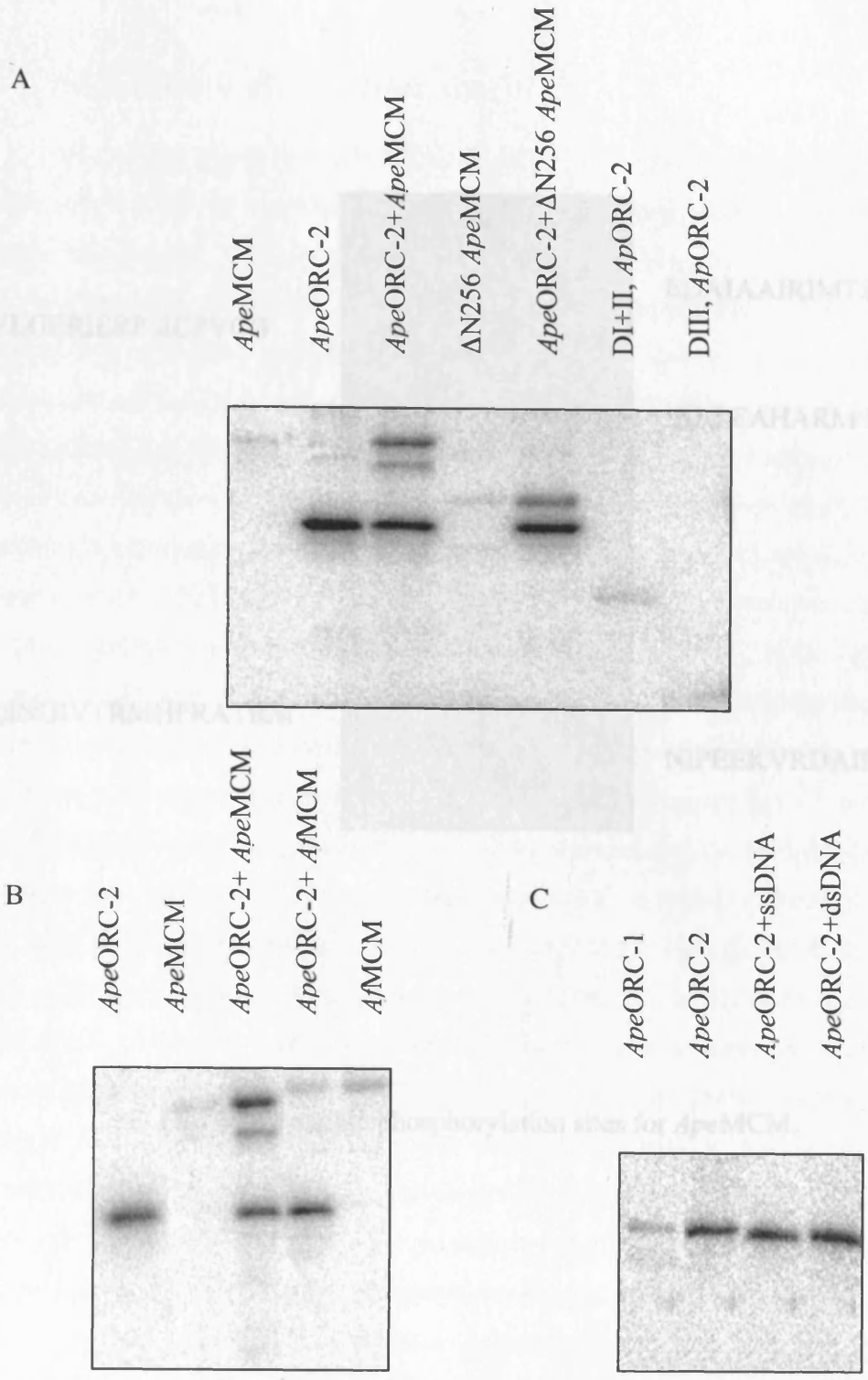


Fig. 4.27 Phosphorylation assays. (A) Within *A.pernix* proteins, (B) With MCM from *A.fulgidus*. Only the full length *Ape*ORC-2 phosphorylate itself and *Ape*MCM, but not *Af*MCM. (C) *Ape*ORC-2 did not undergo phosphorylation and ss or ds DNA did not have any effect on the phosphorylation of *Ape*ORC-2.

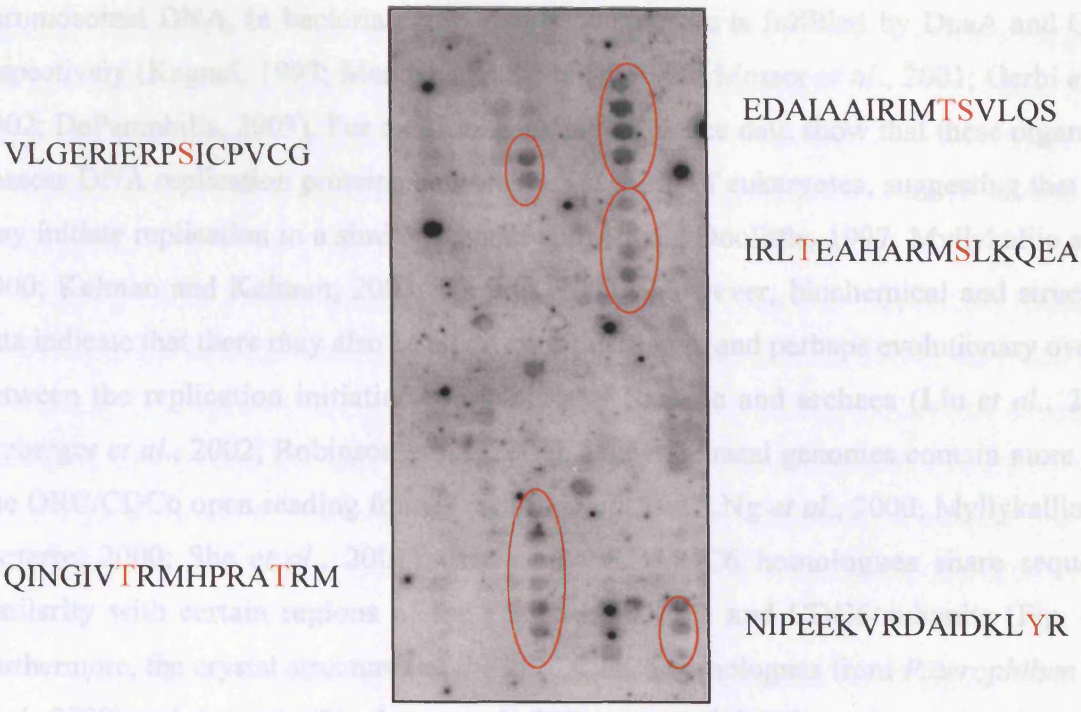


Fig. 4.28 Possible phosphorylation sites for *Ape*MCM.

4.7 Conclusions and future work

All cellular organisms use a defined set of proteins to initiate the replication of chromosomal DNA. In bacteria and eukaryotes, this role is fulfilled by DnaA and ORC, respectively (Kaguni, 1997; Mendez and Stillman, 2000; Messer *et al.*, 2001; Gerbi *et al.*, 2002; DePamphilis, 2003). For archaea, genomic sequence data show that these organisms possess DNA replication proteins homologous to those of eukaryotes, suggesting that they may initiate replication in a similar manner (Edgell and Doolittle, 1997; Myllykallio *et al.*, 2000; Kelman and Kelman, 2003; Giraldo, 2003). However, biochemical and structural data indicate that there may also be significant functional, and perhaps evolutionary overlap between the replication initiation mechanism of bacteria and archaea (Liu *et al.*, 2000; Erzberger *et al.*, 2002; Robinson *et al.*, 2004). Many archaeal genomes contain more than one ORC/CDC6 open reading frames (Smith *et al.*, 1997; Ng *et al.*, 2000; Myllykallio and Forterre, 2000; She *et al.*, 2001). Archaeal ORC/CDC6 homologues share sequence similarity with certain regions of the eukaryotic ORC1 and CDC6 subunits (Fig. 4.2). Furthermore, the crystal structures of the ORC/CDC6 homologues from *P.aerophilum* (Liu *et al.*, 2000) and *A.pernix* (Singleton *et al.*, 2004) revealed that they are composed of an N-terminal AAA+ nucleotide binding module linked to a C-terminal winged helix domain that is believed to be responsible for DNA binding activity (Liu *et al.*, 2000; Singleton *et al.*, 2004). A similar modular organisation was also observed in the three-dimensional structure of the DnaA protein from the bacterium *Aquifex aeolicus* (Erzeberger *et al.*, 2002). The function of the archaeal ORC1/CDC6 homologues remains unknown. It is possible that different ORC1/CDC6 define distinct replication origins within given species. Alternatively, it is conceivable that, although related to both eukaryotic ORC1 and CDC6, individual archaeal ORC/CDC6 factors may have a dual function as the replication initiator (by specifically binding the chromosomal replication origin) and as the DNA-helicase loader (by recruiting the MCM complex at the origin) (Kelman and Kelman, 2003). Finally, it is possible that some of the archaeal ORC1/CDC6 homologues act as regulators. However, the biochemical properties of the archaeal ORC/CDC6 homologues have been poorly investigated so far.

In this chapter, the biochemical characteristics of the two ORC/CDC6 homologues from *A.pernix* (named ORC-1 and ORC-2) were studied. Gel filtration analyses showed that the recombinant *Ape*ORC-1 and *Ape*ORC-2 form monomers in solution at low concentrations. However, more recent analytical ultracentrifugation experiments suggested that *Ape*ORC-1 (but not *Ape*ORC-2) dimerises at higher concentrations (M. Gaudier and S. Westcott). Dimerisation of archaeal ORC homologue was observed with the ORC-2 from *M.thermoautotrophicum* as well (Shin *et al.*, 2003). The significance of this observation is not clear yet.

Both ORC homologues show a high degree of sequence similarity. While the crystallographic structure of *Ape*ORC-2 is available (Singleton *et al.*, 2004), the structure of *Ape*ORC-1 is not, but it is expected to be similar to the ORC-2. Both proteins contain Walker A and Walker B motifs located in the interface between domain I and domain II, and a winged-helix motif located at the C-terminus. The winged-helix domain has been proposed to interact with DNA (Liu *et al.*, 2000; Singleton *et al.*, 2004; Robinson *et al.*, 2004).

To better understand the physical basis of the initiation of DNA replication in *A.pernix*, it is important to define the molecular interactions that occur at the replication origin. Such study has only recently become possible through the identification of an origin of replication in *A.pernix* (Ian Grainge). In the meantime, non-specific DNA of different shapes was used to examine the binding affinity of the proteins. An electrophoretic mobility band shift assays were carried out using a variety of protein concentrations. Both proteins were found to bind single-stranded and double-stranded DNA unlike *Mth*ORC-1 protein, which was found to bind double-stranded, but not single-stranded DNA (Capaldi and Berger, 2004; Kasiviswanathan *et al.*, 2005). The binding curves were fitted to both an independent binding site equation and the Hill equation. Although only ORC-1 was found to form dimers, both ORC proteins of *A.pernix* were found to bind cooperatively to all DNA substrates used (Hill coefficients higher than 1). The DNA binding affinity of Domain I+II and Domain III were also investigated. As expected from the structural organisation, Domain I+II was not able to bind single-stranded or double-stranded DNA, while Domain III alone did bind the different forms of DNA although with a lower affinity.

This observation confirms that the winged-helix domain III is responsible for the DNA binding activity of the proteins.

Recently, a sequence, which may play the role of a replication origin, was detected in *A.pernix* (Ian Grainge). The sequence contained four repeats, named origin of recognition boxes, similar to the ones found in some archaeal species (Robinson *et al.*, 2004; Capaldi and Berger, 2004). The binding affinity of ORC-1 protein to each of the four 26bp-long single boxes was examined. It was found that *Ape*ORC-1 binds cooperatively to ORB1 and ORB4 with a higher affinity than to ORB2 and ORB3. When the full origin was used (all four boxes), the Hill coefficient was higher than 2, suggesting a multimerisation of the protein upon binding to DNA. Future work in the lab will examine further this possibility. Analytical ultracentrifuge experiments using a DNA sequence containing all four ORBs will show if the ORC-1 protein oligomerises further upon binding to origin DNA. Cross-linking together ORC proteins and examination of their oligomerisation state will show if ORC-1 can form higher oligomers alone or the multimerisation occurs as a result of binding to origin DNA.

Next, the ATPase activity of the two ORC homologues was examined. Analysis of the primary structure of both proteins revealed that they contain all sequence motifs found in the proteins of the AAA+ family, and in accordance with this similarity the homogenous recombinant proteins were found to be able to hydrolyse ATP. Both proteins were deficient in ATPase activity when purified from *E.coli*. The crystal structure of *Ape*ORC-2 showed the presence of ADP although the nucleotide was not added to the protein (Singleton *et al.*, 2004). To clear out the bound nucleotide, both ORC proteins were denatured in 6M guanidinium chloride and dialysed against activated charcoal before subsequent renaturation. The apo forms were found to undergo a single turnover of ATP hydrolysis and then suffer from extreme product inhibition due to the tight binding of nucleotide. Although the reason why the proteins bind nucleotide so tightly is not clear, it is likely that there is some biological significance to this observation. One implication could be that ATP turnover controls the activity of the ORC proteins and some additional factors are required to facilitate the release of nucleotide from them. To test if DNA is one of these factors, single-stranded and double-stranded origin and non-specific DNA were added to the

ATPase reaction. It was found that double-stranded (but not single-stranded) DNA inhibits the ATPase activity of both proteins. These results suggest similarity between the activities of the *A.pernix* ORC homologues and the eukaryotic ORC complex (Lee *et al.*, 2000). In *S.cerevisiae*, double-stranded origin DNA inhibits the ATPase activity of the protein, while single-stranded origin DNA was found to stimulate this activity. Therefore, it was proposed that origin unwinding triggers a transition between nucleotide-bound and apo forms (Lee *et al.*, 2000). It is possible that other proteins stimulate the release of nucleotide from the archaeal ORC homologues. Future work in the lab will investigate if the presence of MCM has any effect on the ATPase activity of both ORCs.

Another interesting similarity between eukaryotic ORC and archaeal ORC proteins is the presence of a conserved arginine residue. Such a residue was found in class I, but not class II of the archaeal ORC homologues. Mutation of this residue in *S.cerevisiae* ORC4 eliminated the ATP hydrolysis by ORC, did not support cell viability and showed defects in pre-RC formation (Bowers *et al.*, 2004). Mutation of the corresponding residue in *A.pernix* ORC-1 (R198A) was found to be defective in ATP hydrolysis. Future work in the lab will examine further the role of this arginine. It is interesting to know if the ATP binding activity of the protein is also affected by this residue and if this arginine is important for formation of the pre-replication complex in *A.pernix*.

Several studies on the ORC homologues from *M.thermoautotrophicum* (Grabowski and Kelman, 2001), *S.solfataricus* (De Felice *et al.*, 2003) and *P.aeorophylum* (Grabowski and Kelman, 2001) demonstrated that the proteins undergo autophosphorylation. Although the precise phosphorylation sites are unknown, phosphoamino acid analysis and one-dimensional electrophoresis revealed that both ORC proteins from *M.thermoautotrophicum* are phosphorylated on the serine residues. It was also demonstrated that double-stranded and single-stranded DNA inhibits the autophosphorylation reaction (Grabowski and Kelman, 2001; De Felice *et al.*, 2003). In *A.pernix* the full-length ORC-2, but not *Ape*ORC-1 was found to undergo autophosphorylation, which was not affected by the presence of single-stranded or double-stranded DNA. Although Domain I+II was found to have a similar ATPase activity as the full-length ORC-2, this deletion mutant did not autophosphorylate. This suggests that phosphorylation sites may be located within Domain

III. Moreover, ORC-2 was found to phosphorylate the full-length MCM as well as the deletion mutant ΔN256 MCM. This phosphorylation is species-specific since MCM from *A.fulgidus* was not phosphorylated by ORC-2 from *A.pernix*. This observation suggests that the MCM phosphorylation sites may be located within residues 256-699 of *A.pernix* MCM and may not be conserved between *A.fulgidus* and *A.pernix* MCM homologues. A phosphorylation of MCM by ORC-1 seems to occur in *M.thermoaoutotrophicum* as well although it was not discussed in the article (Kasiviswanathan *et al.*, 2005, Fig. 6, lane 4 of the paper). Interestingly, *Mth*MCM was found to stimulate the autophosphorylation of ORC-1 and inhibit the autophosphorylation of ORC-2 (Kasiviswanathan *et al.*, 2005). In contrast, MCM from *A.pernix* did not affect the level of autophosphorylation of *Ape*ORC-2 and did not provoke a phosphorylation of *Ape*ORC-1. The low level of phosphorylation makes the determination of phosphorylation sites of either ORC-2 or MCM extremely difficult. When a membrane containing overlapping amino acid peptides from *A.pernix* MCM was used in the phosphorylation reaction, several residues were found to be phosphorylated. Future work in the lab involving mutations of these residues will confirm that observation and will investigate further the role of MCM phosphorylation in the formation of the pre-replication complex of *A.pernix*.

CHAPTER 5

BIOCHEMICAL CHARACTERISATION OF *A.pernix* MCM HELICASE

5.1 Introduction

Almost all of the archaea for which the genome is known contain a single MCM homologue (Myllykallio and Forterre, 2000). The archaeal MCM homologues from *M.thermoautotrophicum* (Kelman, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Poplawski *et al.*, 2001; Shin *et al.*, 2003; Chen *et al.*, 2004; Kasiviswanathan *et al.*, 2005), *A.fulgidus* (Grainge *et al.*, 2003) and *S.solfataricus* (Carpentieri *et al.*, 2002; McGeoch *et al.*, 2005) have been extensively studied *in vitro*, while the *P.abyssi* protein was studied *in vivo* (Matsunaga *et al.*, 2001). The archaeal MCM protein can be divided into two functional parts. The N-terminal part is believed to participate in protein multimerisation and ssDNA binding, while the C-terminal part of the protein contains the helicase catalytic domain(s) (Chong *et al.*, 2000; Fletcher *et al.*, 2003). *In vitro* studies revealed that the archaeal MCM helicase possesses biochemical properties similar to those of the eukaryotic MCM helicase (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Tye, 2000; Poplawski *et al.*, 2001; Carpentieri *et al.*, 2002). These include a 3'-5' helicase activity which is dependent upon ATP, ssDNA binding and DNA dependent ATPase activity.

The oligomeric structure of the archaeal MCM complex is still not clear. Whereas the recombinant *M.thermoautotrophicum* MCM was reported to form dodecamers (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Poplawski *et al.*, 2001), the MCM from *S.solfataricus* (Carpentieri *et al.*, 2002) and *A.fulgidus* (Grainge *et al.*, 2003) were both shown to behave as homohexamers in solution. The crystallographic structure of the *Mth*MCM N-terminal portion (residues 2-286) revealed a dodecameric architecture, with two hexameric rings juxtaposed in a head-to-head configuration (Fletcher *et al.*, 2003). A remarkable feature of the *Mth*MCM crystal structure is the presence of a long central channel the surface of which has a considerably high positive charge. A three-dimensional reconstruction of the full sized *Mth*MCM structure by electron microscopy suggests that this central channel runs throughout the entire *Mth*MCM molecule (Pape *et al.*, 2003).

Studies on the biochemical properties of euryarchaeal and crenarchaeal MCMs demonstrated similarities and important differences (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Carpentieri *et al.*, 2002; Grainge *et al.*, 2003). The biochemical

data from euryarchaeal MCMs comes from the characterisation of *M.thermoautotrophicum* MCM (*Mth*MCM) and *A.fulgidus* MCM (*Af*MCM). It has been observed that the single-stranded or double-stranded DNA binding ability of MCM from *M.thermoautotrophicum* and *A.fulgidus* is not affected by the presence or the absence of ATP or divalent metal ions (Fletcher *et al.*, 2003; Kelman and Hurwitz, 2003; Grainge *et al.*, 2003). The ATPase activity of the euryarchaeal MCMs is stimulated by single-stranded and double-stranded DNA, while the ATPase activity of the crenarchaeal *Ss*MCM is not stimulated by DNA, but is much higher than that of *Mth*MCM (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Carpentieri *et al.*, 2002; Grainge *et al.*, 2003). Another distinction between the two archaeal kingdoms is the type of the zinc-binding motif. While the euryarchaeal species have a zinc-binding motif of C_4 type, the crenarchaeal ones have a C_3 type.

The mechanism of MCM loading at the archaeal origins is not yet understood. This process is not well understood in eukaryotes either. It is believed, however, that the overall mechanism will bear at least some similarity to that of bacteria, in which DnaC utilizes ATP binding and hydrolysis to assemble the DnaB helicase at the origin (Lee and Bell, 2000). The main characteristics of the interaction between DnaC and DnaB are the inhibition of DnaB helicase activity when DnaC is bound to ATP and the release of that inhibition on helicase loading and ATP hydrolysis (Wahle *et al.*, 1989a and b; Davey *et al.*, 2002). However, similar activities for the putative eukaryotic helicase loader, CDC6, have not been demonstrated. In archaea, the interactions among the initiation proteins are currently unknown. *In vitro* studies demonstrated that the *M.thermoaoutotrophicus* ORC-1 and ORC-2 proteins inhibit MCM helicase activity and it was therefore suggested that the proteins interact (Shin *et al.*, 2003a). Two-hybrid and Far Western blot analysis has demonstrated the interactions between MCM and the two ORC/CDC6 homologues identified in the *M.thermoautotrophicus* genome (*Mth*CDC6-1 and *Mth*CDC6-2) (Shin *et al.*, 2003a). Although the ORC/CDC6 homologues are very similar in primary amino acid sequence (Giraldo, 2003) and were suggested to have similar structure and domain organisation (Liu *et al.*, 2003; Singleton *et al.*, 2004), they appear to employ different regions for MCM binding. While *Mth*CDC6-1 binds MCM predominately via the WH domain (Shin *et al.*, 2003), this domain of *Mth*CDC6-2 does not interact with the MCM

helicase. Only the full-length *Mth*CDC6-2 protein has been shown to bind MCM. This is similar to the observations made with a CDC6 homologue from *S. sulfataricus* in which an indirect assay suggested that the WH domain of one of the three CDC6 homologues is not required for MCM binding (De Felice *et al.*, 2003). The major contact between *Mth*MCM and *Mth*CDC6 homologues is *via* the N-terminal portion of MCM (Kasiviswanathan *et al.*, 2004; Kasiviswanathan *et al.*, 2005). In contrast, the CDC6 homologue from *P.aerophilum* (*Pa*CDC6) did not bind to MCM (Shin *et al.*, 2003a and b).

