Critical Reviews in Biochemistry and Molecular Biology, 2009; 44(5): 326–342

informa
healthcare

REVIEW ARTICLE

Structural biology of MCM helicases

Alessandro Costa[1], and Silvia Onesti[2]

[1]*Sir William Dunn School of Pathology, University of Oxford, UK, and* [2]*ELETTRA Synchrotron Light Laboratory & International School for Advanced Studies (SISSA), Trieste, Italy*

Abstract

The eukaryotic MCM2-7 complex is recruited onto origins of replication during the G1 phase of the cell cycle and acts as the main helicase at the replication fork during the S phase. Over the last few years a number of structural reports on MCM proteins using both electron microscopy and protein crystallography have been published. The crystal structures of two (almost) full-length archaeal homologs provide the first atomic pictures of a MCM helicase. However one of the structures is at low resolution and the other is of an inactive MCM. Moreover, both proteins are monomeric in the crystal, whereas the activity of the complex is critically dependent on oligomerization. Lower resolution structures derived from electron microscopy studies are therefore crucial to complement the crystallographic analysis and to assemble the multimeric complex that is active in the cell. A critical analysis of all the structural results elucidates the potential conformational changes and dynamic behavior of MCM helicase to provide a first insight into the gamut of molecular configurations adopted during the processes of DNA melting and unwinding.

Keywords: *AAA+ ATPases; MCM2-7; eukaryotic DNA replication; glutamate switch; ploughshare model*

Introduction

Eukaryotic DNA replication is a highly coordinated and tightly regulated process. Due to the size of the eukaryotic genome, initiation of replication takes place at multiple sites throughout chromatin and a complex network of proteins, under strict cell-cycle control, is required to ensure that each origin is used only once and no segment of DNA is left unreplicated or undergoes multiple rounds of replication. Although the details of the process used by the cell to achieve this precise coordination are still not completely understood, the consensus view is that a pre-replicative complex (pre-RC) is assembled in a stepwise manner on the origin during the G1 phase, making chromatin competent ("licensed") for replication (Bell and Dutta, 2002; Sclafani and Holzen, 2007).

A key role in DNA licensing is played by the minichromosome maintenance (MCM) proteins, a family of polypeptides, ranging from 600 to 1200 amino-acid residues, originally discovered and named as factors that support minichromosome maintenance in yeast (Forsburg, 2004; Maiorano *et al.*, 2006; Costa and Onesti, 2008). All eukaryotic genomes possess at least six paralogs (MCM2-7) that work together in a multimeric complex, and their deletion is lethal in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. The MCM proteins have been shown to be chromatin-bound in the G1 phase, displaced during replication in S phase, and absent from chromatin in G2 phase (Diffley and Labib, 2002; Stillman, 2005). The regeneration of replication competence is associated with a new binding event of MCMs to chromatin at the end of mitosis (Blow and Dutta, 2005).

The MCM proteins are therefore present only in proliferating cells and are highly expressed in malignant human cancer cells and pre-cancerous cells undergoing malignant transformation. They are not expressed in differentiated somatic cells that have been withdrawn from the cell cycle. Therefore, these proteins are ideal diagnostic markers for cancer and possibly targets for anti-cancer drug development (Blow and Hodgson, 2002). Several reports have confirmed that MCM proteins can be used as a prognostic marker for the identification of malignant and pre-malignant cells for a variety of cancers. Their expression has been compared with routinely used proliferation markers, such as Ki-67 (MIB-1) and proliferating cell nuclear antigen (PCNA) and it has been shown that they are excellent candidates as novel

Address for Correspondence: Silvia Onesti, Sincrotrone Trieste S.C.P.A., Area Science Park, 34149 Basovizza, Trieste, Italy. Tel: +39 040 375 8451. E-mail: silvia.onesti@elettra.trieste.it

(Received 17 June 2009; revised 09 July 2009; accepted 15 July 2009)

ISSN 1040-9238 print/ISSN 1549-7798 online © 2009 Informa UK Ltd
DOI: 10.1080/10409230903186012

biomarkers for routine use in clinical practice (Laskey, 2005; Lei, 2005).

The eukaryotic MCM protein family

The MCM2-7 complex

A six-subunit multiprotein complex, called the origin recognition complex (ORC), binds DNA to specify origin sites (Wigley, 2009) and serves as a platform for the assembly of the pre-RC (Figure 1A). In addition to ORC, at least two other proteins, Cdc6 and Ctd1, are essential for loading the MCM complex onto chromatin (Lei and Tye, 2001; Speck *et al.*, 2005). Once this has been accomplished, ORC and Cdc6 can be removed without interfering with subsequent steps in DNA replication, suggesting that the primary role of the preRC is MCM loading (Donovan *et al.*, 1997). Chromatin, thus licensed for replication, is guided into the S phase by cell cycle regulated protein kinases, which trigger a series of events leading to the activation of the MCM complex, the melting of origin DNA, the recruitment of proteins involved in DNA replication and the establishment of the replication forks (Diffley, 2004). MCM proteins have a role during the initiation of DNA replication, but are also essential in the elongation step. Detailed studies, involving the specific degradation of each of the *S. cerevisiae* MCM2-7 homologs at different stages throughout the cell cycle, have shown that not only are all the six proteins equally important for entry in S phase, but their subsequent inactivation irreversibly blocks the progression of the replication fork (Labib *et al.*, 2000; 2001).

Within all the MCM proteins, a central region of about 250 amino-acid residues shows the highest degree of conservation and belongs to the AAA+ superfamily of ATPases (Iyer *et al.*, 2004) (Figure 1B); many other proteins are involved in DNA replication, repair and recombination, including Cdc6, the origin recognition proteins, such as DnaA, Orc1, Orc4, Orc5, and proteins involved in recombination and Holliday junctions resolution, such as RuvB and RecG (Davey *et al.*, 2002).

Numerous lines of evidence suggested that the cellular role of MCM protein was DNA unwinding during DNA replication. However a number of observations did not fit the canonical model for a replicative helicase. For a long time it was thought that the purified MCM2-7 did not display any helicase activity, with only a weak helicase activity found associated with the MCM4/6/7 sub-complex (Ishimi, 1997; You *et al.*, 1999; Lee and Hurwitz, 2001). However, a recent biochemical study does identify specific buffer conditions in which the helicase activity of the recombinant MCM2-7 complex is measurable (Bochman and Schwacha, 2008). Moreover, a stable complex comprising Cdc45, MCM2-7 and GINS can be purified from *Drosophila* embryo extracts and displays an ATP-dependent helicase activity (Moyer *et al.*, 2006) (Figure 1A), thus suggesting that Cdc45 and GINS may provide an allosteric switch which triggers the activation of the replicative helicase. All these data converge to indicate that, although the MCM2-7 complex is indeed a helicase, its activity is not robust and most likely requires protein activators and/or essential post-translational modifications in the context of the cell cycle. A second perplexing observation (known as the "MCM paradox") is the fact that the MCM complexes loaded onto chromatin largely exceed the number of replication origins (Edwards *et al.*, 2002) and do not localize only with the replication fork (Hyrien *et al.*, 2003; Takahashi *et al.*, 2005). Experiments in *Drosophila* (Crevel *et al.*, 2007), *Xenopus* (Woodward *et al.*, 2006; Ge *et al.*, 2007) and human cells (Ibarra *et al.*, 2008) have shown that DNA replication can proceed with a very reduced amount of MCM complex (down to 5% of the normal cellular levels). However in these conditions cells lose the ability to activate dormant origins required to recover from replication stress, and as a consequence accumulate DNA damage and display genome instability.

Other eukaryotic MCM proteins

Two eukaryotic paralogs have been described and named MCM8 and MCM9 (Gozuacik *et al.*, 2003; Johnson *et al.*, 2003; Lutzmann *et al.*, 2005; Yoshida, 2005). A detailed bioinformatics analysis has shown that these proteins, initially identified only in higher eukaryotes, have a widespread distribution and are present in the majority of eukaryotic genomes analyzed, but not in *S. cerevisiae*. The presence of these two factors is generally highly correlated (with the exception of *Drosophila* species, which possess MCM8 but lack MCM9) and is most likely due to a shared loss event, strongly suggesting that MCM8 and MCM9 have interdependent functional roles (Liu *et al.*, 2009). The functional characterization of MCM8 in *Xenopus* and human cells has led to contrasting results, with evidence for a role in the initiation phase (Volkening and Hoffmann, 2005; Kinoshita *et al.*, 2008) or in the elongation phase (Gozuacik *et al.*, 2003; Maiorano *et al.*, 2005). Despite some ambiguity, it is likely that MCM8 does not form a complex with MCM2-7 and has a separate role. To further complicate the picture, *Drosophila* MCM8 has been shown to be involved in meiotic recombination (Matsubayashi and Yamamoto, 2003; Blanton *et al.*, 2005), as well as in DNA replication (Crevel *et al.*, 2007). However both the unusual presence of MCM8 alone, and the highly divergent nature of the sequence, suggests some unique role for this protein in *Drosophila* species (Liu *et al.*, 2009). Much less is known regarding the role of MCM9: the only functional study shows that it forms a complex with Cdt1 and is required for MCM2-7 recruitment (Lutzmann and Mechali, 2008).

Finally, a novel component of the human MCM complex (named MCM-BP) has been found which copurifies with subunits MCM3-7, replacing MCM2 (Sakwe *et al.*, 2007). The role of this component is currently unknown, although the protein appears to be conserved in a number of eukaryotes and is distantly related to the MCM2-9 family. Although it was initially thought to be limited to multicellular organisms, MCM-BP is also widely spread, being present in some fungi (such