5.2 Aim of the project

The analysis of the *A.pernix* genome sequence revealed the presence of a single open reading frame coding for a putative homologue of the eukaryotic MCM proteins (APE0188) (Kawarabayasi *et al.*, 1999). Using the computer program ClustalW, a sequence alignment of archaeal MCMs and MCM4 from *H.sapiens* and *S.cerevisiae* was performed. It was found that *Ape*MCM shares a 32% identity with the human and yeast MCM4, respectively. *Ape*MCM does not contain the N-terminal extension of about 160 amino acid residues found in the eukaryotic counterparts and the cyclin-dependent kinase phosphorylation sites that are clustered in the N-terminal region of the eukaryotic MCM4 (Kearsey and Labib, 1998; Tye *et al.*, 1999; Shechter *et al.*, 2000). Moreover, the core region of the archaeal and eukaryal sequences shows a high level of similarity because it contains the four amino acid motifs typically found in DNA helicases (Koonin, 1993b). As observed in Fig. 5.1, the sequence boxes A and B correspond to the Walker A and B motifs that are involved in for nucleotide binding and hydrolysis, respectively (Neuwald *et al.*, 1999). An arginine finger motif was also identified (Neuwald *et al.*, 1999) (Fig.5.1). In addition, the *Aeropyrum* and *Sulfolobus* MCM sequences contain a zinc binding motif of the HC_3-type (Poplawski *et al.*, 2001), whereas the MCM proteins from eukaryotic organisms (Kearsey and Labib, 1998; Tye, 1999) and the euryarchaeal species (Poplawski *et al.*, 2001) possess a zinc-binding motif of the HC_4-type (Fig.5.2).

All known helicases provide the free energy required to break hydrogen bonds in duplex DNA by coupling this reaction to the hydrolysis of nucleoside triphosphates. The goal of the work described in this chapter is to extend our understanding of the biochemical properties of the archaeal MCM protein. These properties were investigated by mutating conserved residues within MCM and clarifying the roles of several individual residues in the biochemical properties of MCM. The biochemical characteristics of the mutant proteins have been compared to those of wild type MCM.

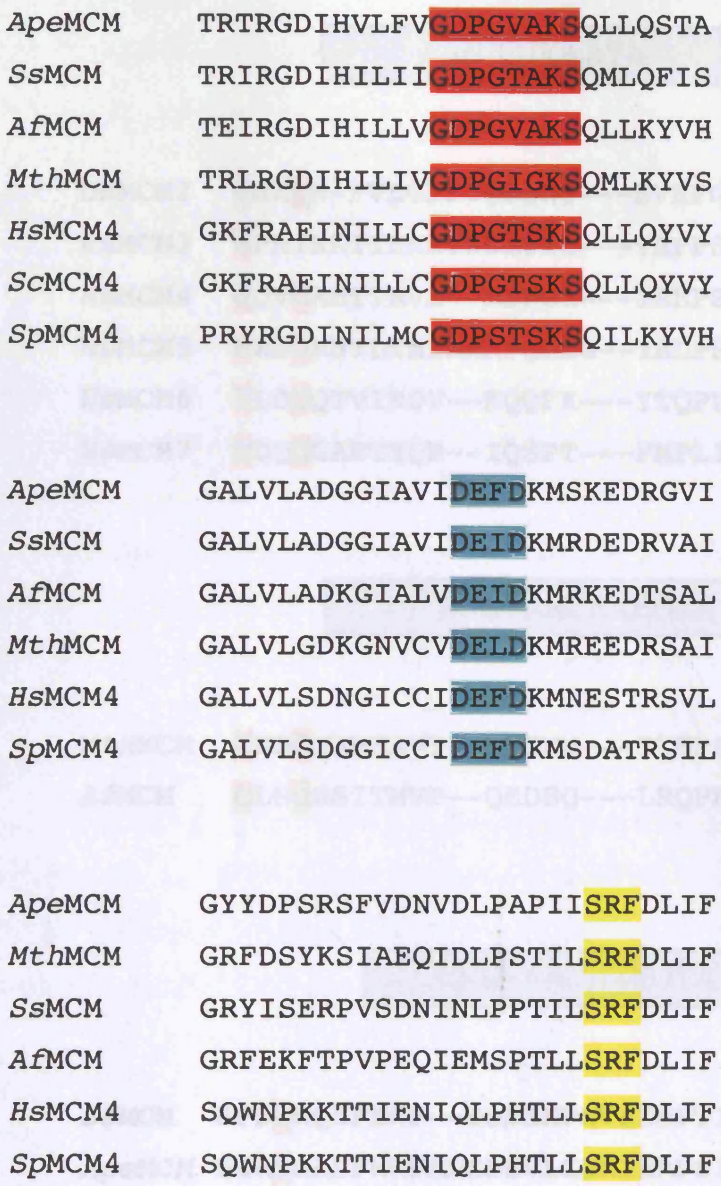


Fig. 5.1 Sequence alignment of archaeal MCM and eukaryal MCM4 proteins showing the Walker A (in red), Walker B (in green) and arginine finger (in yellow) motifs (Walker *et al.*, 1982; Koonin *et al.*, 1993a). Eukaryotic species: *Hs, Homo sapiens.* Euryarchaeal species: *Mth*, *Methanobacterium thermoautotrophicum*; *Af*, *Archaeoglobus fulgidus*. Crenarchaeal species: *Ss*, *Sulfolobus solfataricus*; *Ape*, *Aeropyrum pernix.*

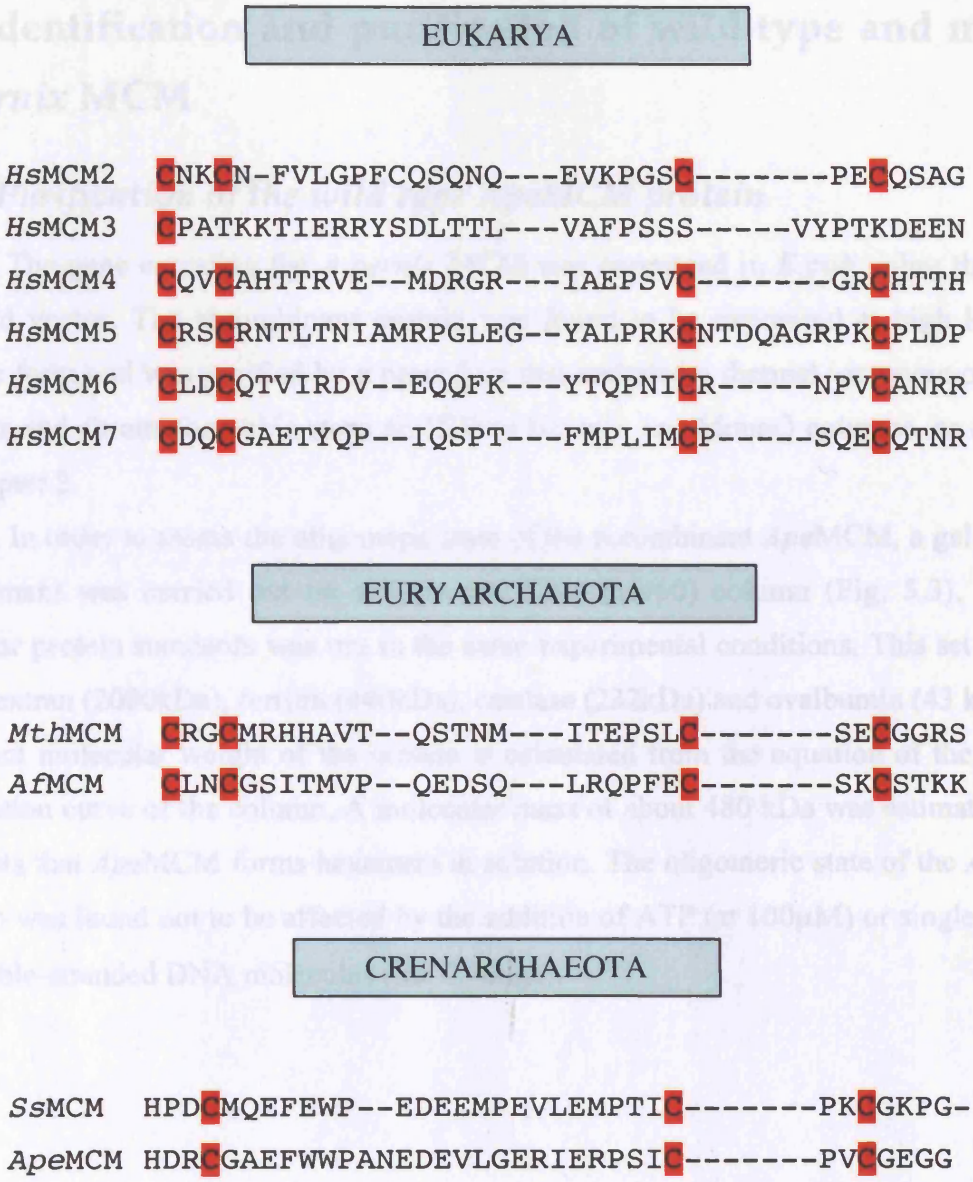


Fig. 5.2 Sequence alignment of archaeal MCM and eukaryal MCM4 proteins showing the zinc-binding domain. The residues capable of coordinating zinc are shown in red. Eukaryotic species: *Hs, Homo sapiens.* Euryarchaeal species: *Mth, Methanobacterium thermoautotrophicum*; *Af, Archaeoglobus fulgidus.* Crenarchaeal species: *Ss, Sulfolobus solfataricus*; *Ape, Aeropyrum pernix.*

5.3 Identification and purification of wild-type and mutant *A.pernix* MCM

5.3.1 Purification of the wild type ApeMCM protein

The gene encoding the *A.pernix* MCM was expressed in *E.coli* using the pET28 plasmid vector. The recombinant protein was found to be expressed at high level in a soluble form and was purified by a procedure that included a thermal treatment of the cell extracts and chromatographic steps on HiTrap Heparin and MonoQ columns, as described in Chapter 2.

In order to assess the oligomeric state of the recombinant *Ape*MCM, a gel filtration experiment was carried out on a Superdex S200 (16/60) column (Fig. 5.3). A set of globular protein standards was run in the same experimental conditions. This set included blue dextran (2000kDa), ferritin (440kDa), catalase (232kDa) and ovalbumin (43 kDa). The apparent molecular weight of the protein is calculated from the equation of the standard calibration curve of the column. A molecular mass of about 480 kDa was estimated which suggests that *Ape*MCM forms hexamers in solution. The oligomeric state of the *Ape*MCM protein was found not to be affected by the addition of ATP (at 100μM) or single-stranded or double-stranded DNA molecules (Ian Grainge).

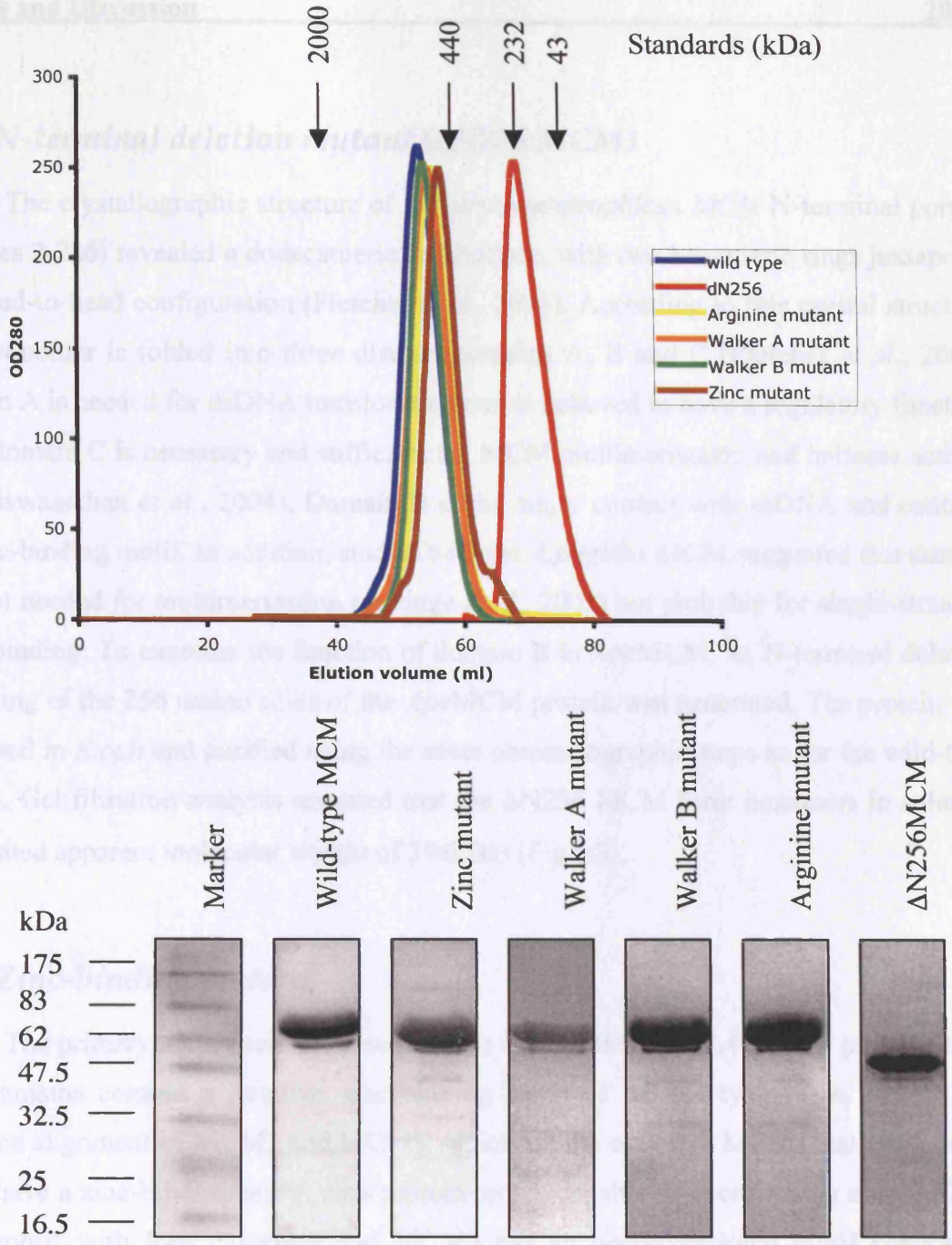


Fig. 5.3 Gel filtration analysis of *Ape*MCM protein and its mutants. (A) Elution profile of the purified proteins from Superdex S200 (16/60) gel filtration column. The arrows indicate the position of the protein markers used to calibrate the column (Blue dextran; Ferritin, Catalase and Ovalbumin). (B) Coomassie blue-stained SDS-PAGE of the wild type, zinc, Walker A, Walker B, arginine and ΔN256 mutant proteins after GF purification procedure.

5.3.2 *N-terminal deletion mutant (ΔN256 MCM)*

The crystallographic structure of *M.thermoautotrophicus MCM* N-terminal portion (residues 2-286) revealed a dodecameric architecture, with two hexameric rings juxtaposed in a head-to-head configuration (Fletcher *et al.*, 2003). According to this crystal structure, each monomer is folded into three distinct domains A, B and C (Fletcher *et al.*, 2003). Domain A is needed for dsDNA translocation and is believed to have a regulatory function, while domain C is necessary and sufficient for MCM multimerisation and helicase activity (Kasiviswanathan *et al.*, 2004). Domain B is the major contact with ssDNA and contains the zinc-binding motif. In addition, studies with the *A.fulgidus* MCM suggested that domain B is not needed for multimerisation (Grainge *et al.*, 2003) but probably for single-stranded DNA binding. To examine the function of domain B in *Ape*MCM, an N-terminal deletion consisting of the 256 amino acids of the *Ape*MCM protein was generated. The protein was expressed in *E.coli* and purified using the same chromatographic steps as for the wild-type protein. Gel filtration analysis revealed that the ΔN256 MCM form hexamers in solution (calculated apparent molecular weight of 294kDa) (Fig.5.3).

5.3.3 *Zinc-binding mutant*

The primary amino acid sequences of the eukaryotic MCM2,4,6 and 7 proteins from all organisms contain a putative zinc-binding motif of the C4 type (Tye, 1999). The sequence alignment of MCM3 and MCM5, which are the only two MCMs that are thought not to have a zinc-binding motif, also contain motifs capable of coordinating zinc. MCM5 has a motif with four cysteines and MCM3 has an unconventional motif containing **CXXTX$_{19}$SX$_7$E**, which can be aligned with the zinc motifs of other MCMs (Fig 5.2) (Yan *et al.*, 1993). All the residues in the proposed zinc motif of MCM3 (threonine, serine and glutamate/aspartate) are capable of participating in zinc atom coordination (Yan *et al.*, 1993).

A zinc-binding motif of the C_4-type has been identified in all archaea of the euryarchaeota kingdom for which the complete genome is known. The crystal structure of

the N-terminus of *Mt*MCM predicts that the zinc-binding motif is important in mediating double hexamer formation (Fletcher *et al*, 2003). A Cys to Ser mutation at the second Cys residue of the zinc motif is reported to result in reduced zinc binding, DNA binding and helicase activity, but has limited effects on the stability of the double hexamer (Poplawski *et al*, 2001). A serine at this position could still coordinate zinc to some extent, as in the case of MCM3 (Yan *et al.*, 1993). More debilitating changes, such as changes to alanine or changes of more than one cysteine residue at a time would have more profound effect on disrupting the hexamer-hexamer formation (Fletcher *et al*, 2003).

In contrast to the euryarchaeal species, the species from the crenarchaeota kingdom, to which *A.pernix* belongs, contain a zinc-binding motif of C_3 type. In order to elucidate the roles played by the zinc-binding domain of this type in *Ape*MCM function, the motif was mutated by substituting the cysteine residue 152 to an alanine. The mutant protein was expressed and purified from *E.coli* cells using the procedure previously described for the purification of the wild type enzyme. There are many examples in which zinc-binding domains participate in protein-protein interactions (Leon and Roth, 2000). The effect of the mutation in the zinc-binding domain on the oligomeric state of *Ape*MCM was determined using gel filtration. SDS-PAGE of the fractions obtained by the gel filtration experiments yielded protein peak at a similar position for the wild type and mutant protein. The protein peak was detected at a position similar to that of the wild-type enzyme (Fig. 5.3). The zinc mutant migrated from the column with an apparent molecular weight of 480kDa. Thus, this suggests that the zinc-binding motif of C_3-type of the *Ape*MCM protein is not needed to stabilise protein-protein interactions within the MCM complex. The same observation was made for the zinc-binding motif of *Mth*MCM (Poplawski *et al.*, 2001).

5.3.4 Walker A and Walker B mutants

The MCM proteins are members of the large AAA+ ATPase family and are required not only for replication initiation but also for elongation (Walter and Newport, 2000; Labib and Diffley, 2001; Forsburg, 2004). In eukaryotes, the intact MCM2-7 complex does not have any detectable biochemical activity (Adachi *et al.*, 1997; Ishimi *et al.*, 1998; Lee and Hurwitz, 2000; Sato *et al.*, 2000). However, a heterohexamer formed from a dimer of the trimer subcomplex MCM4,6,7 possesses *in vitro* single-stranded DNA and ATP binding activities as well as ssDNA-dependent ATPase and 3'-5' DNA helicase activities (Ishimi *et al.*, 1998, 2000; Lee and Hurwitz, 2000, 2001). Mutational analysis within the Walker A and Walker B motifs of mouse MCM4 and MCM6 has demonstrated the importance of these domains for the *in vitro* DNA helicase activity of the hexameric MCM4,6,7 complex (You *et al.*, 1999).

The Walker A motif has the consensus sequence, including the invariant lysine residue found in the ATP-binding proteins. The Walker A motif sequence in the MCMs contains a slight variation from the consensus observed in other AAA+ proteins. It has an alanine or serine in place of a glycine and some additional conserved residues, which create the MCM specific consensus GDPxx(S/A)KS (Fig.5.1) (Koonin, 1993). On the basis of studies of known ATPases, it is suggested that the lysine in the Walker A motif directly interacts with the phosphatyl group of bound ATP (Walker *et al.*, 1982). The Walker B motif has a consensus IDEFDKM and is believed to be responsible for ATP hydrolysis (Fig.5.1) (Neuwald *et al.*, 1999; Patel and Picha, 2000). Apparently the common denominator for all variants of the B motif is at least one negatively charged residue preceded by a stretch of bulky hydrophobic residues predicted to form a β-strand (Koonin, 1993). To better understand the importance of the ATP binding and hydrolysis motifs, mutations of two residues within the Walker A motif (K349E and S350E) and one residue within the Walker B motif (E408A) were constructed. The mutant proteins were expressed and purified from *E.coli* cells using the procedure previously described for the purification of the wild type enzyme. The mutant proteins were found to form stable hexamers in solution as the wild-type protein (Fig.5.3).