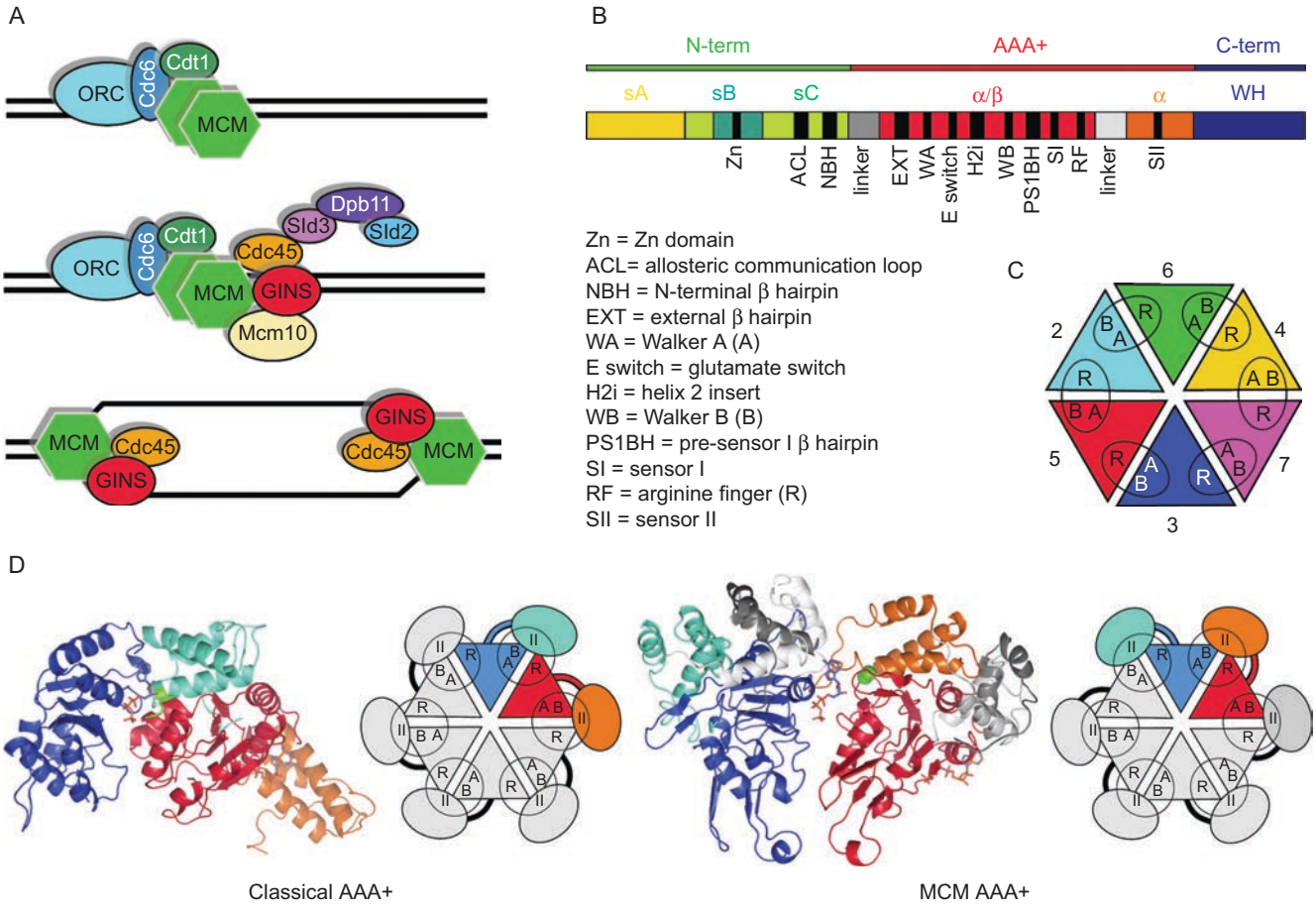

Figure 1. Role and architecture of MCM proteins. (A) A schematic diagram of the steps in the initiation of DNA replication in eukaryotes. *Top panel:* During the G1 phase of the cell cycle the pre-replicative complex (constituted by ORC, Cdc6, Cdt1 and MCM) is recruited at origins of replication. *Middle panel*: Upon DDK and CDK mediated activation, a number of additional factors are recruited, to form the pre-initiation complex. These include MCM10, Cdc45, GINS, Sld2, Sld3 and Dpb11. *Bottom panel*: During the elongation phase of DNA replication a complex made of MCM, GINS and Cdc45 has been found to travel together with the replisome. (B) The primary structure of MCM proteins comprises a Nterminal domain (green), a AAA+ domain (red) and a Cterminal domain (blue). The N-terminal domain can be further subdivided into three subdomains (Fletcher *et al.*, 2003) and includes a number of functional elements, such as the Zn motif, N-terminal β-hairpin (NBH) and ACL loop. The AAA+ domain folds into an α/β subdomain (red) and an α domain (orange); the former contains the Walker A, Walker B, sensor I and R finger motifs found in AAA+ protein, as well as the unique insertions (EXT, h2i and PS1BH) which characterize MCM; the latter contains the sensor II motif. The helical linkers connecting the N-terminal-AAA+ domains, and the α/β-α domains are shown in dark gray and light gray, respectively. The Cterminal domain is predicted to fold into a winged helix (WH) motif. (C) Schematic diagram of the putative MCM2-7 hexamer, showing the relative position of MCM proteins, as derived from biochemical data (Crevel *et al.*, 2001; Davey *et al.*, 2003; Bochman *et al.*, 2008). (D) The unusual architecture of the MCM AAA+ hexamer. The interface between two subunits within a hexamer as seen in the crystal structure of the bacterial transcriptional enhancer PspF bound to ATP (left panel; Rappas *et al.*, 2006) shows the arrangement seen in the majority of AAA+ proteins. In one subunit the α/β domain is in red, with the α domain in orange, whereas in the adjacent subunit the the α/β domain is in blue, with the α domain in cyan; the sensor II arginine highlighted in green; the ATP molecule at the interface between subunits is shown. As schematically illustrated in the diagram, the ATPase active site at the interface between the red and blue subunit is made up of residues mostly belonging to the blue subunit, including Walker A (A), Walker B (B), sensor I, and sensor II (II, in green) acting in *cis*, whereas the arginine finger (R) is contributed in *trans* by the α/β domain of the next-neighboring subunit. On the right panel the same inter-subunit interface is shown for a hexameric model derived from the crystal structure of *Sso*MCM-FL (Brewster *et al.*, 2008) overlapped on the *Mth*MCM N-terminal domain structure (Liu *et al.*, 2008). The same color code as for PspF is used; in addition the connection between the N and C-terminal domain is shown in dark gray, whereas the MCM peculiar insertion between the α/β and α domains is shown in light gray. The effect of this helical insertion is to dramatically reposition the α domain so that it is now located at the adjacent interface. The diagram illustrates how the active site now comprises the Walker A, Walker B, sensor I residues from one subunit (in *cis*) and the R finger and sensor II from the adjacent subunit (in *trans*).

as *S. pombe* and *Ustilago maydis*), and in a variety of other unicellular eukaryotes (including for example *Tetrahymena*, *Paramecium*, *Trypanosoma*, *Trichomonas*, and *Entamoeba*). The report on MCM-BP (Sakwe *et al.*, 2007) opens up the intriguing possibility that there may be alternative forms of the MCM complex, possibly functioning in specific circumstances.

MCM proteins in archaea

Archaeal cells contain a simplified version of the DNA replication apparatus, but maintain the structural peculiarity of the eukaryotic system, including the regulated initiation from multiple sites and the abundance of MCM proteins at origins (Duggin *et al.*, 2008).

While eukaryotes seem to have at least six different albeit homologous MCM proteins working together in a single complex, their number in archaea is variable, with most species possessing only one paralog which forms homomeric rings (Sakakibara *et al.*, 2009). The sequences of the archaeal MCM proteins show a level of sequence identity of roughly 25–35% with their eukaryotic counterparts.

The complexity of the eukaryotic MCM system has made it difficult to obtain large amounts of recombinant active protein, and as a consequence archaeal homologs have become popular model systems for studying the structure and function of the MCM helicases. The first indication that MCM proteins have ATPase and helicase activity came from studies of the *Methanothermobacter thermautotrophicus* (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000) and *Sulfolobus solfataricus* (Carpentieri *et al.*, 2002) proteins, which still now represent the best studied archaeal MCM systems. A large number of biochemical studies have been carried out on both proteins, elucidating the role of the separate domains (Kasiviswanathan *et al.*, 2004; Barry *et al.*, 2007; Pucci *et al.*, 2007), of various functional elements (McGeoch *et al.*, 2005; Jenkinson and Chong, 2006; Sakakibara *et al.*, 2008; Barry *et al.*, 2009), as well as specific residues (Pucci *et al.*, 2004; Moreau *et al.*, 2007; Jenkinson *et al.*, 2009) in DNA binding, ATPase and helicase activity. Details of these biochemical studies have been recently reviewed (Sakakibara *et al.*, 2009).

The simplicity of the archaeal complex has meant that most of the work addressing the structural aspects of the replicative helicase has been carried out on archaeal MCM proteins, as discussed below.

The architecture of MCM proteins

The sequence of a canonical MCM protein can be divided into three domains (Figure 1B). Most of the archaeal proteins conform to this model, whereas some of the eukaryotic MCMs have additional N- or C-terminal extensions.

The N-terminal domain displays a lower level of sequence conservation. The biochemical characterization of this domain in *M. thermautotrophicus* and *S. solfataricus* MCMs showed that it binds both single-stranded and double-stranded DNA (with a preference for ssDNA), is a strong determinant for hexamerization, is able to influence the processivity of the helicase and its polarity and most likely acts as a brake to ensure that the helicase activity is tightly regulated (Kasiviswanathan *et al.*, 2004; Jenkinson and Chong, 2006; Barry *et al.*, 2007; Pucci *et al.*, 2007; Liu *et al.*, 2008).

The "core" of the MCM helicase is the central AAA+ domain, which is responsible for the catalytic activity. The AAA+ superfamily of ATPases shares a number of conserved sequence motifs (Figure 1B), implicated in ATP binding and ATP hydrolysis: these include the Walker A and B motifs (present in most ATPases) and other sequence elements that are unique to the family, such as sensor I, sensor II and arginine finger (R finger). AAA+ proteins share a common fold, with a α/β mononucleotide binding subdomain and a less conserved, α helical sub domain (also known as "lid" domain). Although some AAA+ proteins exist and function as monomers, in a large number of occurrences they have been shown to form oligomeric assemblies, with a strong preference for hexameric structures (Figure 1C). Oligomerization is generally mediated by the mononucleotide binding fold, with an arrangement that places the nucleotide binding pocket at the interface between subunits, with some of the residues provided in *cis* (Walker A, Walker B, sensor I, sensor II) and some provided in *trans* (R finger) (Figure 1C). A systematic analysis of the AAA+ family identified a number of classes characterized by specific structural and functional elements (Iyer *et al.*, 2004). MCM proteins belong to a clade defined by the insertion of a β-hairpin before sensor I (pre-sensor I β-hairpin, PS1BH) and by an additional insertion in the middle of the second helix of the α/β domain (helix-2 insertion, h2i; Figure 1B). Both elements have been shown to have a critical role in DNA unwinding in the *S. solfataricus* and *M. thermautotrophicus* systems (Chong, 2005; Jenkinson and Chong, 2006; McGeoch *et al.*, 2005).

A less well characterized C-terminal domain is present in all the MCM proteins. In the archaeal homologs this domain was shown to have a regulatory effect (Barry *et al.*, 2007; Jenkinson and Chong, 2006). Sequence analysis of the archaeal sequence suggests the presence of a winged-helix (WH) domain (Aravind and Koonin, 1999), despite a low level of sequence conservation (Figure 1B). The signature of a WH motif is not detected in the eukaryotic sequences, but the length and hydrophobic pattern of the region may be compatible with a similar fold.

Structural studies of MCM proteins

Crystallographic studies of the N-terminal domain

Despite the attempts of many laboratories worldwide to obtain crystals of MCM proteins diffracting to high resolution, the complex has been rather resilient to crystallization. The first breakthrough was represented by the crystal structure of the N-terminal domain of the MCM protein from *M. thermautotrophicus* (*Mth*MCM-N; Fletcher, 2003) which revealed a double-hexameric assembly with a head-to-head configuration (Fletcher *et al.*, 2003) (Figure 2A). Each monomer folds into three subdomains with the OB fold-like subdomain C (sC) responsible for hexamer formation, subdomain B (sB) found at the double-ring interface and forming a conserved Zn domain and subdomain A (sA) located in the peripheral belt of the MCM ring, giving the structure its characteristic "dumbbell" shape. A large central channel is positively charged and is large enough to accommodate both ssDNA and dsDNA. The channel is constricted by six βhairpins (NBH) that include positively charged amino-acid residues shown by mutagenesis to be involved in the interaction with DNA (Fletcher *et al.*, 2003; McGeoch *et al.*, 2005). More recently, the crystal structure of the equivalent domain from *S. solfataricus* was determined (*Sso*MCM-N (Liu *et al.*, 2008)). The protein forms a single hexameric ring:

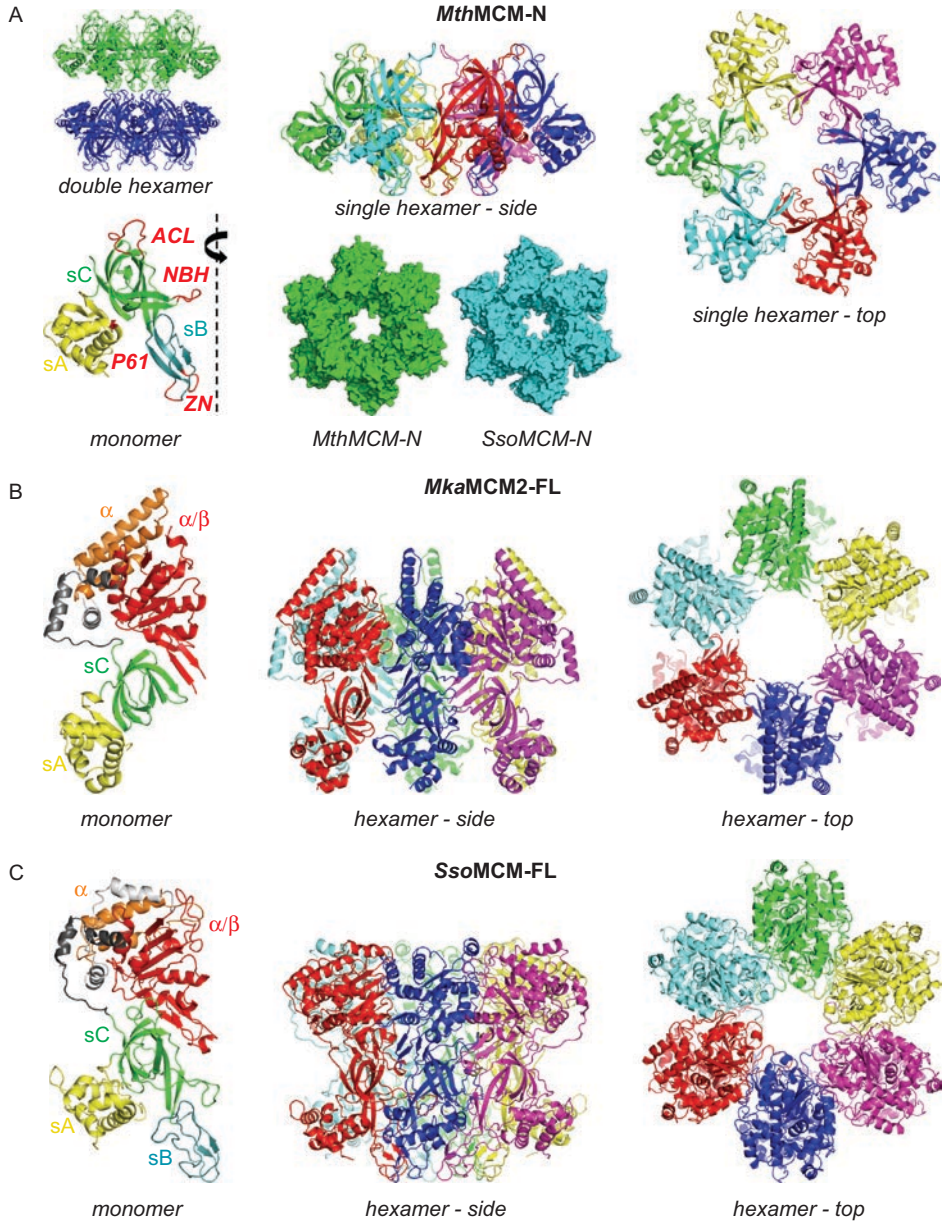

Figure 2. Crystallographic structures of MCM proteins. (A) The crystal structure of the N-terminal domain of *M. thermautotrophicus* MCM (Fletcher *et al.*, 2003). The side view of the double hexameric configuration is shown with the two hexamers highlighted in green and blue, respectively. The top and side views of the *Mth*MCM Nterminal hexamer are color-coded according to the different subunit. The structure contains a channel which is large enough to fit dsDNA. Each monomer folds into three subdomains: subdomain A (sA, yellow) is located in the outer belt of each ring; subdomain B (sB), in teal, coordinates a zinc atom that is located at the interface between the two hexameric rings; subdomain C (sC) is responsible for the inter-subunit contacts within a hexamer and is shown in green. The structural elements that are important for the activity of the protein are highlighted in red. A dashed line indicates the position of the six-fold axis around which the hexameric complex forms. Surface representations of the *Mth*MCM-N and *Sso*MCM-N (Liu *et al.*, 2008) hexamers (in green and cyan, respectively) highlight the differences between the two structures. (B) The crystal structure of the MCM2 of *Methanopyrus kandleri* (Bae *et al.*, 2009). The sequence is unusual for an MCM protein, as it does not have subdomain B and the C-terminal domain, and lacks a number of conserved elements, such as ACL, NBH, Walker A, Walker B, PS1BH and R finger. The protein is a monomer in solution and in the crystal. The AAA+ domain folds into two subdomains (α/β in red and α in orange) and is connected to the N-terminal domain through a long loop wrapping around the AAA+ domain and a couple of helices (dark gray): the N–C connection is intertwined with another helical extension (light gray – partially disordered in the structure) that links the α/β and α domains. A hexamer can be built based on the symmetry of the *Mth*MCM-N structure (C) The crystal structure of the MCM of *Sulfolobus solfataricus* (Brewster *et al.*, 2008). The C-terminal domain is disordered in the crystal and no atomic model could be built. The color code for the monomer is the same as in panels 2A and 2B. Although the protein can form a hexamer in solution, a monomer is present in the crystal; a hexamer can be built based on the symmetry of the *Mth*MCM-N structure.

although the topology of the monomer is similar, there are subtle differences in the overall hexameric architecture, so that when the two hexamers (*Mth*MCM-N and *Sso*MCM-N) are overlapped, the relative position of the corresponding subdomains is significantly shifted. The equivalent NBH is considerably longer in the *Sulfolobus* sequence (five extra residues) and contains positively charged amino acids; however the conformation of the loop is such that it folds

back onto itself making the channel just marginally smaller in *Sso*MCM-N than in *Mth*MCM-N (17 rather than 23 Å). The difference is however large enough to make it unlikely that dsDNA could be accommodated in the channel (Figure 2A).

Crystallographic studies of (almost) full-length MCM proteins

Only very recently has structural information for two full length MCM proteins become available, finally providing a model for a MCM AAA+ domain.

The crystal structure of an unusual MCM protein from *Methanopyrus kandleri* was obtained (Bae *et al.*, 2009) (Figure 2B). *M. kandleri* possesses two open reading frames that resemble MCM, one coding for a polypeptide whose sequence conforms to the canonical architecture of MCM proteins and comprises all the structural and functional elements of an active helicase (*Mka*MCM1), whereas the second (*Mka*MCM2) lacks subdomain B of the N-terminal domain (the Zn domain), the C-terminal WH domain, as well as the majority of functional motifs, such as Walker A, Walker B, sensor I and R finger. *Mka*MCM2 is always monomeric in solution and displays no helicase activity. However the structure of *Mka*MCM2-FL determined to 1.9 Å resolution provides a high resolution structural framework for understanding MCM architecture. A hexameric model could be built by overlapping the *Mka*MCM2-FL monomer onto the *Mth*MCM-N hexameric structure (Figure 2B). Minor clashes between the Walker A and R finger of neighboring subunits can be easily relieved by a small re-orientation of the AAA+ domains to reconfigures a putative ATPase site.

At the same time the crystal structure of the MCM protein from *S. solfataricus* was determined to 4.35 Å resolution (*Sso*MCM-FL), providing a low resolution picture of an active MCM helicase (Brewster *et al.*, 2008). Despite the fact that *Sso*MCM can form a hexamer in solution, a monomer is seen in the crystal (Figure 2C). Two constructs have been crystallized, one with the full-length sequence and one with a deletion of the C-terminal domain; however only weak and partial electron density can be detected for the C-terminal domain, and no atomic model could therefore be built. The N-terminal domain of the full-length *Sso*MCM structure is similar to the two MCM-N structural models, but when the hexameric arrangement seen in the *Sso*MCM-N structure is extrapolated to the full-length model, steric clashes arise between the AAA+ domains. However, when the symmetry of *Mth*MCM-N is applied, the AAA+ domain assumes a configuration that is compatible with the formation of ATPase sites at the subunit interface. This suggests that the differences seen in the hexameric architecture between the structures of the two N-terminal domains may not be species-specific but rather reveals a conformational plasticity that could have functional significance.

When the two structures of the MCM AAA+ domains are superposed, the similarity between the two folds is evident, particularly considering the relatively low level of sequence identity due to the peculiarity of the *Mka*MCM2 sequence. It has to be stressed that the *S. solfataricus* AAA+ domain has been built into a 4.35 Å map. At such low resolution it is extremely difficult to correctly fit the sequence, even when the overall topology is partially known. Only a polyalanine model has therefore been refined, yet the sequence was assigned. A detailed comparison between the two AAA+ domains reveals that in certain regions the sequence of the *Sso*MCM-FL structure may be out of register (in particular the N-C linker, part of the Walker A, α/β-α linker and α domain). Despite these ambiguities, this is the first atomic model of a functional MCM protein and reveals a number of features that have a biological significance.

In addition to the canonical AAA+ sequence motifs, a feature called the "glutamate switch" (E switch) has recently been described that may provide a way to regulate the ATPase activity of this superfamily of proteins (Zhang and Wigley, 2008). A polar residue (typically an asparagine) can form a hydrogen bond with the glutamate residues of the Walker B motif (D**E**xD), which in turn is critically involved in orienting and polarizing a water molecule for the nucleophilic attack on the γphosphate of the ATP substrate (Figure 3A). This ensures that the ATPase activity is switched off unless the system is correctly assembled. The RuvBlike and HslUV/ClpX families possess an unusual version of the E switch, involving a Thr/Ser residue shifted by two residues from the standard Asn position. A three-dimensional model based on the high resolution MCM structure from *M. kandleri* suggests the intriguing possibility that the switch may be even more complex in MCM proteins, involving a conserved Thr/Ser residue located at the end of the second strand of the AAA+ fold core, in the canonical position of the usual Asn, as well as an additional Arg/Lys located two residues further along the chain, where the RuvB-like switch is positioned. This residue is highly conserved either as an arginine or lysine, with the exception of MCM2 proteins which have a conserved glutamine. The equivalent residue (K366) has been mutated in *Sso*MCM and the mutation completely abolishes the helicase activity and displays a modest reduction in the rate of ATP hydrolysis (Moreau *et al.*, 2007). This is consistent with a role in regulating the helicase activity and coupling ATP hydrolysis with substrate binding. The loop containing both the Thr/Ser and the Arg/Lys residues immediately precedes helix 2 and the h2i that is postulated to interact directly with the DNA inside the channel, providing a possible route to switch on the ATPase activity upon substrate binding (Figure 3B). In the transcriptional enhancer PspF a similar connection has been observed between the E switch and two loops (L1 and L2) that have been proposed to bind to the σ54 transcription factor (Rappas *et al.*, 2005); these loops are topologically equivalent to the h2i and PS1BH elements in MCM.