5.3.5 *Arginine mutant*

In some AAA+ proteins The ATP sites are located at the interface of subunits (Davey *et al.*, 2003; Ogura *et al.*, 2004; Neuwald *et al.*, 2005). Residues from both subunits are thought to be required for ATP hydrolysis (Davey *et al.*, 2003). An arginine residue of one subunit is required for hydrolysis of ATP bound to a neighbouring subunit (Davey *et al.*, 2003). A catalytic arginine residue that reaches across a subunit interface to promote nucleotide hydrolysis is referred to as an arginine finger, first described in the RAS GTP-binding protein and its cognate GAP (GTPase activating protein) (Ahmadian *et al.*, 1997). When bound to GTP, RAS is active, but is inactive when bound to GDP. RAS hydrolyses GTP very slowly and requires an arginine residue from GAP to promote rapid hydrolysis. Structural and biochemical studies have suggested that the arginine finger in GAP stabilises the transition state thereby greatly increasing the catalytic rate of the RAS GTPases (Ahmadian *et al.*, 1997).

Studies on the individual MCM subunits from the eukaryotic heterohexameric MCM2-7 complex demonstrated that no individual MCM protein contains significant ATPase activity even though each one has an ATP binding site. ATPase activity is produced by a combination of at least two MCM proteins, implying that ATPase activity requires residues from both subunits (e.g. a catalytic arginine from one subunit and an ATP binding site in another) (Davey *et al.*, 2003).

Sequence alignment of archaeal and eukaryotic MCMs (Fig.5.1) revealed that the conserved arginine finger motif (SRF) in eukaryotic MCMs is present in the archaeal homologues. To investigate the impact of the arginine finger motif on the biochemical properties of the *A.pernix* MCM helicase, the motif was mutated by substituting the arginine residue 476 to an alanine. The mutant protein was expressed and purified from *E.coli* cells using the procedure described for the purification of the wild type enzyme. The mutant protein form hexamers in solution as the wild-type *Ape*MCM (Fig.5.3).

5.4. Biochemical properties

5.4.1 DNA binding activity

Some DNA helicases possess strong affinity for ssDNA (Lohman and Bjornson, 1996; Hippel and Delagoutte, 2001; Shin *et al.*, 2003). The polar nature of the unwinding reaction is expected to require a polar interaction with ssDNA, and this has been demonstrated in DnaB, MCM2-7 and archaeal MCM as an asymmetric binding of radioactively labelled DNA molecules (Biswas *et al.*, 2002; Shin *et al.*, 2003). Interactions with single-stranded DNA are commonly required for initiation of the helicase reaction *in vitro*. However, it should be noted that ssDNA tails will rarely exist *in vivo* and that the helicase is presumably loaded onto its substrate by accessory proteins. Some helicases display a specificity for a particular DNA structure (Shin *et al.*, 2003). Allosteric regulation of the DNA binding properties of the MCM helicase is incompletely understood. However, this enzyme presumably works by cycling vectorally through a series of nucleotide-dependent DNA binding states. Therefore, a good understanding of the DNA binding properties of this enzyme will be informative for its mechanism.

The ability of the wild type and mutants of *Ape*MCM protein to bind single-stranded and double-stranded DNA was investigated using the electrophoretic mobility shift assay technique. For these assays, ^{32}P-labelled oligonucleotides were incubated with increasing amounts of wild type and mutant forms of the *Ape*MCM protein in the presence and absence of ATP. A variety of DNA substrates were used to assay the binding preference of the *Ape*MCM and the mutant proteins (see Fig. 4.4, Chapter 4). These encompassed substrates of single-stranded or double-stranded DNA, or substrates containing both. All of the different substrates in the assays were bound by the wild type enzyme and some of the mutants (Fig. 5.4 and Fig.5.5). The extent of the binding was dependent upon the concentration of protein used in the reaction. Furthermore, this binding was not affected by the presence of ATP (Ian Grainge). In substrates containing only 25 bases of single-stranded or double-stranded DNA, a single-band was observed on the polyacrylamide gels (Fig. 5.4 A and B). However when longer substrates were used, a

second slower migrating band was detected (Fig. 5.4 C, D, E and Fig. 5.5). This could be a result of binding of initially one and then two hexamers to a DNA substrate, when the DNA is long enough to accommodate both. To quantitatively analyse the data for these interactions, the binding curves were fitted to an independent binding site equation and Hill equation when possible (Chapter 2). Binding curves were repeated in triplicate and the plots show the average and standard deviations from the three data sets. Fig. 5.4, Fig. 5.5 and show plots of the percentage of bound substrate against the concentration of *Ape*MCM protein. In the plot the amount bound represents substrate shifted by both complexes (both bands). Although binding to a variety of DNA structures was analysed, there was very little variation in binding efficiency between them, with the exception of those containing a single-stranded bubble flanked by duplex DNA. This suggests that a bubble substrate is a preferred ligand for the MCM hexamer (Fig. 5.6). A cooperative binding between hexamers was observed on bubble substrates that are absent on the other DNA substrates. Furthermore, the size of the single-stranded region of the bubble affects this cooperativity. With the increase of the single-stranded bubble region, the degree of cooperativity increased.

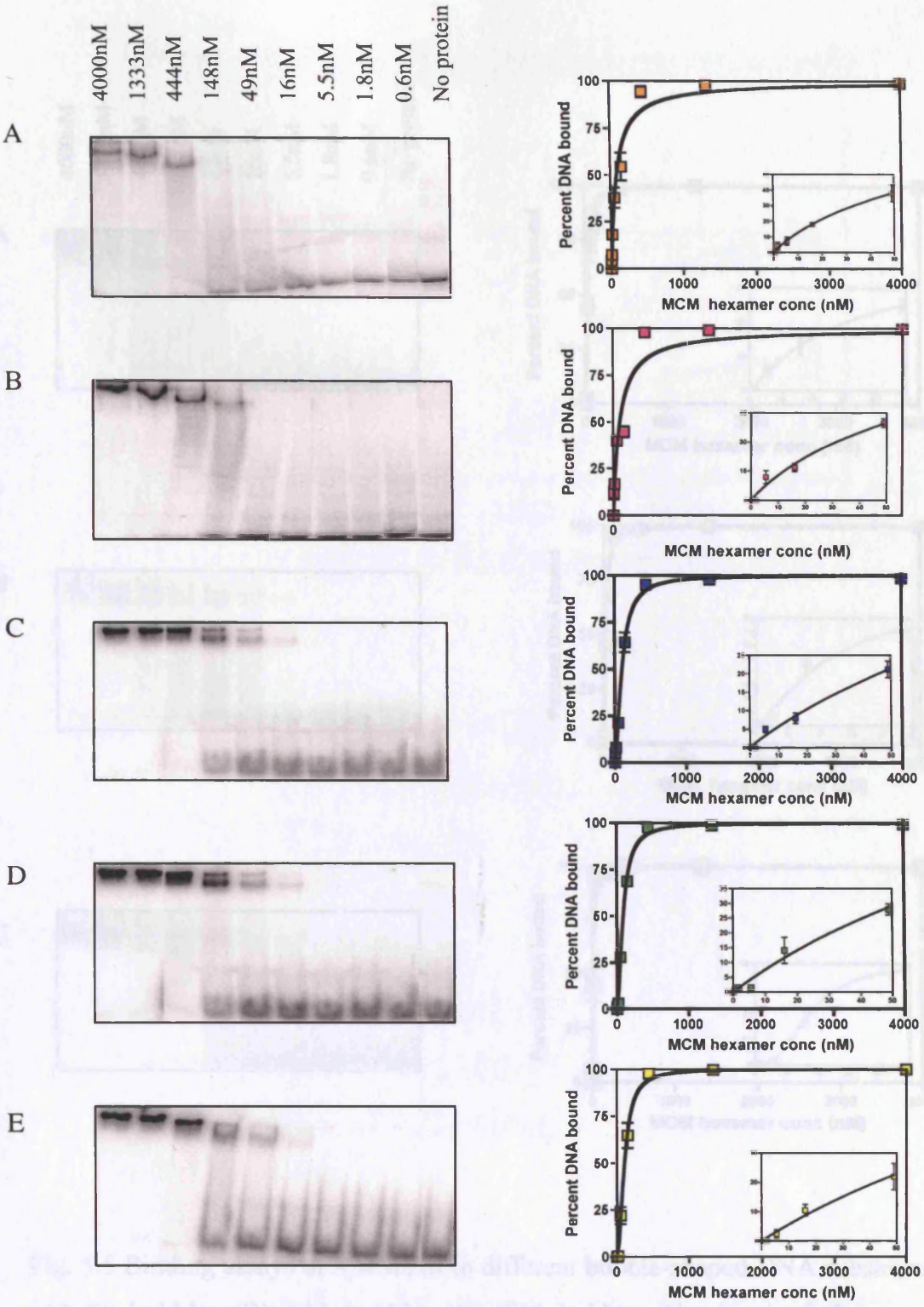


Fig. 5.4 Binding assays of *Ape*MCM to different DNA substrates. (A) Single-stranded; (B) Blunt duplex; (A) 3'-tailed; (D) 5'-tailed; (E) Flayed duplex. The inset of the graph represents the binding affinity of *Ape*MCM at concentrations up to 50nM per hexamer. The percentage bound represents substrate shifted by the two complexes (two bands observed in C, D, A).

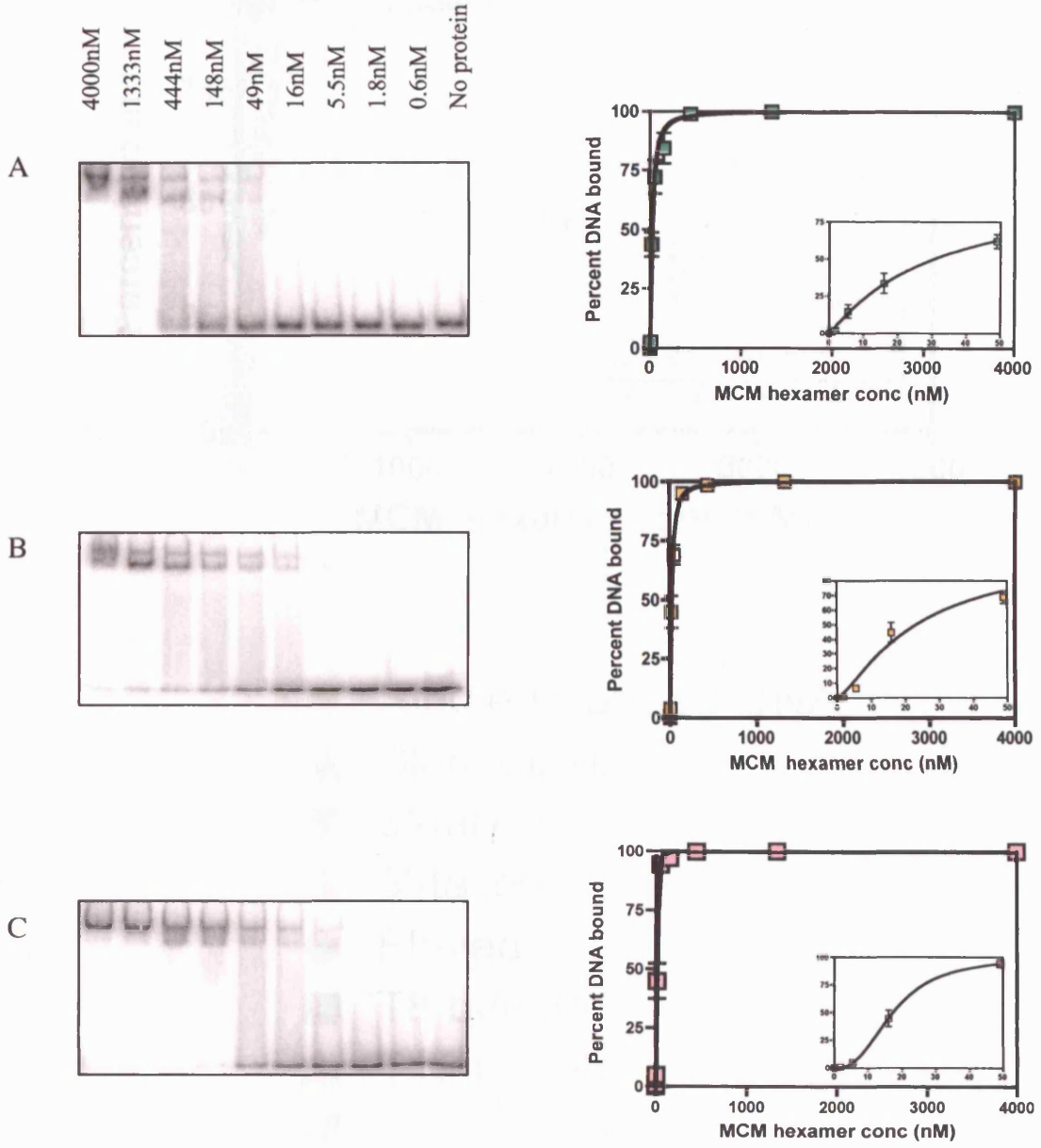


Fig. 5.5 Binding assays of *Ape*MCM to different bubble-shaped DNA substrates. (A) T8 bubble; (B) T15 bubble; (C) T30 bubble. The inset of the graph represents the binding affinity of *Ape*MCM at concentrations up to 50nM per hexamer. The percentage bound represents substrate shifted by the two complexes (two bands). *Ape*MCM binds bubble substrates in a cooperative manner.

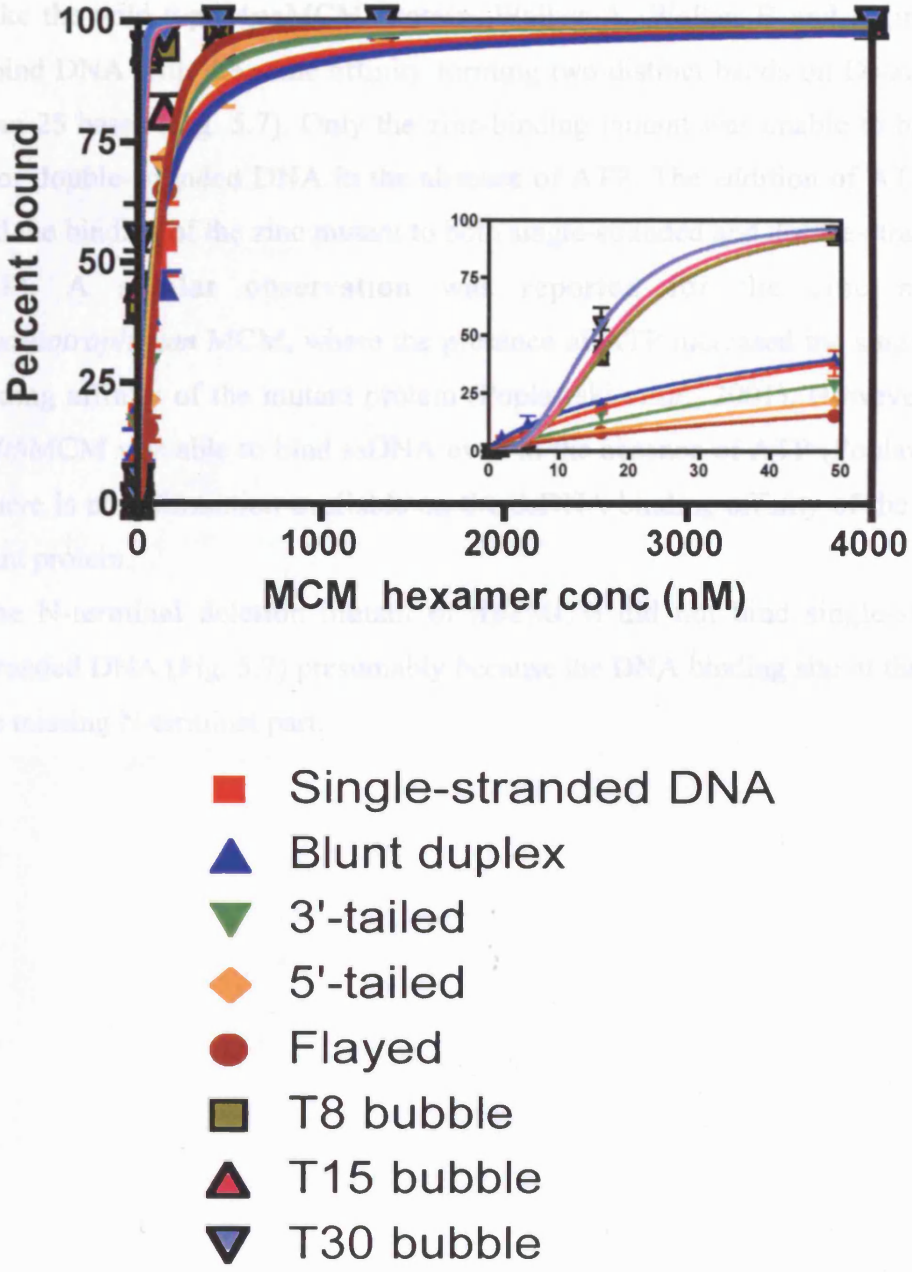


Fig. 5.6 Graphic representation of *Ape*MCM binding to various DNA substrates. (A) Variation in DNA binding at concentrations up to 4000nM per hexamer, (B) Enlarged view of DNA binding at concentrations up to 50nM per hexamer. The percentage bound represents substrate shifted by the two complexes. Cooperative binding between two hexamers was observed only on bubble substrates.

Like the wild type *Ape*MCM protein, Walker A, Walker B and arginine finger mutants bind DNA with the same affinity forming two distinct bands on DNA substrates longer than 25 bases (Fig. 5.7). Only the zinc-binding mutant was unable to bind single-stranded or double-stranded DNA in the absence of ATP. The addition of ATP however stimulated the binding of the zinc mutant to both single-stranded and double-stranded DNA (Fig. 5.8). A similar observation was reported for the zinc mutant of *M.thermoautotrophicum* MCM, where the presence of ATP increased the single-stranded DNA binding affinity of the mutant protein (Poplawski *et al.*, 2001). However, the zinc mutant *Mth*MCM was able to bind ssDNA even in the absence of ATP (Poplawski *et al.*, 2001). There is no information available on the dsDNA binding affinity of the *Mth*MCM zinc mutant protein.

The N-terminal deletion mutant of *Ape*MCM did not bind single-stranded or double-stranded DNA (Fig. 5.7) presumably because the DNA binding site of the protein is within the missing N-terminal part.

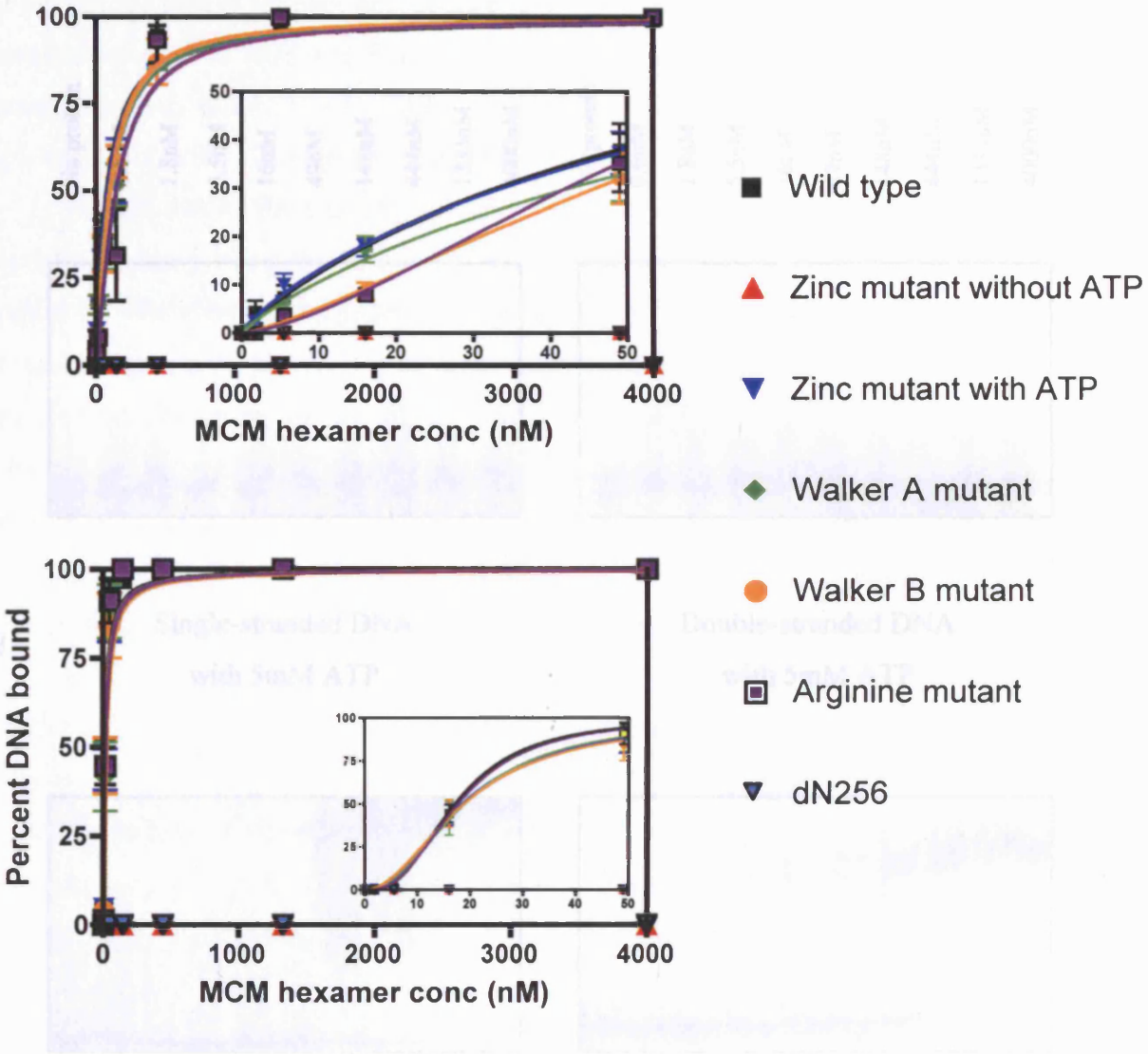


Fig. 5.7 Graphic representation of the DNA binding activity of *Ape*MCM and its mutant proteins. (A) Binding to a single-stranded 25 base long DNA; (B) Binding to T30 bubble. The insets of the graphs represent an enlarged view of DNA binding at concentrations up to 50nM per hexamer. The percentage bound represents substrate shifted by the two complexes. Cooperative binding between two hexamers was observed only on the bubble substrate.