Modeling the MCM hexamer

To understand the action of multimeric helicases it is essential to have an experimental model for the active hexamer, as ATP binding and hydrolysis occur at the subunit interface and therefore the force generation driving the mechanical

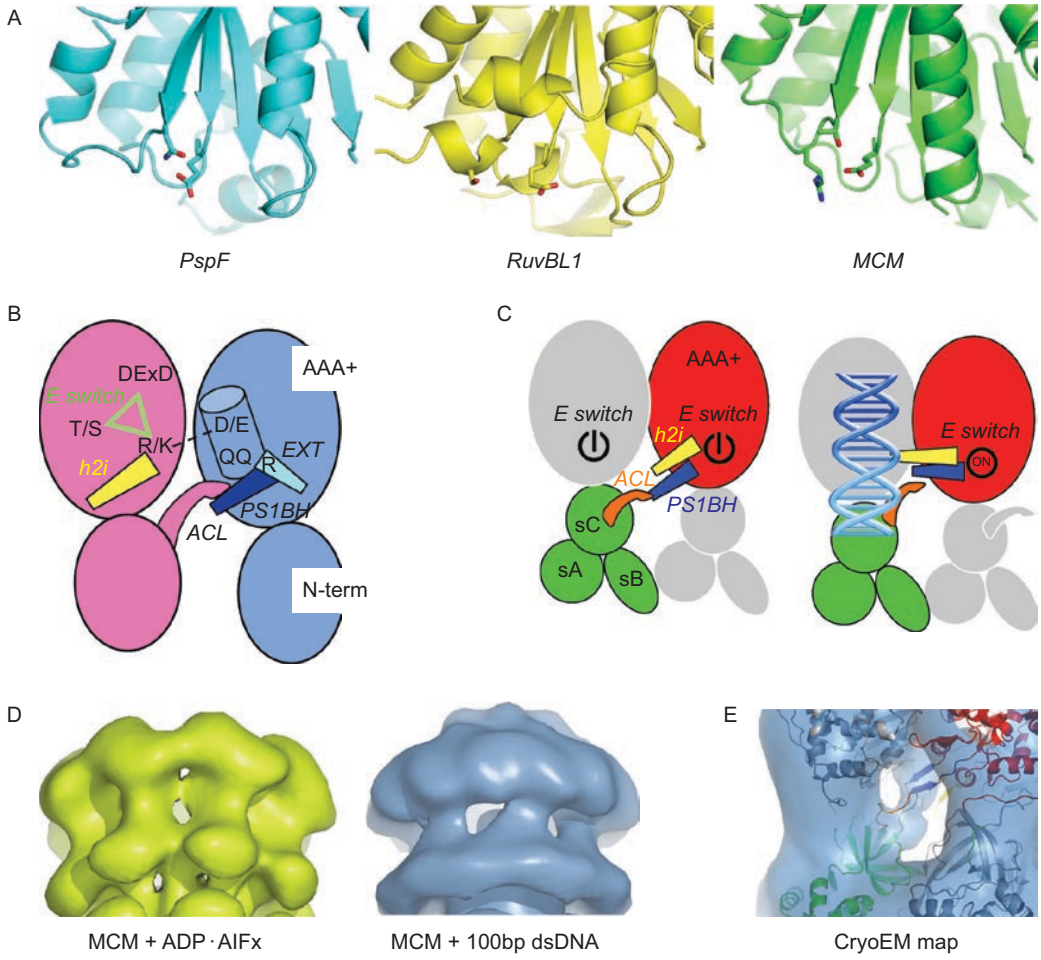

Figure 3. Intra- and inter-subunit communication. (A) The glutamate switch in a variety of AAA+ ATPases (Zhang and Wigley, 2008). On the left panel the canonical switch as seen in the crystal structure of PspF (cyan) bound to the ATP substrate (Rappas *et al.*, 2006); the switch consists of a conserved asparagine (located at the very end of the β-strand following the Walker B motif) holding the Walker B glutamate (DExD) in an inactive conformation in the absence of the substrate. In the RuvB-like family (yellow) the switch involves a threonine or serine residue located two residues downstream from the usual Asn position, as illustrated by the crystal structure of the human RuvBL1 protein (Matias *et al.*, 2006). The *Mka*MCM2 crystal structure (green) shows that the position corresponding to the E switch in PspF is occupied by a residue that is a threonine or serine in the MCM family, whereas an arginine or lysine is located two residues downstream, in the position of the RuvBL1 E switch (Bae *et al.*, 2009). Since the inactive *Mka*MCM2 does not have a Thr/Ser in the first position and has a histidine substituting the Walker B glutamate, the corresponding residues have been mutated in the figure to illustrate the situation found in the majority of the MCM sequence. (B) A schematic diagram illustrating the interactions between the ACL loop of one subunit with the PS1BH and helix 3 (the helix sandwiched between the Walker B and PS1BH) of an adjacent subunit. The interaction involves the absolutely conserved QQ residues at the C-terminal end of the helix (Q423-Q424 in *S. sulfolobus*); close to that end of the helix is a conserved arginine at the base of the EXT loop (R331 in *S. sulfolobus*). The other end of the helix contains a conserved negatively charged residue that is in hydrogen bonding distance from an Arg/Lys residue on an adjacent subunit, which we propose to be part of the E switch. (C) A schematic diagram illustrating the putative cross-talk between the ACL, PS1BH, H2i loops and the E switch, in the absence and presence of DNA bound to the channel. In the absence of the DNA substrate the ACL loop of the N-terminal domain of one subunit makes an interaction with the PS1BH loop of an adjacent subunit, as demonstrated biochemically (Barry *et al.*, 2009) and visualized by EM (Bae *et al.*, 2009). Through a network of hydrogen bonds and direct peptide links, this forces the E switch in the OFF configuration, engaging the Walker B in an orientation that is not productive for hydrolysing ATP. When the DNA substrate is bound, a rigid body rotation moves the h2i towards the center of the channel, with the consequent repositioning of the PS1BH, breaking therefore the interaction with the ACL. This leaves the lateral holes empty and presumably can transmit a signal to the E switch (positioned at the very base of the h2i) to activate the hydrolysis of ATP. (D) A side tilted view of a ring of the double heptameric structure in the presence of ADP·AlF_x (yellow) shows a lateral hole divided in two apertures by the presence of an isthmus of electron density. A tilted view of the upper ring of the double hexameric structure, in the presence of 100 bp of dsDNA (blue) shows a single lateral hole per subunit (Costa *et al.*, 2006a). (E) The interaction between ACL and PS1BH as visualized by electron microscopy. The crystal structures of *Mth*MCM-N (Fletcher *et al.*, 2003) and the AAA+ domains from *Sso*MCM-FL (Brewster *et al.*, 2008) were fitted onto a cryoEM map of *Mth*MCM bound to a long dsDNA molecule, after applying six-fold symmetry (Costa *et al.*, 2008). All the atomic models are in gray, with the exception of on N-terminal domain (green) and the adjacent AAA+ domain (red); the ACL loop is highlighted in orange, the PS1BH in blue and the h2i in yellow. The same isthmus of electron density disrupting the lateral holes is seen in all the EM reconstruction of MCM, with the exception of the double hexameric structures (Figure 4C).

work relies on conformational changes involving neighboring subunits. Unfortunately, whereas the N-terminal domains crystallize as hexamers or double hexamers, the structures of the full-length proteins are monomeric in the crystals and therefore we do not have an experimentally derived picture of a MCM active site.

However, it is possible to assemble a hexamer by applying to the atomic models of the full-length proteins the

hexameric symmetry derived from the *Mth*MCM-N structure (Figures 2B and 2C). This operation places the AAA+ domain in a configuration compatible with the formation of an ATPase site at the interface between subunits. Although the complex interplay between amino-acid conformations that underlines the subtleties of ATP binding and hydrolysis are obviously lost in such a coarse modeling exercise, it is possible to gain some biological insight from the overall architecture of the MCM hexamer.

The first striking observation is that the unusual relative position of the α/β and α domains in MCM proteins strongly affects the architecture of the hexamer and makes the inter-subunit interactions strikingly different from the majority of AAA+ ATPases (Figure 1D). Typically the α subdomain is located at the interface between two neighboring subunits and is oriented in such a way as to contribute the critical sensor II arginine to the active site in *cis* (i.e. to the ATP active site that contains the Walker A, Walker B and sensor I from the same subunit). The presence in MCM of an insertion at the end of the α/β domain dramatically repositions the α domain so that it is now located at the other subunit–subunit interface. This change extensively influences the nature of the inter-subunit contacts within the AAA+ ring assembly. The consequence of such rearrangements is to make the sensor II arginine of the α domain interact with the ATP of the next-neighboring subunit and therefore to act in *trans*, rather than *cis*. This unusual arrangement had already been predicted based on a detailed comparison between AAA+ structures (Erzberger and Berger, 2006), and confirmed by biochemical experiments demonstrating that indeed the sensor II arginine contributes in *trans* to the catalytic activity (Moreau *et al.*, 2007).

Both the hexameric model based on the *Sso*MCM-FL structure (Brewster *et al.*, 2008) and a hybrid model built with the *Mth*MCM-N and the AAA+ from *Mka*MCM2 (Bae *et al.*, 2009) reveal an interaction between the PS1BH of one subunit and a loop in the N-terminal domain of an adjacent subunit (Figure 3E). This loop is very highly conserved in MCM proteins, but absent in the *Mka*MCM2 structure; it has been variably described as β7-β8 loop in the *Mth*MCM (Fletcher *et al.*, 2003; Sakakibara *et al.*, 2008) and *Mka*MCM2 (Bae *et al.*, 2009) systems and ACL (Barry *et al.*, 2009) or L207 (Brewster *et al.*, 2008) in the *Sso*MCM context. Here we will use the ACL nomenclature as it is not dependent on a species-specific numbering scheme. The interaction provides a mean of communication between the N-terminal and Cterminal tiers and also between adjacent subunits.

The hexameric model also reveals that a conserved loop within the AAA+ domain (EXT) lines the lateral holes at the interface between subunits (Brewster *et al.*, 2008). Mutagenesis of residues within this loop affect helicase activity (Moreau *et al.*, 2007; Brewster *et al.*, 2008).

Polymorphism in the MCM complex, as detected by electron microscopy

In parallel to the crystallographic characterization, numerous studies using electron microscopy have provided a low resolution picture of the MCM complex from either negatively-stained or frozen samples. Electron microscopy (EM), although not achieving atomic resolution, is a powerful tool for studying conformational changes and dynamic behavior of large assemblies. EM analysis of MCM proteins has produced a number of snapshots of the complex in different conformational states, which have complemented the crystallographic analysis to tackle the degree of structural polymorphism observed for these complexes and to explore the gamut of molecular configurations adopted during the catalytic process.