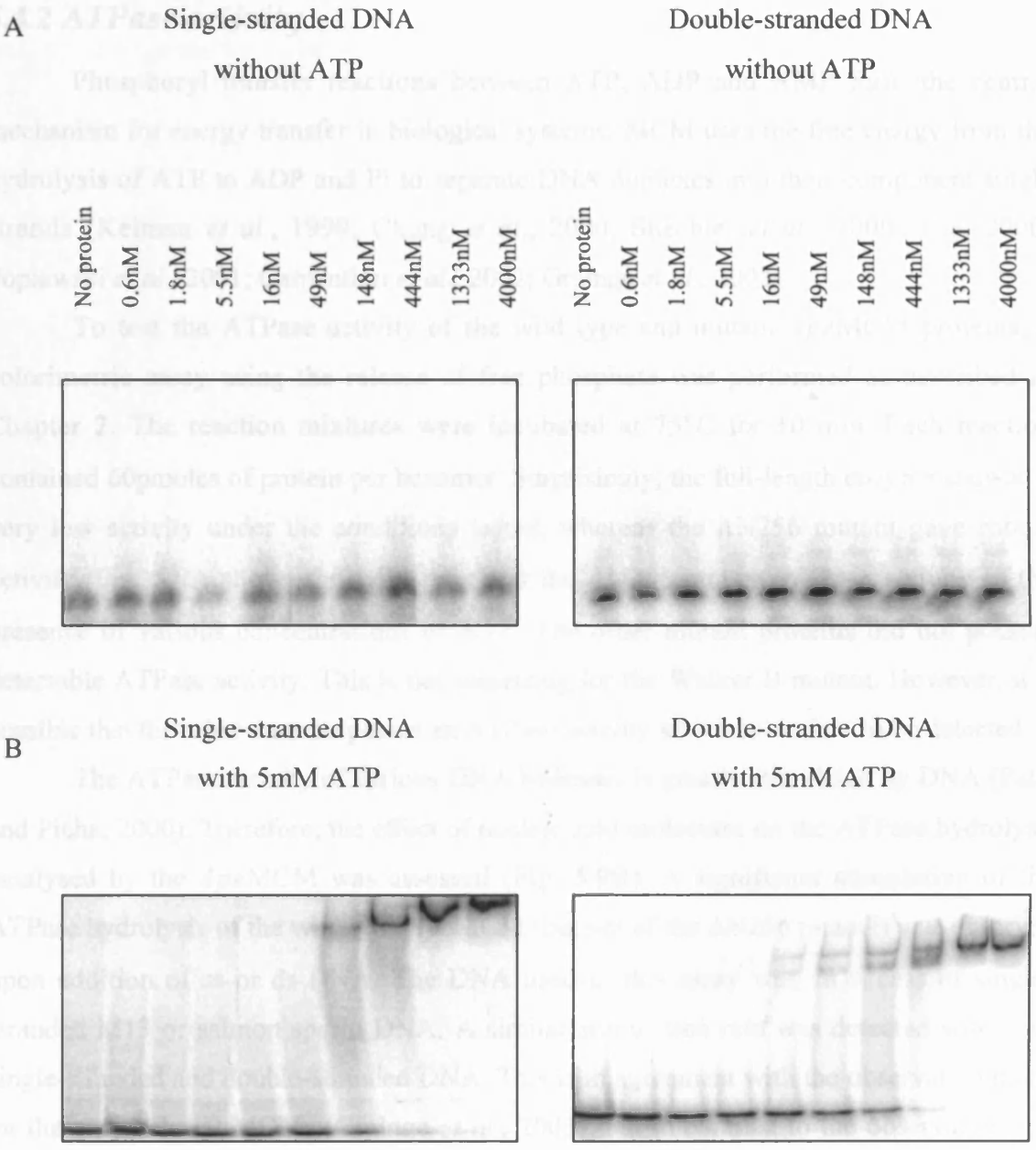


Fig. 5.8 DNA binding assays of the zinc mutant of *Ape*MCM. (A) Binding to a single-stranded and double-stranded T30 bubble without ATP, (B) Binding to the same DNA substrates in the presence of 5mM ATP.

5.4.2 *ATPase activity*

Phosphoryl transfer reactions between ATP, ADP and AMP form the central mechanism for energy transfer in biological systems. MCM uses the free energy from the hydrolysis of ATP to ADP and Pi to separate DNA duplexes into their component single strands (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Tye, 2000; Poplawski *et al.*, 2001; Carpentieri *et al.*, 2002; Grainge *et al.*, 2003).

To test the ATPase activity of the wild type and mutant *Ape*MCM proteins, a colorimetric assay using the release of free phosphate was performed as described in Chapter 2. The reaction mixtures were incubated at 75°C for 10 min. Each reaction contained 60pmoles of protein per hexamer. Surprisingly, the full-length enzyme showed a very low activity under the conditions tested, whereas the ΔN256 mutant gave robust activity. Fig. 5.9A shows the wild type and the ΔN256 mutant ATPase activity in the presence of various concentrations of ATP. The other mutant proteins did not possess detectable ATPase activity. This is not surprising for the Walker B mutant. However, it is possible that the other mutants posses an ATPase activity which is too low to be detected.

The ATPase activity of various DNA helicases is greatly stimulated by DNA (Patel and Picha, 2000). Therefore, the effect of nucleic acid molecules on the ATPase hydrolysis catalysed by the *Ape*MCM was assessed (Fig. 5.9B). A significant stimulation of the ATPase hydrolysis of the wild type *Ape*MCM (but not of the ΔN256 mutant) was observed upon addition of ss or ds DNA. The DNA used in this assay was an excess of single-stranded M13 or salmon sperm DNA. A similar stimulation rate was detected with both single-stranded and double-stranded DNA. This is in agreement with the observation made on the euryarchaeal MCMs (Grainge *et al.*, 2003), but in contrast to the observations on another creanarchaeal protein, the *S.solfataricus* MCM (Carpentieri *et al.*, 2002). A report on the ATPase activity of *S.solfataricus* MCM suggested that the ATP hydrolysis is not affected by the presence of ss or ds DNA (Carpentieri *et al.*, 2002). Surprisingly, ss and ds DNA stimulated the ATPase activity of the zinc mutant protein equally. This is in contrast with the observations on the zinc mutant *Mth*MCM protein (Poplawski *et al.*, 2001). The zinc-binding mutant of *Mth*MCM (C_4 type) was found to possess a very low ATPase

activity in the absence of DNA. Neither ss nor ds DNA, however, stimulated the ATPase activity of the mutant protein.

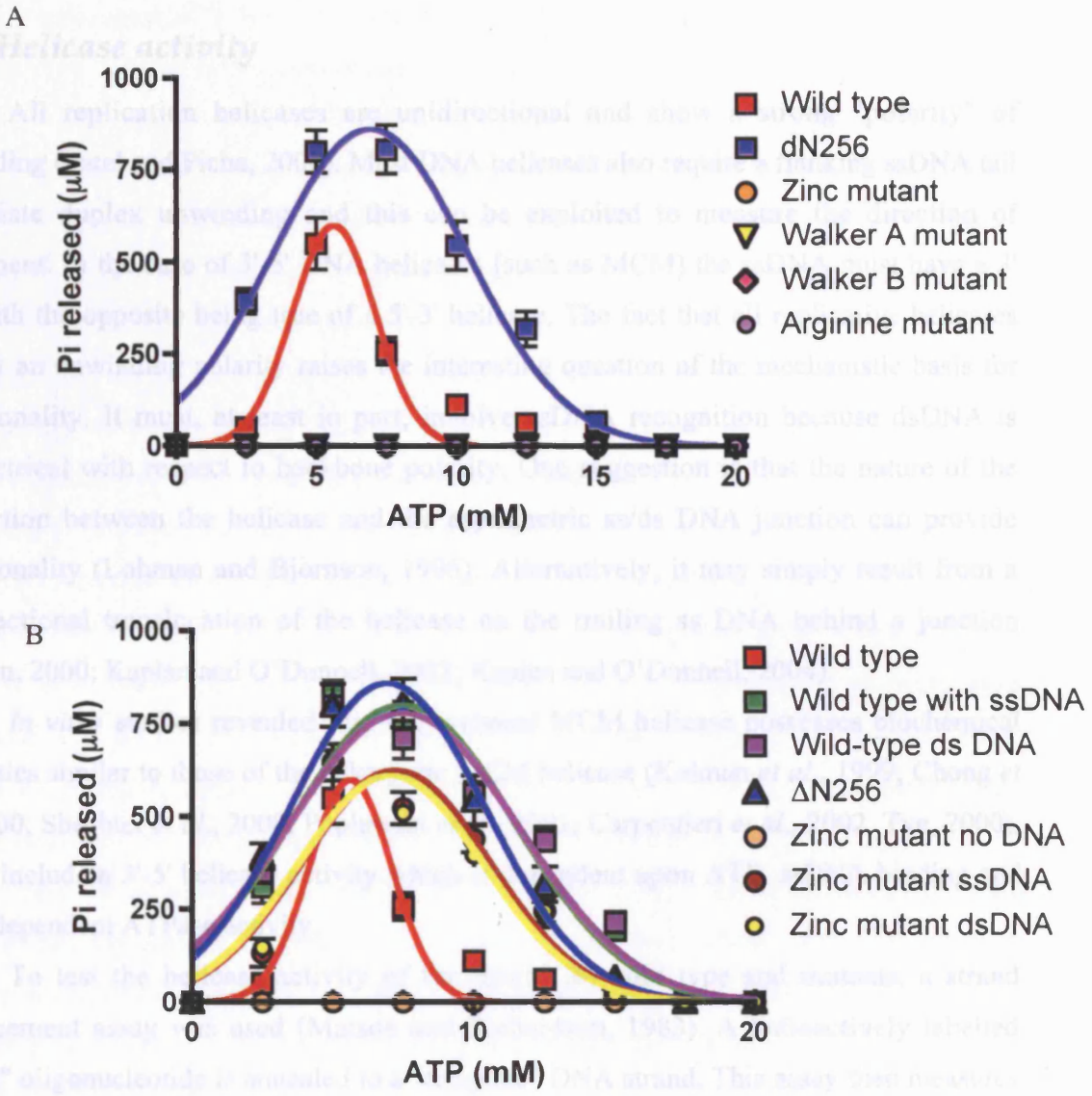


Fig. 5.9 ATPase assay. (A) Comparison of the ATPase activity of the wild type *Ape*MCM and its mutants. (B) The effect of single-stranded and double-stranded DNA on the ATPase activity of the wild type *Ape*MCM and the zinc mutant in comparison to the ATPase activity of ΔN256 mutant.

5.4.3 *Helicase activity*

All replication helicases are unidirectional and show a strong "polarity" of unwinding (Patel and Picha, 2000). Most DNA helicases also require a flanking ssDNA tail to initiate duplex unwinding and this can be exploited to measure the direction of movement. In the case of 3'-5' DNA helicases (such as MCM) the ssDNA must have a 3' end with the opposite being true of a 5'-3' helicase. The fact that all replicative helicases display an unwinding polarity raises the interesting question of the mechanistic basis for directionality. It must, at least in part, involve ssDNA recognition because dsDNA is symmetrical with respect to backbone polarity. One suggestion is that the nature of the interaction between the helicase and the asymmetric ss/ds DNA junction can provide directionality (Lohman and Bjornson, 1996). Alternatively, it may simply result from a unidirectional translocation of the helicase on the trailing ss DNA behind a junction (Kaplan, 2000; Kaplan and O'Donnell, 2002; Kaplan and O'Donnell, 2004).

In vitro studies revealed that the archaeal MCM helicase possesses biochemical properties similar to those of the eukaryotic MCM helicase (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Poplawski *et al.*, 2001; Carpentieri *et al.*, 2002, Tye, 2000). These include a 3'-5' helicase activity which is dependent upon ATP, ssDNA binding and DNA dependent ATPase activity.

To test the helicase activity of the *Ape*MCM wild type and mutants, a strand displacement assay was used (Matson and Richardson, 1983). A radioactively labelled "target" oligonucleotide is annealed to a "template" DNA strand. This assay then measures the ability of the helicase to displace the labelled strand from the template by following the alteration of its gel mobility on non-denaturing gels. Annealed and displaced strands can be quantified using a phosphorimager. The substrate molecule can be modified in a number of ways. For instance, the length of the duplex region can be altered or non-homologous "tails" can be added to the duplex region to investigate requirements for forked DNA substrates. Again, this assay is non-continuous and suffers from the disadvantage of high re-annealing rates after displacement if a short template strand is used. This problem can be circumvented by using longer templates. However, this creates a secondary problem of

extensive uncoupled ATP hydrolysis due to the presence of a large excess of ssDNA. In all cases, *in vitro* helicase substrates are likely to differ substantially from the enzyme's natural substrate, a topological complex DNA that is coated with structural and regulatory proteins.

The recombinant wild type and mutant *Ape*MCM proteins were tested for DNA helicase activity by a strand-displacement assay performed at 60°C for 30 minutes. Forked and blunt duplex DNA was used in these assays. Despite the poor ATPase activity of the full-length protein, it showed robust helicase activity on the forked DNA duplex. As expected, the *Ape*MCM was unable to unwind the blunt DNA duplex (Fig. 5.10). The variation of wild-type MCM helicase activity on the forked substrate with different concentrations of ATP was analysed. Controls consisted of a reaction without ATP, where no activity is observed, and a reaction where the substrate is boiled and placed immediately on ice to maximise denaturation of the DNA. As Fig. 5.10 and Fig. 5.11 shows, helicase activity is very low when ATP is < 5mM but higher activity is seen with the concentration tested at or > 5mM. The helicase activities of the wild-type and mutant proteins were compared. The ΔN256 MCM has a higher activity than the wild-type protein under these conditions, probably because it has a higher ATPase activity as well. These data suggest that the helicase activity remained coupled with the ATPase activity.

Next, the presence of the *Ape*ORC homologues on the helicase activity of the wild type and deletion mutant *Ape*MCM was analysed. First, two-hybrid system was used to assess possible interactions between *Ape*MCM and *Ape*ORC-1 and *Ape*ORC-2 proteins (as described in Chapter 2 and Chapter 3). MCM self-interaction was used as a positive control. Positive interactions were not detected using this approach (Fig. 5.12). Next, the helicase assay with *Ape*MCM was performed in the presence of ORC-1 and ORC-2 proteins. As Fig. 5.13 shows full length ORC-2 and Domain III of ORC-2 inhibit the helicase activity of the wild type MCM, but not the helicase activity of ΔN256 MCM. This is in agreement with the observation that the main contact between MCM and ORC-2 is *via* the N-terminal portion of MCM protein and the winged helix domain of ORC-2 (Kasiviswanathan *et al.*, 2005). ORC-1 protein also inhibited the helicase activity of the wild type MCM, but not of the ΔN256 MCM. However, even Domain I+II of ORC-1

inhibited the helicase activity of MCM. This suggests that may be within Domain I+II there is another MCM interaction site.

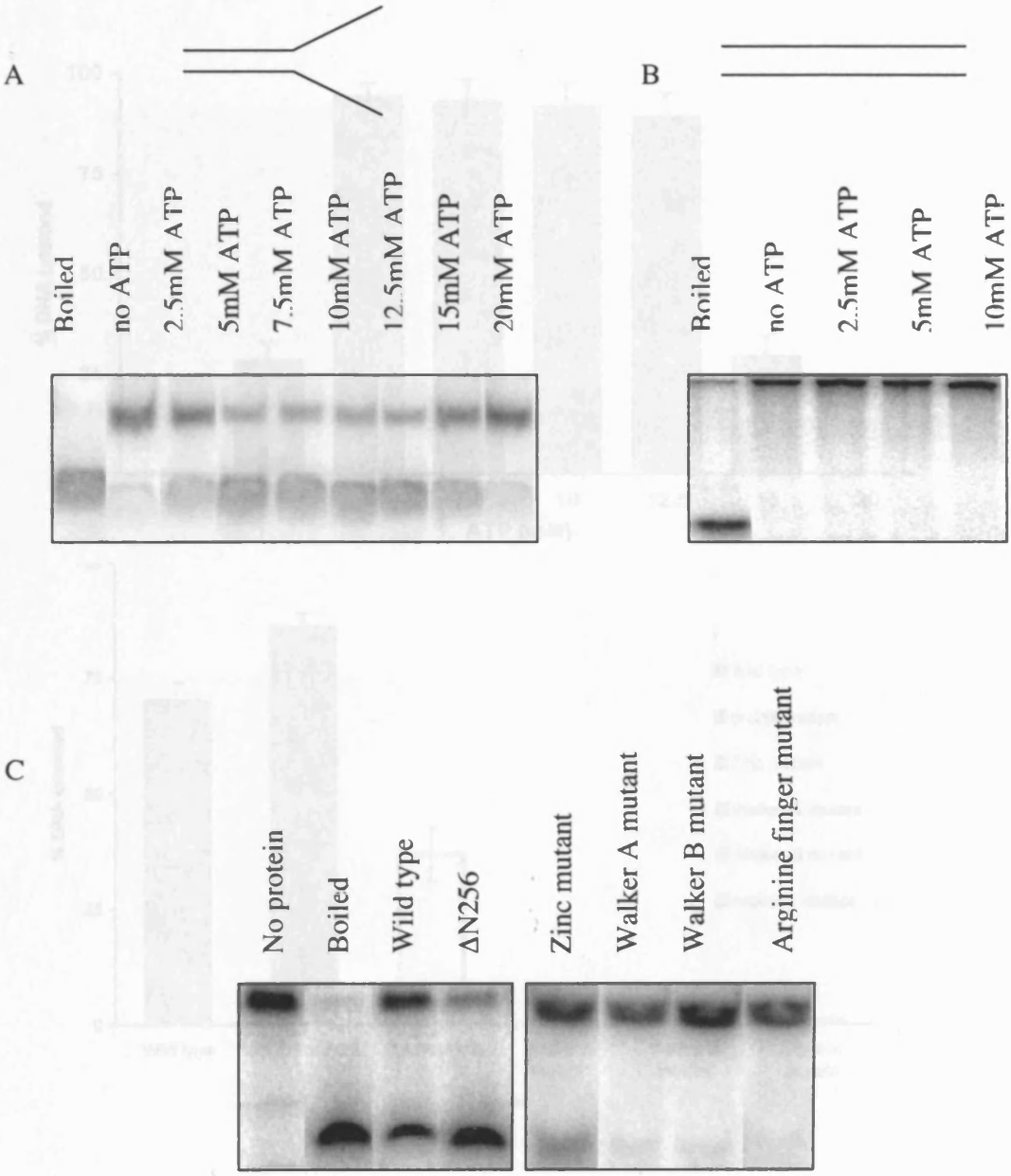


Fig. 5.10 Helicase assay for wild type and mutants of *Ape*MCM. (A) The effect of ATP concentration on the helicase activity of the wild type MCM assayed on flayed duplex and blunt duplex DNA substrates. (B) Comparison of the helicase activity of the wild type and mutant proteins in the presence of 5mM ATP on a flayed duplex DNA.

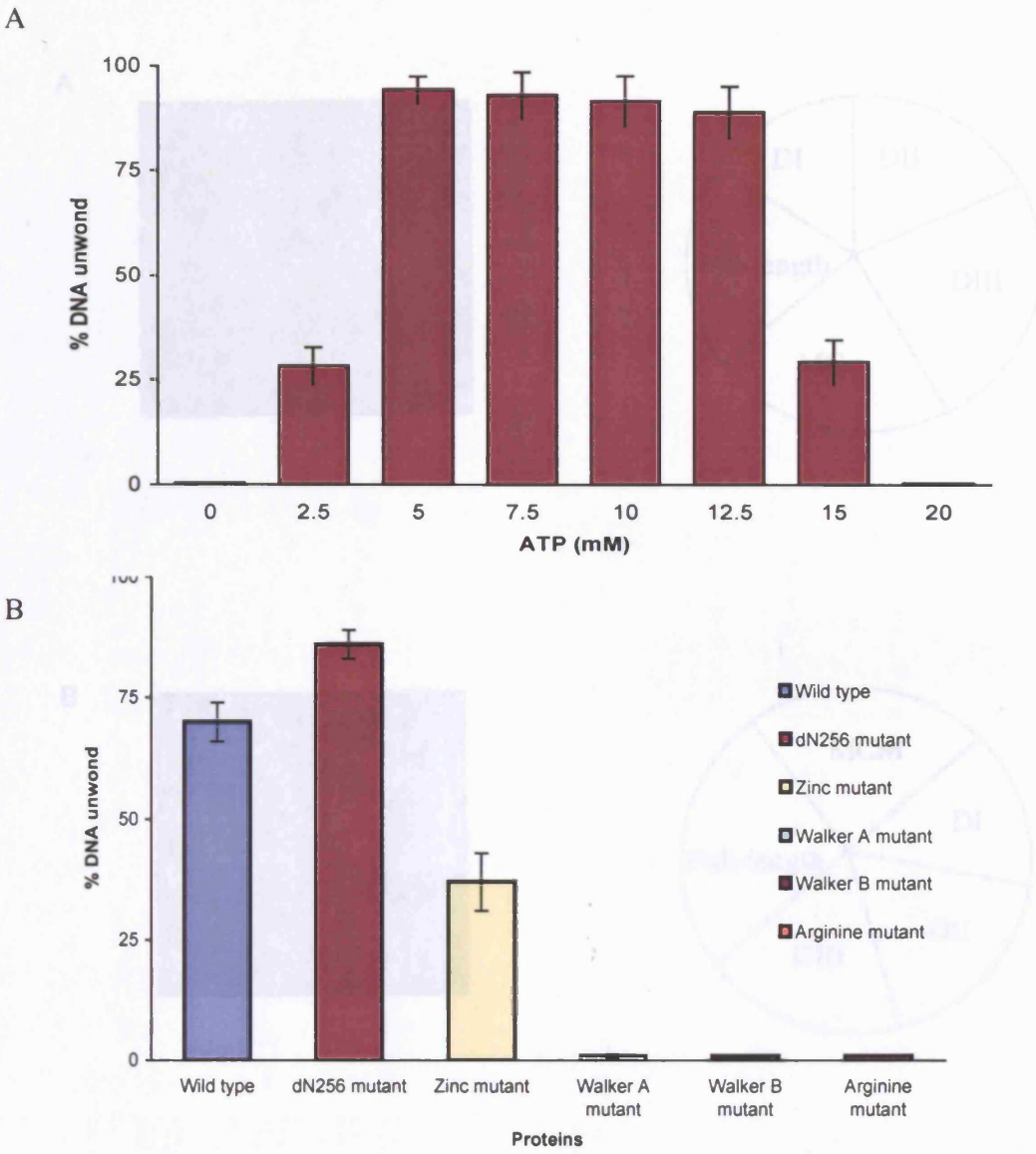
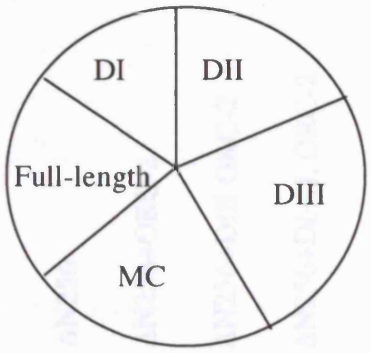
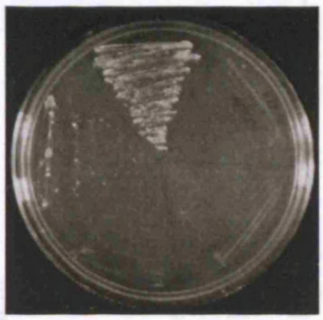


Fig. 5.11 Graphic representation of the helicase activity of wild type and mutant *Ape*MCM proteins on flayed duplex DNA substrate. (A) Effect of ATP concentration on the helicase activity of the wild type enzyme; (B) Helicase activity of the mutant proteins in the presence of 5mM ATP.

A



B



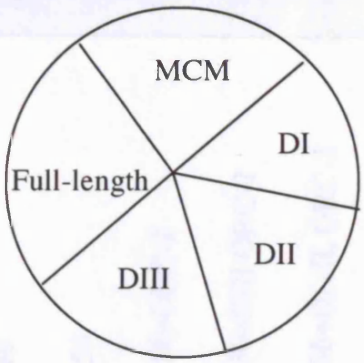


Fig. 5.12 Interaction mating between *Ape*MCM and (A) *Ape*ORC-1, and (B) *Ape*ORC-2 proteins. MCM self-interaction was used as a positive control. Positive interactions were not detected using this approach.

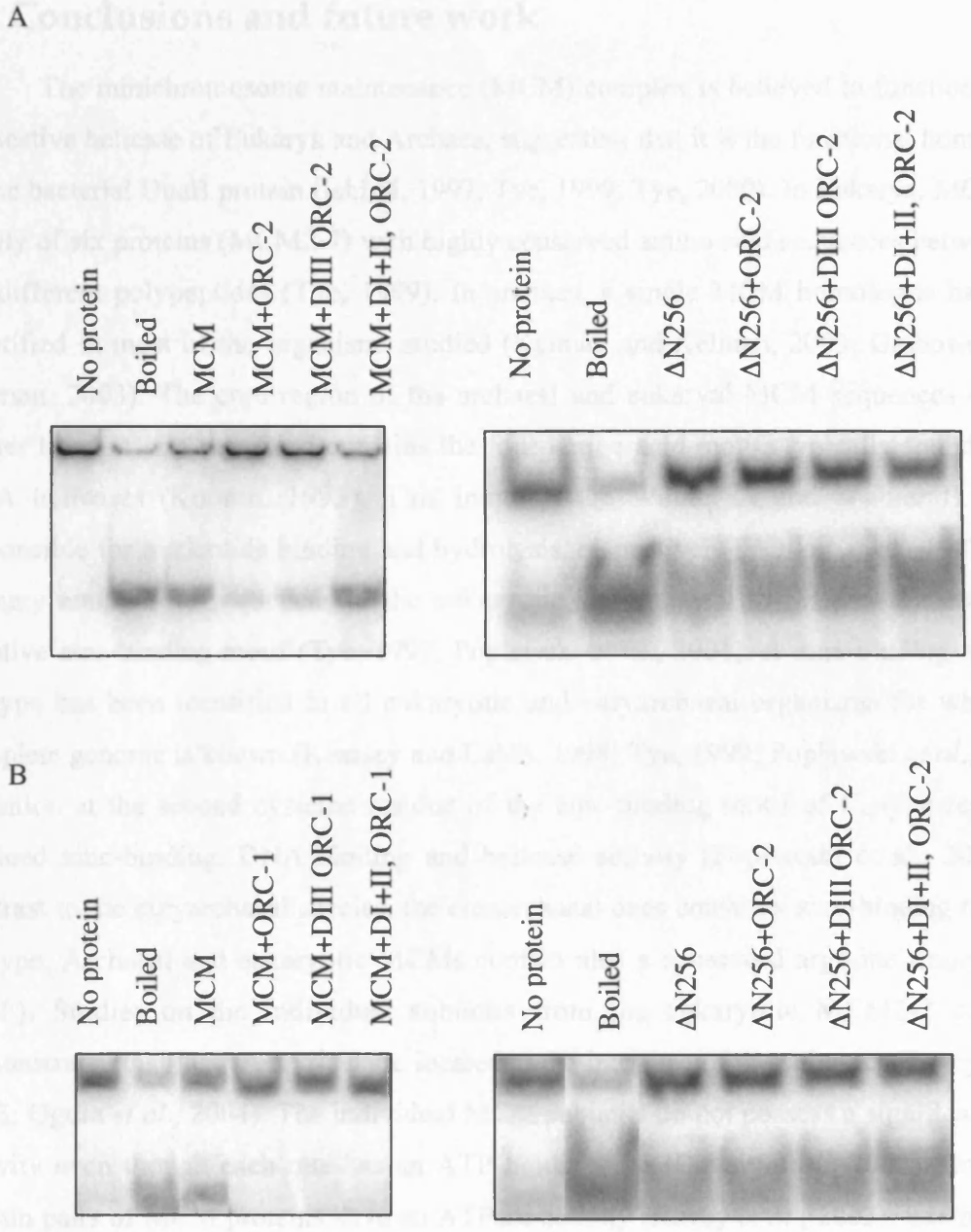


Fig. 5.13 Helicase assay for *Ape*MCM and ΔN256 mutant on a flayed duplex DNA in the presence of 5mM ATP. (A) In the presence of ORC-2, (B) In the presence of ORC-1.

5.5 Conclusions and future work

The minichromosome maintenance (MCM) complex is believed to function as the replicative helicase of Eukarya and Archaea, suggesting that it is the functional homologue of the bacterial DnaB protein (Ishimi, 1997; Tye, 1999; Tye, 2000). In Eukarya, MCM is a family of six proteins (MCM2-7) with highly conserved amino acid sequences between the six different polypeptides (Tye, 1999). In archaea, a single MCM homologue has been identified in most of the organisms studied (Kelman and Kelman, 2003; Grabowski and Kelman, 2003). The core region of the archaeal and eukaryal MCM sequences show a higher level of similarity and contains the four amino acid motifs typically found in the DNA helicases (Koonin, 1993). This includes the Walker A and Walker B motifs responsible for nucleotide binding and hydrolysis, respectively (Walker *et al.*, 1982). The primary amino acid sequences of the eukaryotic and archaeal MCM proteins contain a putative zinc-binding motif (Tye, 1999; Poplawski *et al.*, 2001). A zinc-binding motif of C_4-type has been identified in all eukaryotic and euryarchaeal organisms for which the complete genome is known (Kearsey and Labib, 1998; Tye, 1999; Poplawski *et al.*, 2001). Mutation at the second cysteine residue of the zinc-binding motif of C_4-type results in reduced zinc-binding, DNA binding and helicase activity (Poplawski et al., 2001). In contrast to the euryarchaeal species, the crenarchaeal ones contain a zinc-binding motif of C_3-type. Archaeal and eukaryotic MCMs contain also a conserved arginine finger motif (SRF). Studies on the individual subunits from the eukaryotic MCM2-7 complex demonstrated that their ATP sites are located at the interface of the subunits (Davey *et al.*, 2003; Ogura *et al.*, 2004). The individual MCM subunits do not possess a significant ATP activity even though each one has an ATP binding site (Davey *et al.*, 2003). However, certain pairs of MCM proteins have an ATPase activity (Davey *et al.*, 2003). Examination of these ATPase pairs revealed that one MCM subunit binds ATP and the neighbouring subunit contributes a catalytic arginine residue (Davey *et al.*, 2003).

The analysis of *A.pernix* genome revealed a presence of a single MCM homologue (Kawarabayasi *et al.*, 1999). Sequence alignment of *A.pernix* MCM and MCM homologues from other archaeal organisms confirmed the presence of all four motifs described above.

Conserved residues from these motifs were mutated and the biochemical properties of the mutants were investigated in comparison to the wild type protein. Biochemical studies with the MCM proteins from. *M.thermoautotrophicus* (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Poplawski *et al.*, 2001; Kasiviswanathan *et al.*, 2005); *A.fulgidus* (Grainge *et al.*, 2003) and *S.solfataricus* (Carpentieri *et al.*, 2002) revealed that these enzymes possess biochemical properties similar to those of the eukaryotic MCM4,6,7 complex including 3'-5' helicase, ssDNA binding and DNA-dependent ATPase activities. To investigate the biochemical properties of the wild type and mutant *Ape*MCM proteins, they were expressed in *E.coli* and purified using HiTrap Heparin and MonoQ columns. All purified proteins migrated as a Coomassie blue-stained protein bands of a predicted size (80kDa) in a 12% SDS-polyacrylamide gel. Gel filtration analyses performed on Superdex S200 (16/60) column suggest that wild type and mutant proteins form single hexamers in solution. The oligomeric state of the proteins was not affected by the addition of ATP or single-stranded or double-stranded DNA (Ian Grainge). With the exception of the zinc-binding mutant, all other proteins were able to bind single-stranded and double-stranded DNA of different shapes in the absence of ATP. This suggests that the ATP hydrolysis is required for a function that occurs after the formation of the DNA-MCM complex. One possible requirement for ATP hydrolysis is for the translocation of the protein around the DNA molecule, which may be required for the MCMs to identify an origin marked by ORC. Previous studies on the *M.thermoautotrophicus* MCM have shown that the protein forms dodecamers in solution (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000). The *S.pombe* MCM4,6,7 complex was also shown to form double hexamers but only in the presence of a flayed DNA substrate (Lee and Hurwitz, 2001). Interestingly, gel mobility shift analysis indicated that *S.pombe* MCM forms a faster migrating band that was shown to be a hexamer and slower migrating band that was shown to be a double hexamer (Lee and Hurwitz, 2001). These observations are similar to the gel mobility shift assays presented here, in which wild type and mutant proteins formed faster and a slower migrating band in the presence of DNA substrates longer than 25 bases. Although binding to a variety of DNA structures was analysed, there was very little variation in binding efficiency between them, with the exception of those containing a single-stranded bubble

flanked by duplex DNA. Cooperative binding between hexamers was observed on bubble substrates that are absent on the other DNA substrates. The size of the single-stranded region of the bubble affects the degree of cooperativity. With the increase of the single-stranded bubble region, the degree of cooperativity increased. This suggests that a bubble substrate is a preferred ligand for the MCM hexamer. Unlike the wild type protein, the zinc binding mutant and the N-terminal deletion mutant ΔN256 were unable to bind single- or double-stranded DNA in the absence of ATP. Upon addition of 5mM ATP, however, the zinc-binding mutant was able to bind single-stranded and double-stranded DNA with a similar affinity as the wild-type protein, and when bound to a longer DNA substrate (T30 bubble) it forms two bands. This suggests that nucleotide binding might stabilise the mutant protein. Binding of ΔN256 mutant to single- or double-stranded DNA was not detected. This suggests that the region involved in the ss and ds DNA binding is within the N terminal 1-256 amino acids of the protein. Studies on M*th*MCM revealed that the N-terminus of the protein contains three domains. Domain A is needed for dsDNA translocation and is believed to have a regulatory function, while domain C is necessary and sufficient for MCM multimerisation and helicase activity (Kasiviswanathan *et al.*, 2004). Domain B is the major contact with ssDNA and contains the zinc-binding motif. Sequence alignment of *Mth*MCM and *Ape*MCM indicated that ΔN256 mutant contains only a very small part of the C-domain observed in *Mth*MCM (about 50 amino acids) (Kasiviswanathan *et al.*, 2004). Thus, it is not surprising that the mutant protein did not bind single- or double-stranded DNA.

Next, the ATPase activity of wild type and mutant proteins was investigated using a colorimetric assay, which detects the release of free phosphate (Lanzetta *et al.*, 1979). The wild type protein showed a very low activity under the condition tested, whereas the ΔN256 mutant showed much higher activity. The ATPase activity of the wild type protein was stimulated by the presence of single- and double-stranded DNA. Similar data were obtained for ATP hydrolysis of *Mth*MCM and *Af*MCM (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000; Grainge *et al.*, 2003). In contrast, MCM from the crenarchaeal *S.solfataricus* was found to exhibit a robust ATPase activity, which was not stimulated by the presence of DNA (Carpentieri *et al.*, 2002). This is quite surprising, because the

hydrolysis of ATP (or other nucleotides) by the DNA helicases is usually enhanced if nucleic acids are present (Patel *et al.*, 1992; Washington *et al.*, 1996; Stitt and Xu, 1998; Patel and Picha, 2000). Surprisingly, the ΔN256 mutant showed much higher ATPase activity, which as expected, was not stimulated by the presence of DNA. This suggests that the N-terminal part (residues 1-286) of the protein might conceal the ATP site and this site becomes more accessible when the protein binds DNA or if the N-terminal part is deleted. Similar observations were reported for *A.fulgidus* MCM, where deletion of 111 amino acids of the N-terminus increased the ATPase activity (Grainge *et al.*, 2003). Walker A mutant did not possess any detectable ATPase activity which suggests that nucleotide binding is essential for the ATPase activity of the *Ape*MCM. Similar observations were made with the Walker A mutants of *Mth*MCM (Chong *et al.*, 2000) and *Ss*MCM (Carpentieri *et al.*, 2002). Walker B motif was shown to be important for nucleotide hydrolysis in several enzymes (Pause and Sonenberg, 1992; Story and Steitz, 1992) therefore it was not surprising that the Walker B mutant of *Ape*MCM did not possess an ATPase activity. The arginine finger mutant did not possess any detectable ATPase activity. This suggests that in *Ape*MCM, similarly to the eukaryotic MCM2-7 complex, the ATP binding and/or hydrolysis sites might be located at the subunit interface. One MCM subunit binds ATP and the neighbouring subunit contributes a catalytic arginine residue (Davey *et al.*, 2003). The zinc-binding mutant did not show any detectable ATPase activity in the absence of DNA. The presence of single- and double-stranded DNA, however, stimulated the ATPase activity of the mutant protein. These observations are in contrast to the one made with the zinc-binding mutant of *Mth*MCM (Poplawski *et al.*, 2001). The zinc-binding mutant of *Mth*MCM possesses ATPase activity in the absence of DNA, albeit lower that the wild type (Poplawski *et al.*, 2001). Neither ssDNA nor dsDNA stimulated the ATPases of the mutant protein (Poplawski *et al.*, 2001). Future work in the lab could involve examination of the ATP binding activity of wild type and mutant proteins.

Next, the helicase activity of wild type and mutant *Ape*MCM proteins was examined. Wild type protein showed robust helicase activity on flayed duplex DNA. A variation of MCM helicase activity was observed with this substrate with different concentrations of ATP. Helicase activity was higher when ATP concentrations higher than

5mM were used. The wild type protein was unable to unwind a blunt duplex DNA substrate, which suggests that *Ape*MCM requires a single-stranded tail to unwind DNA. The helicase activities of wild type and mutant proteins were compared using the flayed duplex substrate. Interestingly, the ΔN256 mutant showed a robust helicase activity despite the fact that it did not bind single- or double-stranded DNA. This suggests that the mutant protein might be able to bind DNA but may not form a stable protein-DNA complex, that could be detected by the mobility band shift assay. Walker A, Walker B and arginine mutant proteins did not possess helicase activities. This is not surprising because they did not possess any detectable ATPase activity either. A limited helicase activity was detected with the zinc mutant, which suggests that the zinc-binding site of *Ape*MCM is needed for helicase activity.

Next, the presence of the *Ape*ORC homologues on the helicase activity of the wild type and Δ256 mutant of *Ape*MCM was analysed. Although direct interactions between these proteins were not detected with the yeast two-hybrid system, ORC-1 and ORC-2 inhibited the helicase activity of the wild type protein. Moreover, the winged helix Domain III of both ORC proteins was required for the inhibitory activity. This domain is also needed for interactions with DNA. Several mechanisms of inhibition are possible, including binding of the ORC homologues to MCM and to DNA, thus preventing the helicase from translocating along the DNA, a conformational change in MCM, or destabilising of the MCM interaction with DNA. It is also possible that the two ORC homologues inhibit MCM helicase activity by direct interaction with DNA, thus preventing MCM binding. The requirement for an intact winged-helix motif for efficient inhibition also supports the notion that the inhibition might be due to DNA binding. It is also possible that direct interactions between MCM and the two ORC proteins may be required for the inhibition of helicase activity, although these interactions were not detected by the two-hybrid system. These observations will be similar to the inhibition observed in bacteria in which direct protein-protein interactions between DnaB and DnaC are needed for the inhibition of the helicase activity of the protein (Wahle et al., 1989; Lee and Bell, 2000). Direct interactions between MCM and the ORC homologues were observed in *M.thermoautotrophicum* and *S.solfataricus* (Shin *et al.*, 2003a and b; De Felice *et al.*, 2003). It has been proposed that

the N-terminal portion of MCM interacts with the winged helix domain of ORC-1 and ORC-2 (Kasiviswanathan *et al.*, 2004; Kasiviswanathan *et al.*, 2005). In agreement with this, the helicase activity of the Δ256 mutant was not inhibited neither by ORC-1 nor by ORC-2. Interestingly Domain I+II of ORC-1 was also able to inhibit the helicase activity of the wild type MCM, which suggests that apart from the winged helix domain, there is probably another MCM interaction site within Domain I+II of ORC-1. Future work in the lab could involve isothermal titration calorimetry and analytical ultracentrifuge analyses for a further characterisation of the protein-protein interactions between these archaeal proteins.

REFERENCES

Adams, A. K. and C. Holm (1996). "Specific DNA replication mutations affect telomere length in *Saccharomyces cerevisiae*." *Mol Cell Biol* 16(9): 4614-20.

Aggarwal, B. D. and B. R. Calvi (2004). "Chromatin regulates origin activity in *Drosophila* follicle cells." *Nature* 430(6997): 372-6.

Ahmadian, M.R., Stege, P, Scheffzek, K, Wittinghofer, A. (1997). "Conformation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras". *Nat.Struct.Biol.* 4(9):686-9

Ahnert, P., K. M. Picha and S. S. Patel (2000). "A ring-opening mechanism for DNA binding in the central channel of the T7 helicase-primase protein." *EMBO J* 19(13): 3418-27.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." *J Mol Biol* 215(3): 403-10.

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

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

Austin, R. J., T. L. Orr-Weaver and S. P. Bell (1999). "*Drosophila* ORC specifically binds to ACE3, an origin of DNA replication control element." *Genes Dev* 13(20): 2639-49.

Bahloul, A., F. Boubrik and J. Rouviere-Yaniv (2001). "Roles of Escherichia coli histone-like protein HU in DNA replication: HU-beta suppresses the thermosensitivity of DNAA 46ts." *Biochimie* 83(2): 219-29.

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

Baker, T. A. and S. H. Wickner (1992). "Genetics and enzymology of DNA replication in Escherichia coli." *Annu Rev Genet* 26: 447-77.

Barns, S. M., C. F. Delwiche, J. D. Palmer, S. C. Dawson, K. L. Hershberger and N. R. Pace (1996). "Phylogenetic perspective on microbial life in hydrothermal ecosystems, past and present." *Ciba Found Symp* 202: 24-32.

Beall, E. L., J. R. Manak, S. Zhou, M. Bell, J. S. Lipsick and M. R. Botchan (2002). "Role for a *Drosophila* Myb-containing protein complex in site-specific DNA replication." *Nature* 420(6917): 833-7.

Bell, S. D. and S. P. Jackson (2001). "Mechanism and regulation of transcription in archaea." *Curr Opin Microbiol* 4(2): 208-13.

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

Bell, S. P. and B. Stillman (1992). "ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex." *Nature* 357(6374): 128-34.

Bendixen, C., S. Gangloff and R. Rothstein (1994). "A yeast mating-selection scheme for detection of protein-protein interactions." *Nucleic Acids Res* 22(9): 1778-9.

Berquist, B. R. and S. DasSarma (2003). "An archaeal chromosomal autonomously replicating sequence element from an extreme halophile, Halobacterium sp. strain NRC-1." *J Bacteriol* 185(20): 5959-66.

Bielinsky, A. K., H. Blitzblau, E. L. Beall, M. Ezrokhi, H. S. Smith, M. R. Botchan and S. A. Gerbi (2001). "Origin recognition complex binding to a metazoan replication origin." *Curr Biol* 11(18): 1427-31.

Biswas, E. E. and S. B. Biswas (1999). "Mechanism of DNA binding by the DnaB helicase of *Escherichia coli*: analysis of the roles of domain gamma in DNA binding." *Biochemistry* 38(34): 10929-39.

Biswas, S. B., P. H. Chen and E. E. Biswas (1994). "Structure and function of *Escherichia coli* DnaB protein: role of the N-terminal domain in helicase activity." *Biochemistry* 33(37): 11307-14.