Initial electronmicroscopy studies carried out on the *Mth*MCM complex revealed heptameric (Yu *et al.*, 2002) as well as hexameric (Pape *et al.*, 2003) hollow single rings (Figure 4A). In both structures the ring has two tiers and the central channel is wide enough to bind dsDNA. In addition to the channel, the hexameric structure shows the presence of a large internal cavity that is restricted at two places along the axis, as well as lateral holes (two per each subunit–subunit interface). Both the large cavity and the presence of lateral holes are reminiscent of the structure of the SV40 Large T antigen (LTag), which similarly acts as a DNA helicase (Li *et al.*, 2003). Based on a preliminary fitting of atomic models for the N-terminal and AAA+ domains, a position for the C-terminal domain was suggested at the interface between the Nterminal and AAA+ domains, in a belt of unassigned density (Figure 4A).

Subsequent work confirmed the variability in ring stoichiometry, with numerous samples showing mixed populations of hexameric and heptameric top-views, side views corresponding to double rings, as well as a few open rings (Chen *et al.*, 2005; Gomez-Llorente *et al.*, 2005; Costa *et al.*, 2006b; Jenkinson *et al.*, 2009). Depending on the protein preparation, the construct used or the conditions of incubation, one or the other species was predominant within the particle population. By a careful exercise of particle picking, a double-hexameric reconstruction was generated showing a central channel, with two peaks of density in the middle, and large lateral holes (Gomez-Llorente *et al.*, 2005) (Figure 4B).

A systematic study of the complex in the presence of various substrates showed that treatment with nucleotides (AMP-PNP and ADP·ALF_x) and/or short stretches of dsDNA (100 bp) led to the stabilization of doublering structures, whereas untreated protein tends to form single hexameric or heptameric rings (Costa *et al.*, 2006b). More specifically, the presence of nucleotide analogs triggers the formation of double heptamers, while a consistent shift from a double heptameric to a double hexameric arrangement was shown to be associated with DNA binding (with or without nucleotide). Single particle reconstructions were carried out for two of these samples (Costa *et al.*, 2006a) (Figure 4C). The protein treated with 0.5 mM ADP·ALF_x produced a double heptameric structure, with a wide and empty central channel and symmetry between the two rings. By contrast the sample treated with 100 bp of dsDNA showed a remarkable asymmetry between the two rings, with one ring being capped at the top and with the central channel occupied by

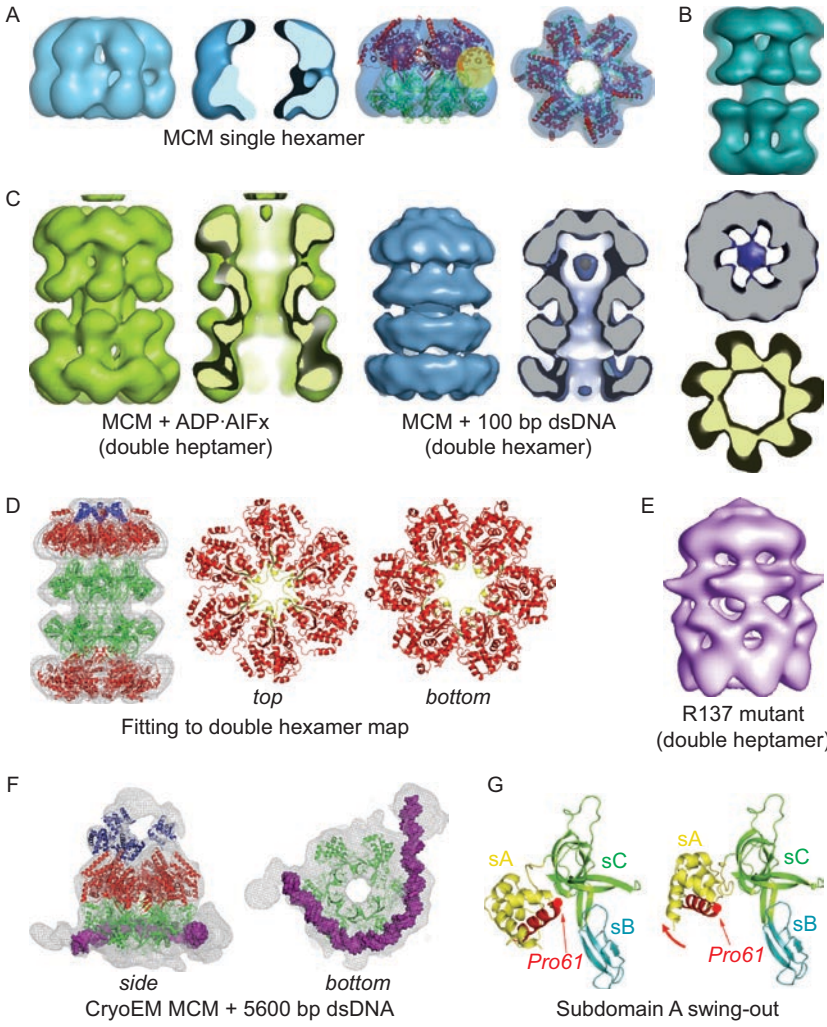

Figure 4. Electron microscopy structures of MCM proteins. (A) Three-dimensional model for MCM from *Methanothermobacter thermautotrophicus* from negative stained electron microscopy data (Pape *et al.*, 2003). Six-fold symmetry was applied throughout the reconstruction process. The structure shows a single hexameric ring with a large central channel connected to the outside by lateral holes (two at each subunit interface). A yellow circle represents the position of the C-terminal domain as predicted in Pape *et al.* (2003). Subsequent fitting of the *Mka*MCM2 crystal structure to the electron density (Bae *et al.*, 2009) does not agree with this prediction. (B) Three dimensional reconstruction of a double hexamer of *Mth*MCM from a negatively stained sample (Gomez-Llorente *et al.*, 2005). (C) Surface rendering of double-ring structures of *Mth*MCM from negatively stained samples in the presence of substrates (Costa *et al.*, 2006a). On the left (in yellow) a side view of the ADP·AlF_x treated double heptamer showing a head-to-head double-ring configuration with lateral holes, and a slice through the center, showing the large uninterrupted channel through the molecule. On the right a side view and a slice-through of the protein treated with 100 bp of dsDNA (in blue). The protein forms a double hexamer with a strong asymmetry between the two rings. The upper ring is closed on top by a cap, and has the central channel obstructed by a peak of electron density, whereas the bottom ring is hollow and uncapped. Seven-fold and six-fold symmetry, respectively, were applied to the reconstructions. A section through the middle of the upper rings in both structures (right panel) shows the striking features of the dsDNA-treated protein, with a peak of electron density connected to the ring by six protrusions. (D) Fitting of atomic models to the double hexameric EM map. The dodecameric *Mth*MCM-N atomic coordinates (Fletcher *et al.*, 2003) were fitted to the central part of the molecule (in green). For the AAA+ tier of the lower ring the *Sso*MCM-FL structure (Brewster *et al.*, 2008) arranged into a hexamer based on the *Mth*MCM-N model matches well the electron density (in red). To fit the upper ring, a rigid body rotation needs to be applied to the AAA+ domain. The conformational change is illustrated by the two central panels, showing the top and bottom AAA+ hexamers. The reorientation of the AAA+ domain in the top ring brings the helix-2 insertions (in yellow) towards the central channel, fitting into the lateral protrusions. The PS1BH elements are shown in green. (E) A side view of the 3D model for a *Mth*MCM mutant where arginine R137, located at the very bottom of subdomain B (in the Zn motif), is mutated to an alanine (Jenkinson *et al.*, 2009). The mutant stabilizes a double heptameric conformation, showing a remarkable swing-out movement of subdomain A. Similarly to the dsDNA-bound structure, the two rings display a large asymmetry, with a cap on the upper but not on the lower ring. (F) Cryo-EM structure of *Mth*MCM bound to long (5600 bp) dsDNA molecules (Costa *et al.*, 2008). The 3D reconstruction was obtained without applying any symmetry constraints. A side view shows the fitting of the N-terminal domain from the *Mth*MCM-N structure (green; Fletcher *et al.*, 2003), the AAA+ domain from the *Mka*MCM2-FL structure (red; Bae *et al.*, 2009), and the C-terminal winged-helix domain from the *Pyrobaculum aerophilum* Cdc6 structure (blue; Liu *et al.*, 2000). The location of the C-terminal domain was confirmed by carrying out a similar 3D reconstruction for a mutant lacking the C-terminal domain. 80 bp of dsDNA (violet) can be fitted to a belt of higher electron density around the Nterminal domain, ending on two filamentous protrusions. To avoid steric clashes and to better fit the electron density, in four out of the six subunits of the ring a conformational change in subdomain A has been modeled. (G) An illustration of the swing-out movement of subdomain A; the movement was extrapolated from the differences between the MthMCM-N (Fletcher *et al.*, 2003) and SsoMCM-N (Liu *et al.*, 2008) crystal structures. The putative recognition helix and the position of the proline that is mutated in the *bob1* mutant are shown in red.

a peak of electron density connected to the body of the protein by six "blades" (Figure 4C). The lack of electron density accounting for the presence of a DNA molecule inside the channel is a common occurrence in similar studies, due to the combination of the mobility of the DNA and the negative staining technique used.

All of the structures described so far have been obtained from negatively-stained samples, and by applying 6fold or 7fold symmetry to the particles. Although this procedure helps to increase the signal-to-noise ratio, it may mask deviations from symmetry that may be intrinsically present in the complex as there is no reason to believe that all the six subunits contain the same nucleotide *in vivo* and they cannot bind DNA in a perfectly symmetric fashion. So the models obtained are indicative of the range of conformations available rather than being exact physiological states of the potentially asymmetric MCM complex. Indeed there is biochemical evidence for asymmetric modes of ATP hydrolysis, with the possibility of tolerating up to three inactive sites (Moreau *et al.*, 2007). Similar observations on the eukaryotic recombinant MCM2-7 complex has suggested a model with some of the site constitutively inactive (Davey *et al.*, 2003; Schwacha and Bell, 2001). However the extremely high degree of sequence conservation within the motifs important for helicase activity in all the MCM2-7 subunits suggests not only ATP binding, but also that hydrolysis can occur at each ATPase site.

A more recent EM analysis of *Mth*MCM was carried out in the presence of long dsDNA molecules (5600 bp). Data were collected from a frozen sample, and a map was refined without applying any symmetry constraint (Costa *et al.*, 2008). Single hexameric rings were visualized, with dsDNA wrapping around the outer belt of the ring and interacting with subdomain A at the N-terminus (Figure 4F).

Conformational flexibility in the AAA+ domain

The asymmetry between the two rings in the double hexamer structure of the dsDNA-treated MCM protein suggests that the protein undergoes a conformational change, so that distinct configurations can be modeled for the AAA+ domain (Costa *et al.*, 2006a).

Modeling the AAA+ domain onto the upper ring shows that the "blades" can be explained by the h2i insertions protruding into the middle of the channel (Figure 4D). These features are neither present in the lower ring of the hexameric complex nor in the heptameric complex, suggesting that a conformational change must take place to move these loops away from the central channel. This change can be modeled by a rigid body rotation of the AAA+, causing a re-orientation of the PS1BH and h2i elements away from their position in the upper ring, so that they assume a more lateral position, in between adjacent subunits. A nucleotide-dependent rigid body rotation of the AAA+ which influences the position of the PS1BH elements has been observed for the SV40 LtAg helicase (Gai *et al.*, 2004). What triggers the movement in the MCM helicase is less clear. Nucleotide binding seems less likely here, as the lower ring of the dsDNA-treated sample (in the absence of nucleotide) and both rings of the ADP·ALF_x-treated sample display a very similar orientation. One possibility is that the trigger for the conformational change is the binding of the DNA substrate to the central channel.

Inter- and intra-subunit communication

When the h2i (and the closely connected PS1BH) are away from the center of the channel, the lateral position of the PS1BH generates an interaction with the ACL loop in the N-terminal domain of an adjacent subunit (Brewster *et al.*, 2008; Bae *et al.*, 2009; Barry *et al.*, 2009) (Figure 3E)

The hexameric models shows that loop may not only interacts with the PS1BH of the adjacent subunit but also with the helix sandwiched between the Walker B and PS1BH motifs, and in particular with two invariant glutamine residues at the C-terminal end of the helix. Mutation of these glutamine residues in *Sulfolobus* (Q423A-Q424A) has been shown to abolish both the ATPase and the helicase activity (Moreau *et al.*, 2007) (Figure 3B). At the other end of the helix is a partially conserved Asp/Glu which is in hydrogen bonding distance with residues on the adjacent subunit, particularly with the Arg/Lys that we suggest may be part of the MCM E switch (Figure 3A). Close to the QQ motif is another residue whose mutation has been shown to have a similar phenotype (R331 in *S. solfataricus*, at the base of the EXT loop). Interestingly the mutational study of Q423-Q424 and R331 (Moreau *et al.*, 2007) shows that both residues act in *trans*, consistently with the model proposed in Figure 3B. The importance of the ACL loop and its critical role in mediating the inter-subunit interactions to ensure cooperativity is well supported by biochemical analysis both in the *S. solfataricus* (Barry *et al.*, 2009) and the *M. thermautotrophicus* systems (Sakakibara *et al.*, 2008).

When a full-length hexameric model is fitted into the EM electron density, the ACL and PS1BH loops interact in such a way as to form a bridge that disrupts the large lateral hole at the interface between subunits (Figure 3E). In the majority of the EM reconstructions the holes are similarly disrupted when the central channel is empty, whereas they are large and continuous when there is electron density in the central channel (presumably due to the DNA substrate binding) (Figure 3D).

The complex network of interactions between ACL, PS1BH, h2i and the E switch may provide the molecular framework to understand the mechanism of action of MCM helicases. In *Mth*MCM and the MCM helicases whose ATPase activity is stimulated by DNA binding, a speculative model can be envisaged in which in the absence of the DNA substrate the AAA+ subunit is rotated so that the PS1BH is away from the central channel and is "trapped" by the interaction with the ACL (Figure 3C). This interaction not only would disrupt the lateral holes, but also contributes to the hydrogen bond network that, through the E switch, holds the Walker B glutamate in an "inactive" conformation (Figure 3B). The presence of DNA in the central channel pulls the h2i and PS1BH towards the middle, breaking the E switch and therefore enhancing the ability of the active site to hydrolyze

ATP. It has to be stressed that the isthmus of electron density that interrupts the lateral holes (deriving from interaction ACL-PS1BH) is also present in the double heptamer structure obtained when *Mth*MCM was treated with ADP·ALF$_x$ (Figure 3D), suggesting that binding of the nucleotide alone is not able to trigger the movement of PS1BH in the absence of the DNA substrate. In *Sso*MCM where the ATPase activity is constitutively elevated and the effect of DNA stimulation is lower, there may be a different activation mechanism, more critically dependent on the nucleotide state, consistent with the fact that the ACL–PS1BH interaction is decreased in the presence of ATP (thereby "releasing" the PS1BH and switching on the activity). Intriguingly in the *Sso*MCM system there is biochemical evidence for a cross-talk between the Nterminal β-hairpins and the ACL loop, suggesting that in the absence of ACL, the NBH are in an inhibitory conformation, that can be modulated via the ACL loop, in a nucleotide-dependent manner (Barry *et al.*, 2009).

Conformational flexibility in the N-terminal domain

Activation of the initiation of DNA replication in yeast requires a phosphorylation event catalyzed by the Dbf4-dependent kinase Cdc7 (DDK). The need for DDK can be bypassed by a mutation of MCM5 (known as the *bob1* mutant) where proline 83 in the N-terminal domain is mutated to a leucine, causing mutant cells to enter S phase prematurely (Hardy *et al.*, 1997). Crystallographic analysis of a *Mth*MCM mutant harboring the equivalent mutation (P61L, Figure 2A) shows a small but significant opening of subdomain A (Fletcher *et al.*, 2003). The result suggests that the effect of DDK-mediated phosphorylation is to open up subdomain A, and the mutation somehow "releases" the interaction between subdomain A with the adjacent subdomain C, making it more prone to swing out. This was confirmed by an EM structure of the *Mth*MCM N-terminal domain alone, showing a dramatic conformational change of subdomain A (Chen *et al.*, 2005). Moreover, when the two N-terminal domain structures are compared, the variations in hexameric assembly are such that in the *Sso*MCM-N structure subdomain A is a more open position.

The cryo-electron microscopy structure of MCM bound to a long dsDNA substrate (Costa *et al.*, 2008) suggests that the movement of subdomain A is functionally necessary for a novel interaction with nucleic acid, where dsDNA wraps around the N-terminal domain (Figure 4F). Structural comparisons reveal that subdomain A displays some similarity with proteins belonging to the family of helix-turn-helix transcription factors. In the context of the "closed" conformation observed in the *Mth*MCM-N crystal structure the putative recognition helix (which starts with the proline residue mutated in the *bob1* mutant) is packed against subdomains C and B. The swing-out movement of subdomain A, previously proposed on the basis of crystallographic and EM evidence (Fletcher *et al.*, 2003; Chen *et al.*, 2005) would allow for productive interaction with dsDNA (Figure 4G). Most likely this represents an initial site of interaction between MCM and dsDNA, distinct from the canonical one in which the protein ring encircles the nucleic acid, prior to the loading and activation of the complex to function as the replicative helicase (Costa and Onesti, 2008) (Figure 5A). This is compatible with the biochemical observation that yeast MCM2-7 complexes display two modalities of chromatin association, defined by the different sensitivity to high salt washes: thus the more labile complex ("associated" MCM) may have DNA wrapped around the N-terminal domain, whereas the more stable complex ("loaded" MCM) would encircle the DNA (Bowers *et al.*, 2004).

A similar conformational change involving a large swing-out movement of subdomain A has also been directly observed in the 3D reconstructions of *Mth*MCM mutant forms where arginines 137 and 160 in subdomain B were mutated to alanine (Jenkinson *et al.*, 2009). The two mutants

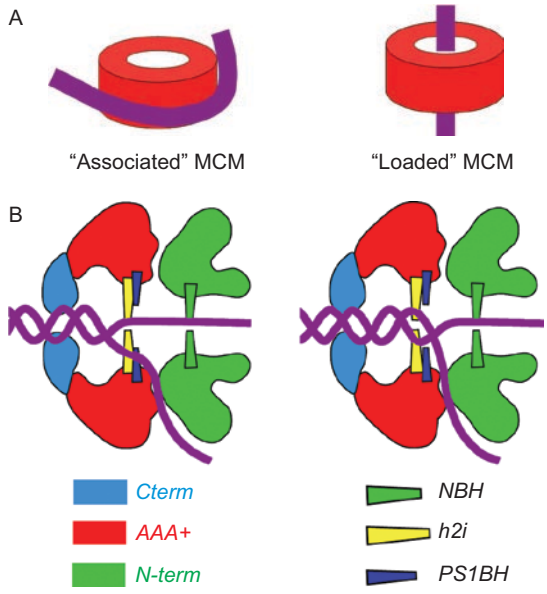

Figure 5. Mechanism of action of MCM proteins. (A) Two possible modes of interaction between MCM proteins and DNA. In the first mode, the DNA wraps around the N-terminal domain, wedged between the "open" conformation of subdomain A and subdomain B. The interaction has been visualized in the cryoEM 3D reconstruction of MCM bound to long dsDNA segments (Figure 4F; Costa *et al.*, 2008), and may represents a model for the "associated" MCM proteins, that can be easily washed from chromatin. In the second mode the DNA is threaded through the hexameric ring channel, with the protein ready to act as helicase. This may represent a model for the (loaded) salt-resistant MCM-chromatin association. (B) Schematic diagrams illustrating two possible configurations of the helicase during elongation, both conforming to the "ploughshare" model proposed by Takahashi *et al.* (2005). The MCM model is based on the upper ring of the dsDNA-treated double hexamer (Figure 4C; Costa *et al.*, 2006a), with the C-terminal domain (blue) forming a cap on top of the AAA+ domain (red), and the h2i (yellow) and PS1BH (dark blue) extending into the central channel. The orientation of the molecule with respect to the fork is based on the observation that *Sso*MCM moves along DNA with the C-terminal motor domain "leading", i.e. being closer to the dsDNA end of the fork (McGeoch *et al.*, 2005). In both diagrams one of the separated strands is extruded through a lateral hole whereas the second strand exits through the N-terminal hexamer (green). In the left panel the h2i elements act as ploughshare, separating the two DNA strands, whereas in the right panel the h2i are involved in dsDNA translocation and the holes directly act as the ploughshare.

(R137A and the double mutant R137A-R160A) form double and single heptameric complexes, respectively, and display reduced DNA binding and helicase activity. The corresponding EM structures show a swing-out movement of subdomain A, which is particularly striking in the double heptameric map of the R137A mutant (Figure 4E). In the 3D model derived from cryoEM data (Costa *et al.*, 2008) (Figure 4G) the dsDNA is sandwiched between the open subdomain A and subdomain B and is close to the R137 position, providing an explanation for the reduced DNA binding ability of the R137A mutant (Figure 4F).

The C-terminal domain

The location of the C-terminal domain within the overall architecture of the MCM proteins has been uncertain, and more than one position been proposed. An initial attempt at locating the C-terminal domain was done in the interpretation of the first hexameric EM structure (Pape *et al.*, 2003), where after fitting the N-terminal domain and a AAA+ domain derived from the RuvB translocase, some density was found in the peripheral belt between the two domains and assigned to the C-terminal domain (Figure 4A). This result was supported by biochemical data from FRET experiments that confirmed that the position was compatible with a measured distance of 75 Å between the N-terminal and C-terminal domains (McGeoch *et al.*, 2005). However, that model relied entirely on the fitting of a homologous AAA+ domain and more recent analyses based on more faithful models for the AAA+ are not consistent with this location.