Blum, S., S. R. Schmid, A. Pause, P. Buser, P. Linder, N. Sonenberg and H. Trachsel (1992). "ATP hydrolysis by initiation factor 4A is required for translation initiation in *Saccharomyces cerevisiae*." *Proc Natl Acad Sci U S A* 89(16): 7664-8.

Bohlke, K., F. M. Pisani, M. Rossi and G. Antranikian (2002). "Archaeal DNA replication: spotlight on a rapidly moving field." *Extremophiles* 6(1): 1-14.

Bowers, J. L., J. C. Randell, S. Chen and S. P. Bell (2004). "ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication." *Mol Cell* 16(6): 967-78.

Boye, E., A. Lobner-Olesen and K. Skarstad (2000). "Limiting DNA replication to once and only once." *EMBO Rep* 1(6): 479-83.

Bramhill, D. and A. Kornberg (1988). "A model for initiation at origins of DNA replication." *Cell* 54(7): 915-8.

Brent, R. and M. Ptashne (1985). "A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor." *Cell* 43(3 Pt 2): 729-36.

Brosh, R. M., Jr. and S. W. Matson (1995). "Mutations in motif II of Escherichia coli DNA helicase II render the enzyme nonfunctional in both mismatch repair and excision repair with differential effects on the unwinding reaction." *J Bacteriol* 177(19): 5612-21.

Brown, G. W., P. V. Jallepalli, B. J. Huneycutt and T. J. Kelly (1997). "Interaction of the S phase regulator cdc18 with cyclin-dependent kinase in fission yeast." *Proc Natl Acad Sci U S A* 94(12): 6142-7.

Brown, G. W. and T. J. Kelly (1998). "Purification of Hsk1, a minichromosome maintenance protein kinase from fission yeast." *J Biol Chem* 273(34): 22083-90.

Brown, G. W. and T. J. Kelly (1999). "Cell cycle regulation of Dfp1, an activator of the Hsk1 protein kinase." *Proc Natl Acad Sci U S A* 96(15): 8443-8.

Bujalowski, W. and M. J. Jezewska (1995). "Interactions of *Escherichia coli* primary replicative helicase DnaB protein with single-stranded DNA. The nucleic acid does not wrap around the protein hexamer." *Biochemistry* 34(27): 8513-9.

Bujalowski, W. and M. J. Jezewska (2000). "Kinetic mechanism of nucleotide cofactor binding to Escherichia coli replicative helicase DnaB protein. stopped-flow kinetic studies using fluorescent, ribose-, and base-modified nucleotide analogues." Biochemistry 39(8): 2106-22.

Bujalowski, W. and M. J. Jezewska (2000). "Kinetic mechanism of the single-stranded DNA recognition by *Escherichia coli* replicative helicase DnaB protein. Application of the matrix projection operator technique to analyze stopped-flow kinetics." *J Mol Biol* 295(4): 831-52.

Bujalowski, W. and M. M. Klonowska (1994). "Close proximity of tryptophan residues and ATP-binding site in Escherichia coli primary replicative helicase DnaB protein. Molecular topography of the enzyme." *J Biol Chem* 269(50): 31359-71.

Bujalowski, W. and M. M. Klonowska (1994). "Structural characteristics of the nucleotide-binding site of Escherichia coli primary replicative helicase DnaB protein. Studies with ribose and base-modified fluorescent nucleotide analogs." *Biochemistry* 33(15): 4682-94.

Bujalowski, W., M. M. Klonowska and M. J. Jezewska (1994). "Oligomeric structure of *Escherichia coli* primary replicative helicase DnaB protein." *J Biol Chem* 269(50): 31350-8.

Bullock, P. A. and Y. P. Myer (1978). "Circular dichroism and resonance Raman studies of cytochrome b562 from *Escherichia coli*." *Biochemistry* 17(15): 3084-91.

Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. Geoghagen and J. C. Venter (1996). "Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*." *Science* 273(5278): 1058-73.

Campbell, J. L. and N. Kleckner (1990). "*E. coli* oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork." *Cell* 62(5): 967-79.

Capaldi, S. A. and J. M. Berger (2004). "Biochemical characterization of Cdc6/Orc1 binding to the replication origin of the euryarchaeon *Methanothermobacter thermoautotrophicus*." *Nucleic Acids Res* 32(16): 4821-32.

Carpenter, P. B., P. R. Mueller and W. G. Dunphy (1996). "Role for a *Xenopus* Orc2-related protein in controlling DNA replication." *Nature* 379(6563): 357-60.

Carpentieri, F., M. De Felice, M. De Falco, M. Rossi and F. M. Pisani (2002). "Physical and functional interaction between the mini-chromosome maintenance-like DNA helicase and the single-stranded DNA binding protein from the crenarchaeon *Sulfolobus solfataricus*." *J Biol Chem* 277(14): 12118-27.

Cassler, M. R., J. E. Grimwade and A. C. Leonard (1995). "Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin." *EMBO J* 14(23): 5833-41.

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

Cheng, L., T. Collyer and C. F. Hardy (1999). "Cell cycle regulation of DNA replication initiator factor Dbf4p." *Mol Cell Biol* 19(6): 4270-8.

Chesnokov, I., M. Gossen, D. Remus and M. Botchan (1999). "Assembly of functionally active *Drosophila* origin recognition complex from recombinant proteins." *Genes Dev* 13(10): 1289-96.

Chesnokov, I., D. Remus and M. Botchan (2001). "Functional analysis of mutant and wild-type *Drosophila* origin recognition complex." *Proc Natl Acad Sci U S A* 98(21): 11997-2002.

Chien, C. T., P. L. Bartel, R. Sternglanz and S. Fields (1991). "The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest." *Proc Natl Acad Sci U S A* 88(21): 9578-82.

Chong, J. P., M. K. Hayashi, M. N. Simon, R. M. Xu and B. Stillman (2000). "A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase." *Proc Natl Acad Sci U S A* 97(4): 1530-5.

Chong, J. P., H. M. Mahbubani, C. Y. Khoo and J. J. Blow (1995). "Purification of an MCM-containing complex as a component of the DNA replication licensing system." *Nature* 375(6530): 418-21.

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

Clarke, A. R. a. Engel., P.C. (1996). "Enzymology Labfax." BIOS Scientific publishers and academic press Chapter 6: 203-204.

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

Coleman, T. R., P. B. Carpenter and W. G. Dunphy (1996). "The *Xenopus* Cdc6 protein is essential for the initiation of a single round of DNA replication in cell-free extracts." *Cell* 87(1): 53-63.

Coue, M., F. Amariglio, D. Maiorano, S. Bocquet and M. Mechali (1998). "Evidence for different MCM subcomplexes with differential binding to chromatin in *Xenopus*." *Exp Cell Res* 245(2): 282-9.

Coue, M., S. E. Kearsey and M. Mechali (1996). "Chromotin binding, nuclear localization and phosphorylation of *Xenopus* cdc21 are cell-cycle dependent and associated with the control of initiation of DNA replication." *EMBO J* 15(5): 1085-97.

Cravchik, A. and A. Matus (1993). "A novel strategy for the immunological tagging of cDNA constructs." *Gene* 137(1): 139-43.

Cunningham, E. L. and J. M. Berger (2005). "Unraveling the early steps of prokaryotic replication." *Curr Opin Struct Biol* 15(1): 68-76.

Dalton, S. and B. Hopwood (1997). "Characterization of Cdc47p-minichromosome maintenance complexes in *Saccharomyces cerevisiae*: identification of Cdc45p as a subunit." *Mol Cell Biol* 17(10): 5867-75.

Davey, M. J., L. Fang, P. McInerney, R. E. Georgescu and M. O'Donnell (2002). "The DnaC helicase loader is a dual ATP/ADP switch protein." *EMBO J* 21(12): 3148-59.

Davey, M. J., D. Jeruzalmi, J. Kuriyan and M. O'Donnell (2002). "Motors and switches: AAA+ machines within the replisome." *Nat Rev Mol Cell Biol* 3(11): 826-35.

De Felice, M., L. Esposito, B. Pucci, F. Carpentieri, M. De Falco, M. Rossi and F. M. Pisani (2003). "Biochemical characterization of a CDC6-like protein from the crenarchaeon *Sulfolobus solfataricus*." *J Biol Chem* 278(47): 46424-31.

De Felice, M., L. Esposito, B. Pucci, M. De Falco, G. Manco, M. Rossi and F. M. Pisani (2004). "Modular organization of a Cdc6-like protein from the crenarchaeon *Sulfolobus solfataricus*." *Biochem J* 381(Pt 3): 645-53.

De Felice, M., L. Esposito, B. Pucci, M. De Falco, M. Rossi and F. M. Pisani (2004). "A CDC6-like factor from the archaea *Sulfolobus solfataricus* promotes binding of the mini-chromosome maintenance complex to DNA." *J Biol Chem* 279(41): 43008-12.

DePamphilis, M. L. (1999). "Replication origins in metazoan chromosomes: fact or fiction?" *Bioessays* 21(1): 5-16.

DePamphilis, M. L. (2003). "Eukaryotic DNA replication origins: reconciling disparate data." *Cell* 114(3): 274-5.

DePamphilis, M. L. (2003). "The 'ORC cycle': a novel pathway for regulating eukaryotic DNA replication." *Gene* 310: 1-15.

Dhar, S. K., L. Delmolino and A. Dutta (2001). "Architecture of the human origin recognition complex." *J Biol Chem* 276(31): 29067-71.

di Guan, C., P. Li, P. D. Riggs and H. Inouye (1988). "Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein." *Gene* 67(1): 21-30.

Donovan, S., J. Harwood, L. S. Drury and J. F. Diffley (1997). "Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast." *Proc Natl Acad Sci U S A* 94(11): 5611-6.

Doran, K. S., D. R. Helinski and I. Konieczny (1999). "Host-dependent requirement for specific DnaA boxes for plasmid RK2 replication." *Mol Microbiol* 33(3): 490-8.

Dubey, D. D., S. M. Kim, I. T. Todorov and J. A. Huberman (1996). "Large, complex modular structure of a fission yeast DNA replication origin." *Curr Biol* 6(4): 467-73.

Dubey, D. D., J. Zhu, D. L. Carlson, K. Sharma and J. A. Huberman (1994). "Three ARS elements contribute to the ura4 replication origin region in the fission yeast, Schizosaccharomyces pombe." *EMBO J* 13(15): 3638-47.

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

Edgell, D. R. and W. F. Doolittle (1997). "Archaea and the origin(s) of DNA replication proteins." *Cell* 89(7): 995-8.

Elsasser, S., F. Lou, B. Wang, J. L. Campbell and A. Jong (1996). "Interaction between yeast Cdc6 protein and B-type cyclin/Cdc28 kinases." *Mol Biol Cell* 7(11): 1723-35.

Erzberger, J. P., M. M. Pirruccello and J. M. Berger (2002). "The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation." *EMBO J* 21(18): 4763-73.

Espejo, A., J. Cote, A. Bednarek, S. Richard and M. T. Bedford (2002). "A protein-domain microarray identifies novel protein-protein interactions." *Biochem J* 367(Pt 3): 697-702.

Felczak, M. M. and J. M. Kaguni (2004). "The box VII motif of Escherichia coli DnaA protein is required for DnaA oligomerization at the *E. coli* replication origin." *J Biol Chem* 279(49): 51156-62.

Felczak, M. M., L. A. Simmons and J. M. Kaguni (2005). "An essential tryptophan of *Escherichia coli* DnaA protein functions in oligomerization at the *E. coli* replication origin." *J Biol Chem* 280(26): 24627-33.

Ferreira, M. F., C. Santocanale, L. S. Drury and J. F. Diffley (2000). "Dbf4p, an essential S phase-promoting factor, is targeted for degradation by the anaphase-promoting complex." *Mol Cell Biol* 20(1): 242-8.

Fields, S. and O. Song (1989). "A novel genetic system to detect protein-protein interactions." *Nature* 340(6230): 245-6.

Fields, S. and R. Sternglanz (1994). "The two-hybrid system: an assay for protein-protein interactions." *Trends Genet* 10(8): 286-92.

Filutowicz, M., W. Ross, J. Wild and R. L. Gourse (1992). "Involvement of Fis protein in replication of the *Escherichia coli* chromosome." *J Bacteriol* 174(2): 398-407.

Findeisen, M., M. El-Denary, T. Kapitza, R. Graf and U. Strausfeld (1999). "Cyclin A-dependent kinase activity affects chromatin binding of ORC, Cdc6, and MCM in egg extracts of *Xenopus* laevis." *Eur J Biochem* 264(2): 415-26.

Finkel, S. E. and R. C. Johnson (1992). "The Fis protein: it's not just for DNA inversion anymore." *Mol Microbiol* 6(22): 3257-65.

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

Forsburg, S. L. (2004). "Eukaryotic MCM proteins: beyond replication initiation." *Microbiol Mol Biol Rev* 68(1): 109-31

.

Forsburg, S. L., D. A. Sherman, S. Ottilie, J. R. Yasuda and J. A. Hodson (1997). "Mutational analysis of Cdc19p, a *Schizosaccharomyces pombe* MCM protein." *Genetics* 147(3): 1025-41.

Forterre, P. (2002). "The origin of DNA genomes and DNA replication proteins." *Curr Opin Microbiol* 5(5): 525-32.

Forterre, P., C. Brochier and H. Philippe (2002). "Evolution of the Archaea." *Theor Popul Biol* 61(4): 409-22.

Fujikawa, N., H. Kurumizaka, O. Nureki, T. Terada, M. Shirouzu, T. Katayama and S. Yokoyama (2003). "Structural basis of replication origin recognition by the DnaA protein." *Nucleic Acids Res* 31(8): 2077-86.

Fujita, M. (1999). "Cell cycle regulation of DNA replication initiation proteins in mammalian cells." *Front Biosci* 4: D816-23.

Fuller, R. S., B. E. Funnell and A. Kornberg (1984). "The DnaA protein complex with the *E. coli* chromosomal replication origin (oriC) and other DNA sites." *Cell* 38(3): 889-900.

Gajiwala, K. S. and S. K. Burley (2000). "Winged helix proteins." *Curr Opin Struct Biol* 10(1): 110-6.

Gajiwala, K. S., H. Chen, F. Cornille, B. P. Roques, W. Reith, B. Mach and S. K. Burley (2000). "Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding." *Nature* 403(6772): 916-21.

Garcia de Viedma, D., R. Giraldo, G. Rivas, E. Fernandez-Tresguerres and R. Diaz-Orejas (1996). "A leucine zipper motif determines different functions in a DNA replication protein." *EMBO J* 15(4): 925-34.

Garcia de Viedma, D., R. Giraldo, M. J. Ruiz-Echevarria, R. Lurz and R. Diaz-Orejas (1995). "Transcription of repA, the gene of the initiation protein of the *Pseudomonas*

plasmid pPS10, is autoregulated by interactions of the RepA protein at a symmetrical operator." *J Mol Biol* 247(2): 211-23.

Garner, J. and E. Crooke (1996). "Membrane regulation of the chromosomal replication activity of *E.coli* DnaA requires a discrete site on the protein." *EMBO J* 15(9): 2313-21.

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

Giraldo, R. and R. Diaz-Orejas (2001). "Similarities between the DNA replication initiators of Gram-negative bacteria plasmids (RepA) and eukaryotes (Orc4p)/archaea (Cdc6p)." *Proc Natl Acad Sci U S A* 98(9): 4938-43.

Grabowski, B. and Z. Kelman (2001). "Autophosphorylation of archaeal Cdc6 homologues is regulated by DNA." *J Bacteriol* 183(18): 5459-64.

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

Grainge, I., S. Scaife and D. B. Wigley (2003). "Biochemical analysis of components of the pre-replication complex of *Archaeoglobus fulgidus*." *Nucleic Acids Res* 31(16): 4888-98.

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

Grimwade, J. E., V. T. Ryan and A. C. Leonard (2000). "IHF redistributes bound initiator protein, DnaA, on supercoiled oriC of *Escherichia coli*." *Mol Microbiol* 35(4): 835-44.

Guenther, B., R. Onrust, A. Sali, M. O'Donnell and J. Kuriyan (1997). "Crystal structure of the delta' subunit of the clamp-loader complex of *E. coli* DNA polymerase III." *Cell* 91(3): 335-45.

Hamlin, J. L. and P. A. Dijkwel (1995). "On the nature of replication origins in higher eukaryotes." *Curr Opin Genet Dev* 5(2): 153-61.

Harper, J. W., G. R. Adami, N. Wei, K. Keyomarsi and S. J. Elledge (1993). "The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases." *Cell* 75(4): 805-16.

Hartwell, L. H. (1973). "Three additional genes required for deoxyribonucleic acid synthesis in *Saccharomyces cerevisiae*." *J Bacteriol* 115(3): 966-74.

Harvey, K. J. and J. Newport (2003). "Metazoan origin selection: origin recognition complex chromatin binding is regulated by CDC6 recruitment and ATP hydrolysis." *J Biol Chem* 278(49): 48524-8.

Hase, M., T. Yoshimi, Y. Ishikawa, A. Ohba, L. Guo, S. Mima, M. Makise, Y. Yamaguchi, T. Tsuchiya and T. Mizushima (1998). "Site-directed mutational analysis for the membrane binding of DnaA protein. Identification of amino acids involved in the functional interaction between DnaA protein and acidic phospholipids." *J Biol Chem* 273(44): 28651-6.

Hendrickson, M., M. Madine, S. Dalton and J. Gautier (1996). "Phosphorylation of MCM4 by cdc2 protein kinase inhibits the activity of the minichromosome maintenance complex." *Proc Natl Acad Sci U S A* 93(22): 12223-8.

Herbig, U., J. W. Griffith and E. Fanning (2000). "Mutation of cyclin/cdk phosphorylation sites in HsCdc6 disrupts a late step in initiation of DNA replication in human cells." *Mol Biol Cell* 11(12): 4117-30.

Herbig, U., C. A. Marlar and E. Fanning (1999). "The Cdc6 nucleotide-binding site regulates its activity in DNA replication in human cells." *Mol Biol Cell* 10(8): 2631-45.

Hiasa, H. and K. J. Marians (1994). "Fis cannot support oriC DNA replication in vitro." *J Biol Chem* 269(40): 24999-5003.

Higgins, D. G. (1994). "CLUSTAL V: multiple alignment of DNA and protein sequences." *Methods Mol Biol* 25: 307-18.

Hingorani, M. M. and M. O'Donnell (2000). "A tale of toroids in DNA metabolism." *Nat Rev Mol Cell Biol* 1(1): 22-30.

Ho, Y., A. Gruhler, A. Heilbut, G. D. Bader, L. Moore, S. L. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier, L. Yang, C. Wolting, I. Donaldson, S. Schandorff, J. Shewnarane, M. Vo, J. Taggart, M. Goudreault, B. Muskat, C. Alfarano, D. Dewar, Z. Lin, K. Michalickova, A. R. Willems, H. Sassi, P. A. Nielsen, K. J. Rasmussen, J. R. Andersen, L. E. Johansen, L. H. Hansen, H. Jespersen, A. Podtelejnikov, E. Nielsen, J. Crawford, V. Poulsen, B. D. Sorensen, J. Matthiesen, R. C. Hendrickson, F. Gleeson, T. Pawson, M. F. Moran, D. Durocher, M. Mann, C. W. Hogue, D. Figeys and M. Tyers (2002). "Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry." *Nature* 415(6868): 180-3.

Hochuli, E., H. Dobeli and A. Schacher (1987). "New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues." *J Chromatogr* 411: 177-84.

Hofmann, J. F. and D. Beach (1994). "Cdt1 is an essential target of the Cdc10/Sct1 transcription factor: requirement for DNA replication and inhibition of mitosis." *EMBO J* 13(2): 425-34.

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

Hopwood, B. and S. Dalton (1996). "Cdc45p assembles into a complex with Cdc46p/Mcm5p, is required for minichromosome maintenance, and is essential for chromosomal DNA replication." *Proc Natl Acad Sci U S A* 93(22): 12309-14.

Hua, X. H. and J. Newport (1998). "Identification of a pre-initiation step in DNA replication that is independent of origin recognition complex and Cdc6, but dependent on Cdk2." *J Cell Biol* 140(2): 271-81.

Huber, H., M. J. Hohn, R. Rachel, T. Fuchs, V. C. Wimmer and K. O. Stetter (2002). "A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont." *Nature* 417(6884): 63-7.

Huber, H., M. J. Hohn, K. O. Stetter and R. Rachel (2003). "The phylum Nanoarchaeota: present knowledge and future perspectives of a unique form of life." *Res Microbiol* 154(3): 165-71.

Husi, H., M. A. Ward, J. S. Choudhary, W. P. Blackstock and S. G. Grant (2000). "Proteomic analysis of NMDA receptor-adhesion protein signaling complexes." *Nat Neurosci* 3(7): 661-9.

Hwang, D. S. and A. Kornberg (1992). "Opening of the replication origin of *Escherichia coli* by DnaA protein with protein HU or IHF." *J Biol Chem* 267(32): 23083-6.

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

Ishimi, Y. and Y. Komamura-Kohno (2001). "Phosphorylation of Mcm4 at specific sites by cyclin-dependent kinase leads to loss of Mcm4,6,7 helicase activity." *J Biol Chem* 276(37): 34428-33.

Ishimi, Y., Y. Komamura-Kohno, Z. You, A. Omori and M. Kitagawa (2000). "Inhibition of Mcm4,6,7 helicase activity by phosphorylation with cyclin A/Cdk2." *J Biol Chem* 275(21): 16235-41.

Izumi, M., K. Yanagi, T. Mizuno, M. Yokoi, Y. Kawasaki, K. Y. Moon, J. Hurwitz, F. Yatagai and F. Hanaoka (2000). "The human homolog of *Saccharomyces cerevisiae* Mcm10 interacts with replication factors and dissociates from nuclease-resistant nuclear structures in G (2) phase." *Nucleic Acids Res* 28(23): 4769-77.

Jacob, F., Brenner, S. and Cuzin, F. (1964). "On the regulation of DNA replication in bacteria." *Cold Spring Harb Symp Quant Biol* 28: 329-347.

Jain, R., M. C. Rivera and J. A. Lake (1999). "Horizontal gene transfer among genomes: the complexity hypothesis." *Proc Natl Acad Sci U S A* 96(7): 3801-6.

Jares, P. and J. J. Blow (2000). "*Xenopus* Cdc7 function is dependent on licensing but not on XOrc, XCdc6, or Cdk activity and is required for XCdc45 loading." *Genes Dev* 14(12): 1528-40.

Jeruzalmi, D. (2004). "Chromosomal DNA replication on a protein "chip"." *Structure* (Camb) 12(12): 2100-2.

Jezewska, M. J. and W. Bujalowski (1996). "Global conformational transitions in *Escherichia coli* primary replicative helicase DnaB protein induced by ATP, ADP, and single-stranded DNA binding. Multiple conformational states of the helicase hexamer." *J Biol Chem* 271(8): 4261-5.

Jezewska, M. J. and W. Bujalowski (1996). "A general method of analysis of ligand binding to competing macromolecules using the spectroscopic signal originating from a reference macromolecule. Application to *Escherichia coli* replicative helicase DnaB protein nucleic acid interactions." *Biochemistry* 35(7): 2117-28.

Jezewska, M. J. and W. Bujalowski (2000). "Interactions of *Escherichia coli* replicative helicase PriA protein with single-stranded DNA." *Biochemistry* 39(34): 10454-67.

Jezewska, M. J., U. S. Kim and W. Bujalowski (1996). "Interactions of *Escherichia coli* primary replicative helicase DnaB protein with nucleotide cofactors." *Biophys* J 71(4): 2075-86.

Jezewska, M. J., U. S. Kim and W. Bujalowski (1996). "Binding of *Escherichia coli* primary replicative helicase DnaB protein to single-stranded DNA. Long-range allosteric conformational changes within the protein hexamer." *Biochemistry* 35(7): 2129-45.

Jezewska, M. J., S. Rajendran, D. Bujalowska and W. Bujalowski (1998). "Does single-stranded DNA pass through the inner channel of the protein hexamer in the complex with the *Escherichia coli* DnaB Helicase? Fluorescence energy transfer studies." *J Biol Chem* 273(17): 10515-29.

Jezewska, M. J., S. Rajendran and W. Bujalowski (1998). "Complex of *Escherichia coli* primary replicative helicase DnaB protein with a replication fork: recognition and structure." *Biochemistry* 37(9): 3116-36.

Jiang, W. and T. Hunter (1997). "Identification and characterization of a human protein kinase related to budding yeast Cdc7p." *Proc Natl Acad Sci U S A* 94(26): 14320-5.

Jiang, W., D. McDonald, T. J. Hope and T. Hunter (1999). "Mammalian Cdc7-Dbf4 protein kinase complex is essential for initiation of DNA replication." *EMBO J* 18(20): 5703-13.

Jiang, W., N. J. Wells and T. Hunter (1999). "Multistep regulation of DNA replication by Cdk phosphorylation of *Hs*Cdc6." *Proc Natl Acad Sci U S A* 96(11): 6193-8.

Kaguni, J. M. (1997). "*Escherichia coli* DnaA protein: the replication initiator." *Mol Cells* 7(2): 145-57.

Kaplan, D. L. (2000). "The 3'-tail of a forked-duplex sterically determines whether one or two DNA strands pass through the central channel of a replication-fork helicase." *J Mol Biol* 301(2): 285-99.

Kaplan, D. L., M. J. Davey and M. O'Donnell (2003). "Mcm4,6,7 uses a "pump in ring" mechanism to unwind DNA by steric exclusion and actively translocate along a duplex." *J Biol Chem* 278(49): 49171-82.

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

Kasiviswanathan, R., J. H. Shin, E. Melamud and Z. Kelman (2004). "Biochemical characterization of the *Methanothermobacter thermautotrophicus* minichromosome maintenance (MCM) helicase N-terminal domains." *J Biol Chem* 279(27): 28358-66.

Katayama, T., T. Kubota, K. Kurokawa, E. Crooke and K. Sekimizu (1998). "The initiator function of DnaA protein is negatively regulated by the sliding clamp of the *E. coli* chromosomal replicase." *Cell* 94(1): 61-71.

Kato, J. and T. Katayama (2001). "Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*." *EMBO J* 20(15): 4253-62.

Kawarabayasi, Y., Y. Hino, H. Horikawa, S. Yamazaki, Y. Haikawa, K. Jin-no, M. Takahashi, M. Sekine, S. Baba, A. Ankai, H. Kosugi, A. Hosoyama, S. Fukui, Y. Nagai, K. Nishijima, H. Nakazawa, M. Takamiya, S. Masuda, T. Funahashi, T. Tanaka, Y. Kudoh, J. Yamazaki, N. Kushida, A. Oguchi, H. Kikuchi and et al. (1999). "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1." *DNA Res* 6(2): 83-101, 145-52.

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

Keegan, L., G. Gill and M. Ptashne (1986). "Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein." *Science* 231(4739): 699-704.

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

Kelman, L. M. and Z. Kelman (2003). "Archaea: an archetype for replication initiation studies?" *Mol Microbiol* 48(3): 605-15.

Kelman, L. M. and Z. Kelman (2004). "Multiple origins of replication in archaea." *Trends Microbiol* 12(9): 399-401.

Kelman, Z. (2000). "DNA replication in the third domain (of life)." *Curr Protein Pept Sci* 1(2): 139-54.

Kelman, Z. (2000). "The replication origin of archaea is finally revealed." *Trends Biochem Sci* 25(11): 521-3.

Kelman, Z. and J. Hurwitz (2003). "Structural lessons in DNA replication from the third domain of life." *Nat Struct Biol* 10(3): 148-50.

Kelman, Z., J. K. Lee and J. Hurwitz (1999). "The single minichromosome maintenance protein of *Methanobacterium thermoautotrophicum* DeltaH contains DNA helicase activity." *Proc Natl Acad Sci U S A* 96(26): 14783-8.

Kihara, M., W. Nakai, S. Asano, A. Suzuki, K. Kitada, Y. Kawasaki, L. H. Johnston and A. Sugino (2000). "Characterization of the yeast Cdc7p/Dbf4p complex purified from insect cells. Its protein kinase activity is regulated by Rad53p." *J Biol Chem* 275(45): 35051-62.

Kim, J. M., M. Yamada and H. Masai (2003). "Functions of mammalian Cdc7 kinase in initiation/monitoring of DNA replication and development." *Mutat Res* 532(1-2): 29-40.

Kitagawa, R., H. Mitsuki, T. Okazaki and T. Ogawa (1996). "A novel DnaA protein-binding site at 94.7 min on the *Escherichia coli* chromosome." *Mol Microbiol* 19(5): 1137-47.

Kitagawa, R., T. Ozaki, S. Moriya and T. Ogawa (1998). "Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for *Escherichia coli* DnaA protein." *Genes Dev* 12(19): 3032-43.

Klemm, R. D., R. J. Austin and S. P. Bell (1997). "Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex." *Cell* 88(4): 493-502.

Klemm, R. D. and S. P. Bell (2001). "ATP bound to the origin recognition complex is important for preRC formation." *Proc Natl Acad Sci U S A* 98(15): 8361-7.

Komori, H., F. Matsunaga, Y. Higuchi, M. Ishiai, C. Wada and K. Miki (1999). "Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 A resolution." *EMBO J* 18(17): 4597-607.

Kong, D. and M. L. DePamphilis (2001). "Site-specific DNA binding of the *Schizosaccharomyces pombe* origin recognition complex is determined by the Orc4 subunit." *Mol Cell Biol* 21(23): 8095-103.

Konieczny, I. and D. R. Helinski (1997). "Helicase delivery and activation by DnaA and TrfA proteins during the initiation of replication of the broad host range plasmid RK2." *J Biol Chem* 272(52): 33312-8.

Konopa, G., A. Szalewska-Palasz, A. Schmidt, S. Srutkowska, W. Messer and G. Wegrzyn (1999). "The presence of two DnaA-binding sequences is required for an efficient interaction of the *Escherichia coli* DnaA protein with each particular weak DnaA box region." *FEMS Microbiol Lett* 174(1): 25-31.

Koonin, E. V. (1993). "A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif." *J Mol Biol* 229(4): 1165-74.

Koonin, E. V. (1993). "A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication." *Nucleic Acids Res* 21(11): 2541-7.

Koonin, E. V., A. R. Mushegian, M. Y. Galperin and D. R. Walker (1997). "Comparison of archaeal and bacterial genomes: computer analysis of protein sequences predicts novel functions and suggests a chimeric origin for the archaea." *Mol Microbiol* 25(4): 619-37.

Kornberg, A. a. B., T.A. (1992). "DNA replication." W.H. Freeman and Co., New York, NY.

Krause, M. and W. Messer (1999). "DnaA proteins of Escherichia coli and *Bacillus subtilis*: coordinate actions with single-stranded DNA-binding protein and interspecies inhibition during open complex formation at the replication origins." *Gene* 228(1-2): 123-32.

Kreitz, S., M. Ritzi, M. Baack and R. Knippers (2001). "The human origin recognition complex protein 1 dissociates from chromatin during S phase in HeLa cells." *J Biol Chem* 276(9): 6337-42.

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

Kumagai, H., N. Sato, M. Yamada, D. Mahony, W. Seghezzi, E. Lees, K. Arai and H. Masai (1999). "A novel growth- and cell cycle-regulated protein, ASK, activates human Cdc7-related kinase and is essential for G1/S transition in mammalian cells." *Mol Cell Biol* 19(7): 5083-95.

Kurokawa, K., T. Mizushima, T. Kubota, T. Tsuchiya, T. Katayama and K. Sekimizu (1998). "A stimulation factor for hydrolysis of ATP bound to DnaA protein, the initiator of chromosomal DNA replication in *Escherichia coli*." *Biochem Biophys Res Commun* 243(1): 90-5.

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

Labib, K., S. E. Kearsey and J. F. Diffley (2001). "MCM2-7 proteins are essential components of prereplicative complexes that accumulate cooperatively in the nucleus during G1-phase and are required to establish, but not maintain, the S-phase checkpoint." *Mol Biol Cell* 12(11): 3658-67.

Lanzetta, P. A., L. J. Alvarez, P. S. Reinach and O. A. Candia (1979). "An improved assay for nanomole amounts of inorganic phosphate." *Anal Biochem* 100(1): 95-7.

Laskey, R. A. and M. A. Madine (2003). "A rotary pumping model for helicase function of MCM proteins at a distance from replication forks." *EMBO Rep* 4(1): 26-30.

LeBowitz, J. H. and R. McMacken (1986). "The *Escherichia coli* dnaB replication protein is a DNA helicase." *J Biol Chem* 261(10): 4738-48.

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

Lee, D. G. and S. P. Bell (1997). "Architecture of the yeast origin recognition complex bound to origins of DNA replication." *Mol Cell Biol* 17(12): 7159-68.

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

Lee, D. G., A. M. Makhov, R. D. Klemm, J. D. Griffith and S. P. Bell (2000). "Regulation of origin recognition complex conformation and ATPase activity: differential effects of single-stranded and double-stranded DNA binding." *EMBO J* 19(17): 4774-82.

Lee, J. K. and J. Hurwitz (2001). "Processive DNA helicase activity of the minichromosome maintenance proteins 4, 6, and 7 complex requires forked DNA structures." *Proc Natl Acad Sci U S A* 98(1): 54-9.

Lee, J. K., K. Y. Moon, Y. Jiang and J. Hurwitz (2001). "The *Schizosaccharomyces pombe* origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by means of the AT-hook domains of the spOrc4 protein." *Proc Natl Acad Sci U S A* 98(24): 13589-94.

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

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

Leipe, D. D., L. Aravind and E. V. Koonin (1999). "Did DNA replication evolve twice independently?" *Nucleic Acids Res* 27(17): 3389-401.

Li, J. J. and I. Herskowitz (1993). "Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system." *Science* 262(5141): 1870-4.

Liu, J., C. L. Smith, D. DeRyckere, K. DeAngelis, G. S. Martin and J. M. Berger (2000). "Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control." *Mol Cell* 6(3): 637-48.

Lobner-Olesen, A., F. G. Hansen, K. V. Rasmussen, B. Martin and P. L. Kuempel (1994). "The initiation cascade for chromosome replication in wild-type and Dam methyltransferase deficient *Escherichia coli* cells." *EMBO J* 13(8): 1856-62.

Lobry, J. R. (1996). "Asymmetric substitution patterns in the two DNA strands of bacteria." *Mol Biol Evol* 13(5): 660-5.

Lohman, T. M. and W. Bujalowski (1994). "Effects of base composition on the negative cooperativity and binding mode transitions of *Escherichia coli* SSB-single-stranded DNA complexes." *Biochemistry* 33(20): 6167-76.

Lopez, P., H. Philippe, H. Myllykallio and P. Forterre (1999). "Identification of putative chromosomal origins of replication in Archaea." *Mol Microbiol* 32(4): 883-6.

Lopez-Girona, A., O. Mondesert, J. Leatherwood and P. Russell (1998). "Negative regulation of Cdc18 DNA replication protein by Cdc2." *Mol Biol Cell* 9(1): 63-73.

Lu, M., J. L. Campbell, E. Boye and N. Kleckner (1994). "SeqA: a negative modulator of replication initiation in *E. coli*." *Cell* 77(3): 413-26.

Lundgren, M., A. Andersson, L. Chen, P. Nilsson and R. Bernander (2004). "Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination." *Proc Natl Acad Sci U S A* 101(18): 7046-51.

Mackiewicz, P., J. Zakrzewska-Czerwinska, A. Zawilak, M. R. Dudek and S. Cebrat (2004). "Where does bacterial replication start? Rules for predicting the oriC region." *Nucleic Acids Res* 32(13): 3781-91.

MacNeill, S. A. (2001). "Understanding the enzymology of archaeal DNA replication: progress in form and function." *Mol Microbiol* 40(3): 520-9.

Madine, M. A., M. Swietlik, C. Pelizon, P. Romanowski, A. D. Mills and R. A. Laskey (2000). "The roles of the MCM, ORC, and Cdc6 proteins in determining the replication competence of chromatin in quiescent cells." *J Struct Biol* 129(2-3): 198-210.

Madiraju, M. V., M. H. Qin, K. Yamamoto, M. A. Atkinson and M. Rajagopalan (1999). "The dnaA gene region of *Mycobacterium avium* and the autonomous replication activities of its 5' and 3' flanking regions." *Microbiology* 145 (Pt 10): 2913-21.

Maiorano, D., J. M. Lemaitre and M. Mechali (2000). "Stepwise regulated chromatin assembly of MCM2-7 proteins." *J Biol Chem* 275(12): 8426-31.

Majka, J., J. Zakrzewska-Czerwinska and W. Messer (2001). "Sequence recognition, cooperative interaction, and dimerization of the initiator protein DnaA of *Streptomyces*." *J Biol Chem* 276(9): 6243-52.

Makise, M., S. Mima, T. Tsuchiya and T. Mizushima (2001). "Molecular mechanism for functional interaction between DnaA protein and acidic phospholipids: identification of important amino acids." *J Biol Chem* 276(10): 7450-6.

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

Margulies, C. and J. M. Kaguni (1996). "Ordered and sequential binding of DnaA protein to oriC, the chromosomal origin of *Escherichia coli*." *J Biol Chem* 271(29): 17035-40.

Margulies, C. and J. M. Kaguni (1998). "The FIS protein fails to block the binding of DnaA protein to oriC, the *Escherichia coli* chromosomal origin." *Nucleic Acids Res* 26(22): 5170-5.

Marszalek, J. and J. M. Kaguni (1994). "DnaA protein directs the binding of DnaB protein in initiation of DNA replication in *Escherichia coli*." *J Biol Chem* 269(7): 4883-90.

Marszalek, J., W. Zhang, T. R. Hupp, C. Margulies, K. M. Carr, S. Cherry and J. M. Kaguni (1996). "Domains of DnaA protein involved in interaction with DnaB protein, and in unwinding the *Escherichia coli* chromosomal origin." *J Biol Chem* 271(31): 18535-42.

Masai, H. and K. Arai (2000). "Dbf4 motifs: conserved motifs in activation subunits for Cdc7 kinases essential for S-phase." *Biochem Biophys Res Commun* 275(1): 228-32.

Masai, H. and K. Arai (2000). "Regulation of DNA replication during the cell cycle: roles of Cdc7 kinase and coupling of replication, recombination, and repair in response to replication fork arrest." *IUBMB Life* 49(5): 353-64.

Masai, H. and K. Arai (2002). "Cdc7 kinase complex: a key regulator in the initiation of DNA replication." *J Cell Physiol* 190(3): 287-96.

Mastrangelo, I. A., P. V. Hough, J. S. Wall, M. Dodson, F. B. Dean and J. Hurwitz (1989). "ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication." *Nature* 338(6217): 658-62.

Masuda, T., S. Mimura and H. Takisawa (2003). "CDK- and Cdc45-dependent priming of the MCM complex on chromatin during S-phase in *Xenopus* egg extracts: possible activation of MCM helicase by association with Cdc45." *Genes Cells* 8(2): 145-61.

Matsui, M., A. Oka, M. Takanami, S. Yasuda and Y. Hirota (1985). "Sites of dnaA protein-binding in the replication origin of the *Escherichia coli* K-12 chromosome." *J Mol Biol* 184(3): 529-33.

Matsunaga, F., P. Forterre, Y. Ishino and H. Myllykallio (2001). "*In vivo* interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin." *Proc Natl Acad Sci* U S A 98(20): 11152-7.

McGarry, K. C., V. T. Ryan, J. E. Grimwade and A. C. Leonard (2004). "Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP." *Proc Natl Acad Sci U S A* 101(9): 2811-6.

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

Mendez, J. and B. Stillman (2003). "Perpetuating the double helix: molecular machines at eukaryotic DNA replication origins." *Bioessays* 25(12): 1158-67.

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

Messer, W. (1987). "Initiation of DNA replication in *Escherichia coli*." *J Bacteriol* 169(8): 3395-9.

Messer, W. (2002). "The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication." *FEMS Microbiol Rev* 26(4): 355-74.

Messer, W., F. Blaesing, D. Jakimowicz, M. Krause, J. Majka, J. Nardmann, S. Schaper, H. Seitz, C. Speck, C. Weigel, G. Wegrzyn, M. Welzeck and J. Zakrzewska-Czerwinska (2001). "Bacterial replication initiator DnaA. Rules for DnaA binding and roles of DnaA in origin unwinding and helicase loading." *Biochimie* 83(1): 5-12.

Messer, W., W. Seufert, C. Schaefer, A. Gielow, H. Hartmann and M. Wende (1988). "Functions of the DnaA protein of *Escherichia coli* in replication and transcription." *Biochim Biophys Acta* 951(2-3): 351-8.

Messer, W. and J. Zakrzewska-Czerwinska (2002). "*Streptomyces* and *Escherichia coli*, model organisms for the analysis of the initiation of bacterial chromosome replication." *Arch Immunol Ther Exp* 50(6): 393-8.

Michaelis, L. a. M., M.L. (1913). Biochem Zeitschr 49(333).

Mimura, S. and H. Takisawa (1998). "Xenopus Cdc45-dependent loading of DNA polymerase alpha onto chromatin under the control of S-phase Cdk." *EMBO J* 17(19): 5699-707.

Mizushima, T. (1998). "Function of DnaA protein, the initiator for chromosomal DNA replication in E. coli." Yakugaku Zasshi 118(7): 257-71.

Mizushima, T. (2000). "Site-directed mutational analysis of DnaA protein, the initiator of chromosomal DNA replication in *E. coli*." *J Biochem* 127(1): 1-7.

Mizushima, T., N. Takahashi and B. Stillman (2000). "Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro." *Genes Dev* 14(13): 1631-41.

Mizushima, T., T. Takaki, T. Kubota, T. Tsuchiya, T. Miki, T. Katayama and K. Sekimizu (1998). "Site-directed mutational analysis for the ATP binding of DnaA protein. Functions of two conserved amino acids (Lys-178 and Asp-235) located in the ATP-binding domain of DnaA protein in vitro and in vivo." *J Biol Chem* 273(33): 20847-51.

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

Morigen, E. Boye, K. Skarstad and A. Lobner-Olesen (2001). "Regulation of chromosomal replication by DnaA protein availability in *Escherichia coli*: effects of the datA region." *Biochim Biophys Acta* 1521(1-3): 73-80.

Mourez, M., M. Hofnung and E. Dassa (1997). "Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permeases defines an important site of interaction with the ATPase subunits." *EMBO J* 16(11): 3066-77.

Musahl, C., D. Schulte, R. Burkhart and R. Knippers (1995). "A human homologue of the yeast replication protein Cdc21. Interactions with other Mcm proteins." *Eur J Biochem* 230(3): 1096-101.

Muzi Falconi, M., G. W. Brown and T. J. Kelly (1996). "cdc18+ regulates initiation of DNA replication in *Schizosaccharomyces pombe*." *Proc Natl Acad Sci U S A* 93(4): 1566-70.

Muzi-Falconi, M., G. W. Brown and T. J. Kelly (1996). "Controlling initiation during the cell cycle. DNA replication." *Curr Biol* 6(3): 229-33.

Myllykallio, H. and P. Forterre (2000). "Mapping of a chromosome replication origin in an archaeon: response." *Trends Microbiol* 8(12): 537-9.

Myllykallio, H., P. Lopez, P. Lopez-Garcia, R. Heilig, W. Saurin, Y. Zivanovic, H. Philippe and P. Forterre (2000). "Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon." *Science* 288(5474): 2212-5.

Natale, D. A., C. J. Li, W. H. Sun and M. L. DePamphilis (2000). "Selective instability of Orc1 protein accounts for the absence of functional origin recognition complexes during the M-G(1) transition in mammals." *EMBO J* 19(11): 2728-38.

Nelson, C. C., S. C. Hendy and P. J. Romaniuk (1995). "Relationship between P-box amino acid sequence and DNA binding specificity of the thyroid hormone receptor. The effects of half-site sequence in everted repeats." *J Biol Chem* 270(28): 16981-7.