Subsequent EM work on the dsDNA-treated double hexamer of *Mth*MCM offered the advantage of showing the position of the h2i protruding into the channel and therefore provided a more accurate positioning of the AAA+ domain (Costa *et al.*, 2006a). It was therefore clear that the previously proposed location of the C-terminal domain needed to be revised. In the double hexamer map, the only unexplained electron density, once the N-terminal and AAA+ domains are fitted, is at the cap which is present on one of the two rings. This density was therefore assigned to the C-terminal domain (Figures 4C and 4D). The fact that the cap is often absent in 3D reconstructions suggests that the C-terminal domain undergoes either a conformational change or an order-disorder transition. The new proposed position is still consistent with a distance of 75 Å from the Nterminal domain (McGeoch *et al.*, 2005).

More recent work by cryo-electron microscopy on a frozen sample bound to long dsDNA segments showed the presence of some protrusion on the top of the AAA+ domain (Figure 4F). A 3D reconstruction was carried out from cryoEM data for a sample without the C-terminal domain and compared with the full-length structure, definitely confirming the location of the C-terminal domain at the protrusions on top of the AAA+ ring (Costa *et al.*, 2008). The fact that only four out of six C-term domains could be built in the electron density further confirms its intrinsic mobility.

One more location for the C-terminal domain has been proposed in the middle of the central channel based on the presence of unexplained electron density in the double hexameric 3D reconstruction (Gomez-Llorente *et al*. 2005). Although it is possible that such a peak may be due to contaminating DNA, weak and unexplained electron density is seen in the crystal structure of the full-length *Sso*MCM protein at a similar position (Brewster *et al.*, 2008).

The archaeal C-terminal domain has been predicted to fold into a WH motif. This raises the obvious question of whether the domain does bind DNA. Although no clear biochemical evidence for the isolated domain has been shown (Jenkinson and Chong, 2006; Barry *et al.*, 2007; Pucci *et al.*, 2007), the ordering of the domain into the cap upon dsDNA interaction (Costa *et al.*, 2006a) suggests that there may be some form of cooperative binding. Moreover the cryoEM structure (Costa *et al.*, 2008) shows some unexplained electron density adjacent to the C-terminal domain that may represent a fragment of dsDNA binding to the winged helix domain (Figure 4F). It is interesting that the deletion of the C-terminal domain showed an increase in DNA-stimulated ATPase activity, but a decrease in helicase activity (Jenkinson and Chong, 2006) suggesting that it may act as a brake (as the N-terminal domain) onto uncontrolled ATP hydrolysis in the absence of DNA substrate.

Structural studies of eukaryotic MCM proteins

Highly defocused two-dimensional electron microscopy images available for the eukaryotic MCM complex suggest the presence of a hexamer. The *S. pombe* MCM proteins have been purified as a complex with an apparent molecular weight of 560 kDa, containing equimolar amounts of each of the six subunits, consistent with the formation of a heterohexamer and EM studies showed that the complex has a globular shape with a central cavity (Adachi *et al*., 1997). The structure of the human MCM4/6/7 complex also showed a toroidal shape with a central channel (Sato *et al*., 2000) and a similar structure has been observed for the MCM2/4/6/7 complex (Yabuta et al., 2003). Negative stained micrographs of recombinant MCM2-7 and MCM4/6/7 complexes also reveal a hollow ring, probably with hexameric symmetry (Bochman and Schwacha, 2007).

Physiological relevance of structural results

Functional role of single and double rings

Recombinant *Mth*MCM forms large multimeric complexes, with a preference for double rings as characterized by size-exclusion chromatography and EM analysis (Kelman *et al.*, 1999; Chong *et al.*, 2000; Shechter *et al.*, 2000). Consistently, the crystal structure of *Mth*MCM-N shows a double hexamer, held together by interactions involving a conserved Zn motif at the bottom of subdomain B (Fletcher *et al.*, 2003) (Figure 2A). In the absence of substrates, the double rings are not stable enough to be easily visualized by electron microscopy (Yu *et al.*, 2002; Pape *et al.*, 2003) (Figure 4). Treating the protein with either nucleotide analogs or short dsDNA substrates leads to the stabilization of double ring configurations, specifically double heptamers in the presence of

ATP analogs and double hexamers in the presence of DNA fragments (Costa *et al.*, 2006b) (Figure 4C). In contrast with *Mth*MCM, MCM proteins from other archaeal species are found to assemble into single rings in solution (Sakakibara *et al.*, 2009).

There are conflicting results as to the need of a double hexamer for the helicase activity of *Mth*MCM, as a disruption of the hexamer–hexamer interface has been reported to decrease or abolish helicase activity (Fletcher *et al.*, 2005; Jenkinson *et al.*, 2009), whereas a more recent analysis shows that the *Mth*MCM protein is a single ring at temperatures and salt conditions comparable to those found in *M. thermautotrophicus* cells, with the helicase activity decreasing in conditions that favor the formation of double rings (Shin *et al.*, 2009).

Whether the protein is active as a single or double hexamer strongly influences the helicase mechanisms that can be envisaged (Takahashi *et al.*, 2005). Whereas the canonical steric-exclusion mechanism implies a single ring moving along ssDNA, displacing the other strand, the model proposed for the SV40 LtAg helicase requires a double ring sitting at the origin, with ssDNA being threaded through lateral holes (Li *et al.*, 2003). It has to be stressed that the design of the biochemical assays for helicase activity is based on the steric exclusion model and may not reflect the exact situation in the cellular context; moreover, the activity of the protein *in vivo* may be influenced by a variety of factors (such as Cdc45 and the GINS complex) that affect its function in vivo (Gambus *et al.*, 2006; Moyer *et al.*, 2006; Aparicio *et al.*, 2009).

Although the eukaryotic MCM2-7 complex is invariably found as a single ring in solution, preliminary structural data on the yeast complex shows that the protein loads onto dsDNA as a double ring (J. Diffley, D. Remus and D. Wigley, personal communication) that resembles the double hexamer arrangement seen for the *Mth*MCM protein treated with 100 bp of dsDNA (Costa *et al.*, 2006a) (Figure 4C). This information is consistent with footprinting experiments carried out on the MCM4/6/7 complex indicating that the protein protects a DNA region corresponding to the size of a double hexamer (You and Masai, 2005) as well as with the size of the pre-replicative footprinting (Santocanale and Diffley, 1996; Labib *et al.*, 2001).

Whether the double hexamer has simply a role in the initial stage of DNA replication (helicase loading, initial melting), or is necessary for the helicase activity, remains to be established. One difficulty with an SV40-like mechanism where the double hexamer is the active helicase is that its topology is hard to reconcile with the commonly accepted model for the establishment of sister chromatid cohesion during the S phase, suggesting that the replication fork passes through a single cohesin ring (Takahashi *et al.*, 2005; Haering *et al.*, 2008).

Functional role of heptameric rings

More puzzling is the observation that archaeal MCM proteins form heptameric and double heptameric rings, raising the question of the physiological significance of these configurations. One possibility is that the heptamer is an *in vitro* artifact. This would be consistent with the fact that in eukaryotes the minimal MCM requirement is a complex made up by six different subunits (MCM2-7): although the exact stoichiometry is not known, the most likely arrangement would be a heterohexamer with one copy of each subunit, as suggested by electron microscopy analysis (Bochman and Schwacha, 2007).

However, the formation of double heptameric rings is systematically triggered in the *Mth*MCM system by the addition of ATP analogs, such as AMP-PNP and ADP·ALF_x (Costa *et al.*, 2006b) (Figure 4B). As nucleotide binding occurs in the AAA+ domain and the double ring interface involves the bottom of the Nterminal domain this result indicates long range inter- and intra-subunit communication. Heptameric rings are also stabilized by mutations at the very bottom of subdomain B; in particular the R137A mutant shows a double heptamer configuration with a remarkable swing-out of subdomain A in one of the two rings (Jenkinson *et al.*, 2009) (Figure 4E). This raises the intriguing possibility that the mutation "traps" a more labile transient intermediate state that is necessary for the loading of the complex onto DNA.

Although it is more difficult to extend this model to the eukaryotic hexameric MCM2-7, initial data on EM analysis of MCM complexes purified from eukaryotic cells show hints of heptameric symmetry. This observation, although still very preliminary, has been made independently by two different groups (J. Diffley, D. Remus and D. Wigley; S. Onesti, A. Patwardhan and J. Blow, unpublished observations) working on two diverse eukaryotic systems. Further work is required to confirm the result and establish the biochemical nature of the putative heptamer.

Mechanism of action of MCM helicases

Role in the initiation of DNA replication

The identification of an unexpected modality of contact between MCM and dsDNA, with the nucleic acid wrapping around the N-terminal half of the protein (Figure 4F, Costa *et al.*, 2008) hints at the possibility that MCM protein may be actively involved in the initiation of DNA replication (Costa and Onesti, 2008) (Figure 5A). The topological changes in the DNA that are a consequence of this interaction could assist the first steps in origin melting by changing the degree of supercoiling of the adjacent DNA regions. This is reminiscent of the mechanism proposed for the bacterial initiator DnaA (Erzberger *et al.*, 2006), involving a helical structure around which the initiator DNA can be wrapped. To reinforce the analogy, *Mth*MCM proteins have also been reported to form fibrous and helical assemblies (Chen *et al.*, 2005), which seem to be dependent on the presence of DNA (Costa *et al.*, 2008).

Role in fork elongation

Although the function of MCM proteins as replicative helicases is generally accepted, the way in which the unwinding

of the DNA double helix is achieved is still contentious, with a number of different models being proposed.

The steric-exclusion model represents the simpler mechanism (Enemark and Joshua-Tor, 2008), where a helicase ring encircles a single DNA strand and translocates along it; the other strand is displaced by being excluded from the ring. This mechanism probably applies to the hexameric helicases belonging to the SF3 and SF4 families (Singleton *et al.*, 2007), including the bacterial DnaB-like helicases, the bacteriophage T7 helicase (Singleton *et al.*, 2000) and the papillomavirus E1 helicase. The crystal structure of E1 bound to ssDNA and ADP strongly suggests a sequential mode of ATP hydrolysis linked to DNA translocation via the correlated movements of the PS1BH hairpins (Enemark and Joshua-Tor, 2006).

Takahashi *et al.* (2005) have pointed out that in eukaryotes one essential element to consider is that the helicase is loaded onto chromatin much before the helicase activity is actually needed (i.e. in the G1 rather than the S phase of the cell cycle). MCM2-7 proteins have been shown to interact with chromatin in two different ways: as a labile "associated complex" that is disrupted by high salt, and as a "loaded complex" that is salt-resistant (Donovan *et al.*, 1997; Bowers *et al.*, 2004). As presumably the loaded complex is topologically linked to DNA (Figure 5A), the lack of evidence for origin melting before the beginning of S phase suggests that the complex is loaded onto dsDNA. Although it cannot be ruled out that after initial loading onto dsDNA, MCM may undergo a rearrangement in such a way as to encircle ssDNA upon activation, it is easier to envisage alternative mechanisms that maintain the topological link between the protein ring and dsDNA.