Neuwald, A. F., L. Aravind, J. L. Spouge and E. V. Koonin (1999). "AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes." *Genome Res* 9(1): 27-43.

Newlon, C. S. and J. F. Theis (1993). "The structure and function of yeast ARS elements." *Curr Opin Genet Dev* 3(5): 752-8.

Ng, W. V., S. P. Kennedy, G. G. Mahairas, B. Berquist, M. Pan, H. D. Shukla, S. R. Lasky, N. S. Baliga, V. Thorsson, J. Sbrogna, S. Swartzell, D. Weir, J. Hall, T. A. Dahl, R. Welti, Y. A. Goo, B. Leithauser, K. Keller, R. Cruz, M. J. Danson, D. W. Hough, D. G. Maddocks, P. E. Jablonski, M. P. Krebs, C. M. Angevine, H. Dale, T. A. Isenbarger, R. F. Peck, M. Pohlschroder, J. L. Spudich, K. W. Jung, M. Alam, T. Freitas, S. Hou, C. J. Daniels, P. P. Dennis, A. D. Omer, H. Ebhardt, T. M. Lowe, P. Liang, M. Riley, L. Hood and S. DasSarma (2000). "Genome sequence of *Halobacterium* species NRC-1." *Proc Natl Acad Sci U S A* 97(22): 12176-81.

Nishitani, H., Z. Lygerou, T. Nishimoto and P. Nurse (2000). "The Cdt1 protein is required to license DNA for replication in fission yeast." *Nature* 404(6778): 625-8.

Nougarede, R., F. Della Seta, P. Zarzov and E. Schwob (2000). "Hierarchy of S-phase-promoting factors: yeast Dbf4-Cdc7 kinase requires prior S-phase cyclin-dependent kinase activation." *Mol Cell Biol* 20(11): 3795-806.

Ogawa, Y., T. Takahashi and H. Masukata (1999). "Association of fission yeast Orp1 and Mcm6 proteins with chromosomal replication origins." *Mol Cell Biol* 19(10): 7228-36.

Ogura, T., S. W. Whiteheart and A. J. Wilkinson (2004). "Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases." *J Struct Biol* 146(1-2): 106-12.

Ogura, T. and A. J. Wilkinson (2001). "AAA+ superfamily ATPases: common structure--diverse function." *Genes Cells* 6(7): 575-97.

Okuno, Y., A. J. McNairn, N. den Elzen, J. Pines and D. M. Gilbert (2001). "Stability, chromatin association and functional activity of mammalian pre-replication complex proteins during the cell cycle." *EMBO J* 20(15): 4263-77.

Okuno, Y., H. Satoh, M. Sekiguchi and H. Masukata (1999). "Clustered adenine/thymine stretches are essential for function of a fission yeast replication origin." *Mol Cell Biol* 19(10): 6699-709.

Olsen, G. J. and C. R. Woese (1997). "Archaeal genomics: an overview." *Cell* 89(7): 991-4.

Pabo, C. O. and R. T. Sauer (1992). "Transcription factors: structural families and principles of DNA recognition." *Annu Rev Biochem* 61: 1053-95.

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

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

Pati, U. K. (1992). "Novel vectors for expression of cDNA encoding epitope-tagged proteins in mammalian cells." *Gene* 114(2): 285-288.

Pause, A. and N. Sonenberg (1992). "Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A." *EMBO J* 11(7): 2643-54.

Pelizon, C., M. A. Madine, P. Romanowski and R. A. Laskey (2000). "Unphosphorylatable mutants of Cdc6 disrupt its nuclear export but still support DNA replication once per cell cycle." *Genes Dev* 14(19): 2526-33.

Perkins, G. and J. F. Diffley (1998). "Nucleotide-dependent prereplicative complex assembly by Cdc6p, a homolog of eukaryotic and prokaryotic clamp-loaders." *Mol Cell* 2(1): 23-32.

Phizicky, E. M. and S. Fields (1995). "Protein-protein interactions: methods for detection and analysis." *Microbiol Rev* 59(1): 94-123.

Polaczek, P. (1990). "Bending of the origin of replication of E. coli by binding of IHF at a specific site." *New Biol* 2(3): 265-71.

Poplawski, A., B. Grabowski, S. E. Long and Z. Kelman (2001). "The zinc finger domain of the archaeal minichromosome maintenance protein is required for helicase activity." *J Biol Chem* 276(52): 49371-7.

Rao, H. and B. Stillman (1995). "The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators." *Proc Natl Acad Sci U S A* 92(6): 2224-8.

Reeve, J. N. (2003). "Archaeal chromatin and transcription." *Mol Microbiol* 48(3): 587-98.

Remus, D., E. L. Beall and M. R. Botchan (2004). "DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC-DNA binding." *EMBO J* 23(4): 897-907.

Robinson, N. P., I. Dionne, M. Lundgren, V. L. Marsh, R. Bernander and S. D. Bell (2004). "Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*." *Cell* 116(1): 25-38.

Romanowski, P., M. A. Madine, A. Rowles, J. J. Blow and R. A. Laskey (1996). "The *Xenopus* origin recognition complex is essential for DNA replication and MCM binding to chromatin." *Curr Biol* 6(11): 1416-25.

Romanowski, P., J. Marr, M. A. Madine, A. Rowles, J. J. Blow, J. Gautier and R. A. Laskey (2000). "Interaction of *Xenopus* Cdc2 x cyclin A1 with the origin recognition complex." *J Biol Chem* 275(6): 4239-43.

Rost, B., C. Sander and R. Schneider (1994). "PHD--an automatic mail server for protein secondary structure prediction." *Comput Appl Biosci* 10(1): 53-60.

Rowles, A., J. P. Chong, L. Brown, M. Howell, G. I. Evan and J. J. Blow (1996). "Interaction between the origin recognition complex and the replication licensing system in *Xenopus*." *Cell* 87(2): 287-96.

Rowles, A., S. Tada and J. J. Blow (1999). "Changes in association of the *Xenopus* origin recognition complex with chromatin on licensing of replication origins." *J Cell Sci* 112 (Pt 12): 2011-8.

Rowley, A., J. H. Cocker, J. Harwood and J. F. Diffley (1995). "Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC." *EMBO* J 14(11): 2631-41.

Ryan, V. T., J. E. Grimwade, J. E. Camara, E. Crooke and A. C. Leonard (2004). "Escherichia coli pre-replication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA." *Mol Microbiol* 51(5): 1347-59.

Ryan, V. T., J. E. Grimwade, C. J. Nievera and A. C. Leonard (2002). "IHF and HU stimulate assembly of pre-replication complexes at *Escherichia coli* oriC by two different mechanisms." *Mol Microbiol* 46(1): 113-24.

Saha, P., J. Chen, K. C. Thome, S. J. Lawlis, Z. H. Hou, M. Hendricks, J. D. Parvin and A. Dutta (1998). "Human CDC6/Cdc18 associates with Orc1 and cyclin-cdk and is selectively eliminated from the nucleus at the onset of S phase." *Mol Cell Biol* 18(5): 2758-67.

Sambrook, J. a. D. W. R. (2001). "Molecular clonning: a laboratory manual." New York Cold Spring Harbour Laboratory press.

Samitt, C. E., F. G. Hansen, J. F. Miller and M. Schaechter (1989). "*In vivo* studies of DnaA binding to the origin of replication of *Escherichia coli*." *EMBO* J 8(3): 989-93.

Sandman, K. and J. N. Reeve (2000). "Structure and functional relationships of archaeal and eukaryal histones and nucleosomes." *Arch Microbiol* 173(3): 165-9.

Santocanale, C. and J. F. Diffley (1996). "ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in *Saccharomyces cerevisiae*." *EMBO J* 15(23): 6671-9.

Sato, M., T. Gotow, Z. You, Y. Komamura-Kohno, Y. Uchiyama, N. Yabuta, H. Nojima and Y. Ishimi (2000). "Electron microscopic observation and single-stranded DNA binding activity of the Mcm4,6,7 complex." *J Mol Biol* 300(3): 421-31.

Sato, N., K. Arai and H. Masai (1997). "Human and Xenopus cDNAs encoding budding yeast Cdc7-related kinases: in vitro phosphorylation of MCM subunits by a putative human homologue of Cdc7." *EMBO J* 16(14): 4340-51.

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

Schwacha, A. and S. P. Bell (2001). "Interactions between two catalytically distinct MCM subgroups are essential for coordinated ATP hydrolysis and DNA replication." *Mol Cell* 8(5): 1093-104.

Sclafani, R. A., R. J. Fletcher and X. S. Chen (2004). "Two heads are better than one: regulation of DNA replication by hexameric helicases." *Genes Dev* 18(17): 2039-45.

Segurado, M., A. de Luis and F. Antequera (2003). "Genome-wide distribution of DNA replication origins at A+T-rich islands in *Schizosaccharomyces pombe*." *EMBO Rep* 4(11): 1048-53.

Seitz, H., C. Weigel and W. Messer (2000). "The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli*." *Mol Microbiol* 37(5): 1270-9.

Sekimizu, K., D. Bramhill and A. Kornberg (1987). "ATP activates DnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome." *Cell* 50(2): 259-65.

Sekimizu, K., D. Bramhill and A. Kornberg (1988). "Sequential early stages in the in vitro initiation of replication at the origin of the *Escherichia coli* chromosome." *J Biol Chem* 263(15): 7124-30.

Sekimizu, K., B. Y. Yung and A. Kornberg (1988). "The DnaA protein of *Escherichia coli*. Abundance, improved purification, and membrane binding." *J Biol Chem* 263(15): 7136-40.

She, Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. C. Chan-Weiher, I. G. Clausen, B. A. Curtis, A. De Moors, G. Erauso, C. Fletcher, P. M. Gordon, I. Heikamp-de Jong, A. C. Jeffries, C. J. Kozera, N. Medina, X. Peng, H. P. Thi-Ngoc, P. Redder, M. E. Schenk, C. Theriault, N. Tolstrup, R. L. Charlebois, W. F. Doolittle, M. Duguet, T. Gaasterland, R. A. Garrett, M. A. Ragan, C. W. Sensen and J. Van der Oost (2001). "The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2." *Proc Natl Acad Sci U S A* 98(14): 7835-40.

Shechter, D. F., C. Y. Ying and J. Gautier (2000). "The intrinsic DNA helicase activity of *Methanobacterium thermoautotrophicum* delta H minichromosome maintenance protein." *J Biol Chem* 275(20): 15049-59.

Sherman, M. B., T. Soejima, W. Chiu and M. van Heel (1998). "Multivariate analysis of single unit cells in electron crystallography." *Ultramicroscopy* 74(4): 179-99.

Shin, J. H., B. Grabowski, R. Kasiviswanathan, S. D. Bell and Z. Kelman (2003). "Regulation of minichromosome maintenance helicase activity by Cdc6." *J Biol Chem* 278(39): 38059-67.

Shin, J. H., Y. Jiang, B. Grabowski, J. Hurwitz and Z. Kelman (2003). "Substrate requirements for duplex DNA translocation by the eukaryal and archaeal minichromosome maintenance helicases." *J Biol Chem* 278(49): 49053-62.

Shiomi, Y., J. Usukura, Y. Masamura, K. Takeyasu, Y. Nakayama, C. Obuse, H. Yoshikawa and T. Tsurimoto (2000). "ATP-dependent structural change of the eukaryotic clamp-loader protein, replication factor C." *Proc Natl Acad Sci U S A* 97(26): 14127-32.

Simmons, D. T. (2000). "SV40 large T antigen functions in DNA replication and transformation." *Adv Virus Res* 55: 75-134.

Simmons, L. A., M. Felczak and J. M. Kaguni (2003). "DnaA Protein of Escherichia coli: oligomerization at the *E. coli* chromosomal origin is required for initiation and involves specific N-terminal amino acids." *Mol Microbiol* 49(3): 849-58.

Singleton, M. R., R. Morales, I. Grainge, N. Cook, M. N. Isupov and D. B. Wigley (2004). "Conformational changes induced by nucleotide binding in Cdc6/ORC from *Aeropyrum pernix*." *J Mol Biol* 343(3): 547-57.

Skarstad, K., T. A. Baker and A. Kornberg (1990). "Strand separation required for initiation of replication at the chromosomal origin of *E.coli* is facilitated by a distant RNA--DNA hybrid." *EMBO J* 9(7): 2341-8.

Skarstad, K., E. Boye and H. B. Steen (1986). "Timing of initiation of chromosome replication in individual *Escherichia coli* cells." *EMBO J* 5(7): 1711-7.

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

Speck, C. and W. Messer (2001). "Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA." *EMBO J* 20(6): 1469-76.

Speck, C., C. Weigel and W. Messer (1999). "ATP- and ADP-dnaA protein, a molecular switch in gene regulation." *EMBO J* 18(21): 6169-76.

Stillman, B., S. P. Bell, A. Dutta and Y. Marahrens (1992). "DNA replication and the cell cycle." *Ciba Found Symp* 170: 147-56; discussion 156-60.

Stinchcomb, D. T., K. Struhl and R. W. Davis (1979). "Isolation and characterisation of a yeast chromosomal replicator." *Nature* 282(5734): 39-43.

Studier, F. W. and B. A. Moffatt (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." *J Mol Biol* 189(1): 113-30.

Story, R.M. and Steitz, T.A. (1992). "Structure of the recA protein-ADP complex". *Nature* 355 (6358) 374-6

Sutton, M. D., K. M. Carr, M. Vicente and J. M. Kaguni (1998). "Escherichia coli DnaA protein. The N-terminal domain and loading of DnaB helicase at the *E. coli* chromosomal origin." *J Biol Chem* 273(51): 34255-62.

Tada, S., A. Li, D. Maiorano, M. Mechali and J. J. Blow (2001). "Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin." *Nat Cell Biol* 3(2): 107-13.

Takeda, T., K. Ogino, E. Matsui, M. K. Cho, H. Kumagai, T. Miyake, K. Arai and H. Masai (1999). "A fission yeast gene, him1(+)/dfp1(+), encoding a regulatory subunit for Hsk1 kinase, plays essential roles in S-phase initiation as well as in S-phase checkpoint control and recovery from DNA damage." *Mol Cell Biol* 19(8): 5535-47.

Tanaka, T., D. Knapp and K. Nasmyth (1997). "Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs." *Cell* 90(4): 649-60.

Theisen, P. W., J. E. Grimwade, A. C. Leonard, J. A. Bogan and C. E. Helmstetter (1993). "Correlation of gene transcription with the time of initiation of chromosome replication in *Escherichia coli*." *Mol Microbiol* 10(3): 575-84.

Thome, K. C., S. K. Dhar, D. G. Quintana, L. Delmolino, A. Shahsafaei and A. Dutta (2000). "Subsets of human origin recognition complex (ORC) subunits are expressed in non-proliferating cells and associate with non-ORC proteins." *J Biol Chem* 275(45): 35233-41.

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

Tye, B.K and Sawyer, S. (2000). "The hexameric eukaryotic MCM helicase:building symmetry from nonidentical parts". *J Biol Chem* 275 (45):34833-6

Vas, A., W. Mok and J. Leatherwood (2001). "Control of DNA rereplication via Cdc2 phosphorylation sites in the origin recognition complex." *Mol Cell Biol* 21(17): 5767-77.

von Freiesleben, U., K. V. Rasmussen and M. Schaechter (1994). "SeqA limits DnaA activity in replication from oriC in *Escherichia coli*." *Mol Microbiol* 14(4): 763-72.

Wahle, E., R. S. Lasken and A. Kornberg (1989). "The DnaB-DnaC replication protein complex of *Escherichia coli*. I. Formation and properties." *J Biol Chem* 264(5): 2463-8.

Walker, J. E., M. Saraste, M. J. Runswick and N. J. Gay (1982). "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold." *EMBO J* 1(8): 945-51.

Walker, S. S., S. C. Francesconi and S. Eisenberg (1990). "A DNA replication enhancer in *Saccharomyces cerevisiae*." *Proc Natl Acad Sci U S A* 87(12): 4665-9.

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

Walter, J. C. (2000). "Evidence for sequential action of cdc7 and cdk2 protein kinases during initiation of DNA replication in *Xenopus* egg extracts." *J Biol Chem* 275(50): 39773-8.

Wang, H. and S. J. Elledge (1999). "DRC1, DNA replication and checkpoint protein 1, functions with DPB11 to control DNA replication and the S-phase checkpoint in *Saccharomyces cerevisiae*." *Proc Natl Acad Sci U S A* 96(7): 3824-9.

Weigel, C., W. Messer, S. Preiss, M. Welzeck, Morigen and E. Boye (2001). "The sequence requirements for a functional *Escherichia coli* replication origin are different for the chromosome and a minichromosome." *Mol Microbiol* 40(2): 498-507.

Weigel, C., A. Schmidt, B. Ruckert, R. Lurz and W. Messer (1997). "DnaA protein binding to individual DnaA boxes in the *Escherichia coli* replication origin, oriC." *EMBO J* 16(21): 6574-83.

Weigel, C., A. Schmidt, H. Seitz, D. Tungler, M. Welzeck and W. Messer (1999). "The N-terminus promotes oligomerization of the *Escherichia coli* initiator protein DnaA." *Mol Microbiol* 34(1): 53-66.

Weigelt, J., S. E. Brown, C. S. Miles, N. E. Dixon and G. Otting (1999). "NMR structure of the N-terminal domain of *E. coli* DnaB helicase: implications for structure rearrangements in the helicase hexamer." *Structure Fold Des* 7(6): 681-90.

Weinreich, M., C. Liang and B. Stillman (1999). "The Cdc6p nucleotide-binding motif is required for loading mcm proteins onto chromatin." *Proc Natl Acad Sci U S A* 96(2): 441-6.

Weinreich, M. and B. Stillman (1999). "Cdc7p-Dbf4p kinase binds to chromatin during S phase and is regulated by both the APC and the RAD53 checkpoint pathway." *EMBO J* 18(19): 5334-46.

Whittaker, A. J., I. Royzman and T. L. Orr-Weaver (2000). "Drosophila double parked: a conserved, essential replication protein that co-localizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts." *Genes Dev* 14(14): 1765-76.

Wilmes, G. M. and S. P. Bell (2002). "The B2 element of the *Saccharomyces cerevisiae* ARS1 origin of replication requires specific sequences to facilitate pre-RC formation." *Proc Natl Acad Sci U S A* 99(1): 101-6.

Woese, C. R. (2000). "Interpreting the universal phylogenetic tree." *Proc Natl Acad Sci U S A* 97(15): 8392-6.

Woese, C. R. and G. E. Fox (1977). "Phylogenetic structure of the prokaryotic domain: the primary kingdoms." *Proc Natl Acad Sci U S A* 74(11): 5088-90.

Woese, C. R., O. Kandler and M. L. Wheelis (1990). "Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya." *Proc Natl Acad Sci U S A* 87(12): 4576-9.

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

Wohlschlegel, J. A., B. T. Dwyer, S. K. Dhar, C. Cvetic, J. C. Walter and A. Dutta (2000). "Inhibition of eukaryotic DNA replication by geminin binding to Cdt1." *Science* 290(5500): 2309-12.

Wold, S., K. Skarstad, H. B. Steen, T. Stokke and E. Boye (1994). "The initiation mass for DNA replication in *Escherichia coli* K-12 is dependent on growth rate." *EMBO J* 13(9): 2097-102.

Yabuta, N., N. Kajimura, K. Mayanagi, M. Sato, T. Gotow, Y. Uchiyama, Y. Ishimi and H. Nojima (2003). "Mammalian Mcm2/4/6/7 complex forms a toroidal structure." *Genes Cells* 8(5): 413-21.

Yan, H., Merchant, A.M. and Tye, B.K. (1993). Cell cycle-regulated nuclear localization of MCM2 and MCM3, which are required for the initiation of DNA synthesis at chromosomal replication origins in yeast". *Genes Dev* 7(11): 2149-60.

You, Z., Y. Komamura and Y. Ishimi (1999). "Biochemical analysis of the intrinsic Mcm4-Mcm6-mcm7 DNA helicase activity." *Mol Cell Biol* 19(12): 8003-15.

Young, M. R. and B. K. Tye (1997). "Mcm2 and Mcm3 are constitutive nuclear proteins that exhibit distinct isoforms and bind chromatin during specific cell cycle stages of *Saccharomyces cerevisiae*." *Mol Biol Cell* 8(8): 1587-601.

Yu, X., M. M. Hingorani, S. S. Patel and E. H. Egelman (1996). "DNA is bound within the central hole to one or two of the six subunits of the T7 DNA helicase." *Nat Struct Biol* 3(9): 740-3.

Yu, X., M. J. Jezewska, W. Bujalowski and E. H. Egelman (1996). "The hexameric *E. coli* DnaB helicase can exist in different quaternary states." *J Mol Biol* 259(1): 7-14.

Yu, X., M. S. VanLoock, A. Poplawski, Z. Kelman, T. Xiang, B. K. Tye and E. H. Egelman (2002). "The *Methanobacterium thermoautotrophicum* MCM protein can form heptameric rings." *EMBO Rep* 3(8): 792-7.

Zawilak, A., M. C. Durrant, P. Jakimowicz, S. Backert and J. Zakrzewska-Czerwinska (2003). "DNA binding specificity of the replication initiator protein, DnaA from Helicobacter pylori." *J Mol Biol* 334(5): 933-47.

Zhang, R. and C. T. Zhang (2003). "Multiple replication origins of the archaeon *Halobacterium* species NRC-1." *Biochem Biophys Res Commun* 302(4): 728-34.

Zheng, W., Z. Li, K. Skarstad and E. Crooke (2001). "Mutations in DnaA protein suppress the growth arrest of acidic phospholipid-deficient *Escherichia coli cells*." *EMBO J* 20(5): 1164-72.

Zou, L., J. Mitchell and B. Stillman (1997). "CDC45, a novel yeast gene that functions with the origin recognition complex and Mcm proteins in initiation of DNA replication." *Mol Cell Biol* 17(2): 553-63.

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

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