The peculiarity of the initiation of DNA replication has therefore led to a number of different suggestions which are all linked by having a MCM complex encircling (or partially encircling) dsDNA. One model was put forward to try to explain the "MCM paradox" through a pumping mechanism, where many MCM complexes loaded onto chromatin and physically attached to a nuclear matrix could translocate dsDNA through the ring causing a torsional stress that would promote DNA unwinding (Laskey and Madine, 2003). Another scheme was derived from the proposed model for the SV40 LTag helicase. This protein is supposed to bind to the viral origin of replication as a double hexamer and melt the DNA by extruding ssDNA through lateral holes (Li *et al.*, 2003), although the similarity between SV40 LTag and E1 helicases makes it difficult to reconcile such different mechanisms. The last model proposed for MCM function advocates a hexameric helicase moving along dsDNA dragging a "ploughshare" that physically separates the two strands (Takahashi *et al.*, 2005). The LTag mechanism can be classified in this category if the two rings that compose the double hexamer come apart.

The classical helicase assay is based on the steric exclusion model, but it is not easy to design alternative assays in the absence of a way to load the protein with the correct topology. The available biochemical results are compatible with multiple models, showing that MCM complexes can displace a 5′ tail, translocate along dsDNA, or unwind a heterologous junction (Shin *et al.*, 2003; Kaplan and O'Donnell, 2004). MCM proteins can bind to both dsDNA and ssDNA (the latter with slightly higher affinity), but tend to have a higher affinity for forked and bubble substrates (Grainge *et al.*, 2003; You *et al.*, 2003; De Felice *et al.*, 2004). Moreover the protein appears to have multiple DNA binding sites, including NBH, Zn domain, subdomain A, PS1BH and h2i (reviewed in Sakakibara *et al.*, 2009). Site directed mutagenesis studies of PS1BH in *Sso*MCM (McGeoch *et al.*, 2005) and h2i in *Mth*MCM (Jenkinson and Chong, 2006) have demonstrated their crucial role for DNA binding and helicase activity, with fluorescence studies suggesting a DNA-triggered movement of the h2i. However, movement of the PS1BH and h2i elements can be envisaged in all the models proposed.

The structural results do not strongly favor a specific model either. The central channel, as visualized by the crystal structure of *Mth*MCM-N (Figure 2A) and EM studies (Figure 4) can easily accommodate dsDNA. In contrast, the crystal structure of *Sso*MCM-N shows a smaller channel. One explanation is that the Nterminal domain preferentially binds ssDNA. The other is that the crystal structure has trapped an "inactive" conformation, possibly the NBH inhibitory conformation that is observed biochemically when the ACL loop is deleted (Barry *et al.*, 2009). The EM structure of *Mth*MCM bound to dsDNA shows electron density features from the internal walls of the protein ring towards the center of the DNA channel (Figure 4C), that can be explained by the h2i insertion protruding into the channel, possibly to interact with DNA.

On the other hand the considerations summarized in Takahashi (2005) make a ploughshare model the most likely candidate to describe the activity of MCM proteins. Although the ploughshare could be an element of the MCM proteins or a separate protein (such as GINS or Cdc45), the fact that both archaeal and eukaryotic proteins display helicase activity in the absence of accessory factors argues in favor of an "internal" ploughshare (Figure 5B). Of the two loops that can be modeled to project into the channel (PS1BH and h2i), the strongest contender for the ploughshare is the helix-2 insertion. This conclusion is based on biochemical data (Chong, 2005; Jenkinson and Chong, 2006) as well as structural information, showing that the h2i loop is longer and therefore more apt to extend into the wide channel and in most MCM proteins contains a conserved aromatic residue which could interact with DNA (Costa *et al.*, 2006a; Brewster *et al.*, 2008). An alternative model is that the h2i loop works by translocating dsDNA inside the channel (with a mechanism similar to the PS1BH-mediated translocation of ssDNA in the E1 helicase) and the "ploughshare" is simply represented by the lateral holes through which the ssDNA is extruded (Figure 5B).

Even though it can be imagined that both single strands are extruded by the lateral holes, a model in which one strand exits through a lateral hole while the other emerges through

the N-terminal hexamer is more consistent with biochemical data, showing that the NBH has an important role in the regulation of helicase activity (Barry *et al.*, 2009) and that the N-terminal domain has a marked preference for ssDNA (Liu *et al.*, 2008). It is more difficult to conceive a ploughshare model where the lateral holes are not used, as complementary single strands in the narrow N-terminal channel would immediately re-anneal. Moreover the position, sequence conservation and mutagenesis of the EXT loops (Brewster *et al.*, 2008) support a role of the holes in helicase activity. From EM single particle work (Costa *et al.*, 2008), the lateral holes in MCM are often blocked by an isthmus of electron density that can be modeled by the inter-subunit interaction between the PS1BH and ACL loop (Figure 3E). This interaction is disrupted in the double hexamer structure, where presumably DNA binding has caused a re-orientation of the h2i towards the center of the channel and a disengagement of the PS1BH from the ACL, possibly providing a possible mechanism to "free" the hole for extruding the ssDNA.

To structurally validate the proposed models, high resolution crystallographic data on an active hexameric MCM helicase are crucial, in the presence of various nucleotide and DNA substrates. To avoid the possibility of crystallographic artifacts, it will be crucial to obtain a hexameric model in which all helicase subunits are crystallographically independent (i.e. a whole helicase complex in the asymmetric unit). To complicate things even further, the important role of Cdc45 and GINS in stabilizing the helicase activity of the MCM complex means that a true representation of the active helicase would require functional and structural information of the interactions within the CMG complex. As obtaining all of the necessary crystal forms will obviously be a challenge, electron microscopy is likely to still play an important role in elucidating a variety of functional states, therefore providing snapshots of the helicase reaction.

Acknowledgement

We are grateful to Peter Brick, Ivet Krastanova and Barbara Medagli for critical reading of the manuscript.

***Declaration of interest*:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Adachi Y, Usukura J and Yanagida M. 1997. A globular complex formation by Nda1 and the other five members of the MCM protein family in fission yeast. Genes Cells 2:467–479.

Aparicio T, Guillou E, Coloma J, Montoya G and Mendez J. 2009. The human GINS complex associates with Cdc45 and MCM and is essential for DNA replication. Nucleic Acids Res 37:2087–2095.

Aravind L and Koonin EV. 1999. DNA-binding proteins and evolution of transcription regulation in the archaea. Nucleic Acids Res 27:4658–4670.

Bae B, Chen YH, Costa A, Onesti S, Brunzelle JS, Lin Y, Cann IK and Nair SK. 2009. Insights into the architecture of the replicative helicase from the structure of an archaeal MCM homolog. Structure 17:211–22.

Barry ER, McGeoch AT, Kelman Z and Bell SD. 2007. Archaeal MCM has separable processivity, substrate choice and helicase domains. Nucleic Acids Res 35:988–998.

Barry ER, Lovett JE, Costa A, Lea SM and Bell SD. 2009. Intersubunit allosteric communication mediated by a conserved loop in the MCM helicase. Proc Natl Acad Sci USA 106:1051–1056.

Bell SP and Dutta A. 2002. DNA replication in eukaryotic cells. Annu Rev Biochem 71:333–374.

Blanton HL, Radford SJ, McMahan S, Kearney HM, Ibrahim JG and Sekelsky J. 2005. REC, Drosophila MCM8, drives formation of meiotic crossovers. PLoS Genet 1:e40.

Blow JJ and Dutta A. 2005. Preventing re-replication of chromosomal DNA. Nat Rev Mol Cell Biol 6:476–486.

Blow JJ and Hodgson B. 2002. Replication licensing—defining the proliferative state? Trends Cell Biol 12:72–78.

Bochman ML and Schwacha A. 2007. Differences in the single-stranded DNA binding activities of MCM2-7 and MCM467:MCM2 and MCM5 define a slow ATP-dependent step. J Biol Chem 282:33795–33804.

Bochman ML and Schwacha A. 2008. The Mcm2-7 complex has in vitro helicase activity. Mol Cell 31:287–293.

Bochman ML, Bell SP and Schwacha A. 2008. Subunit organization of Mcm2-7 and the unequal role of active sites in ATP hydrolysis and viability. Mol Cell Biol 28:5865–5873.

Bowers JL, Randell JC, Chen S and Bell SP. 2004. ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. Mol Cell 16:967–978.

Brewster AS, Wang G, Yu X, Greenleaf WB, Carazo JM, Tjajadia M, Klein MG and Chen XS. 2008. Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase. Proc Natl Acad Sci USA 105:20191–20196.

Carpentieri F, De Felice M, De Falco M, Rossi M and Pisani FM. 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 277:12118–12127.

Chen YJ, Yu X, Kasiviswanathan R, Shin JH, Kelman Z and Egelman EH. 2005. Structural polymorphism of Methanothermobacter thermautotrophicus MCM. J Mol Biol 346:389–394.

Chong JP. 2005. Learning to unwind. Nat Struct Mol Biol 12:734–736.

Chong JP, Hayashi MK, Simon MN, Xu RM and Stillman B. 2000. A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. Proc Natl Acad Sci USA 97:1530–1535.

Costa A and Onesti S. 2008. The MCM complex: (just) a replicative helicase? Biochem Soc Trans 36:136–140.

Costa A, Pape T, van Heel M, Brick P, Patwardhan A and Onesti S. 2006a. Structural basis of the *Methanothermobacter thermautotrophicus* MCM helicase activity. Nucleic Acids Res 34:5829–5838.

Costa A, Pape T, van Heel M, Brick P, Patwardhan A and Onesti S. 2006b. Structural studies of the archaeal MCM complex in different functional states. J Struct Biol 156:210–219.

Costa A, van Duinen G, Medagli B, Chong J, Sakakibara N, Kelman Z, Nair SK, Patwardhan A and Onesti S. 2008. Cryo-electron microscopy reveals a novel DNA-binding site on the MCM helicase. EMBO J 27:2250–2258.

Crevel G, Ivetic A, Ohno K, Yamaguchi M and Cotterill S. 2001. Nearest neighbour analysis of MCM protein complexes in Drosophila melanogaster. Nucleic Acids Res 29:4834–4842.

Crevel G, Hashimoto R, Vass S, Sherkow J, Yamaguchi M, Heck MM and Cotterill S. 2007. Differential requirements for MCM proteins in DNA replication in Drosophila S2 cells. PLoS ONE 2:e833.

Davey MJ, Jeruzalmi D, Kuriyan J and O'Donnell M. 2002. Motors and switches: AAA+ machines within the replisome. Nat Rev Mol Cell Biol 3:826–835.

Davey MJ, Indiani C and O'Donnell M. 2003. Reconstitution of the mcm2-7p heterohexamer, subunit arrangement, and ATP site architecture. J Biol Chem 278:4491–4499.

De Felice M, Esposito L, Pucci B, De Falco M, Rossi M and Pisani FM. 2004. A CDC6-like factor from the archaea Sulfolobus solfataricus promotes binding of the mini-chromosome maintenance complex to DNA. J Biol Chem 279:43008–43012.

Diffley JF. 2004. Regulation of early events in chromosome replication. Curr Biol 14:R778–786.

Diffley JF and Labib K. 2002. The chromosome replication cycle. J Cell Sci. 115:869–872.

Donovan S, Harwood J, Drury LS and Diffley JF. 1997. Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. Proc Natl Acad Sci USA 94:5611–5616.

Duggin IG, McCallum SA and Bell SD. 2008. Chromosome replication dynamics in the archaeon Sulfolobus acidocaldarius. Proc Natl Acad Sci USA 105:16737–16742.

Edwards MC, Tutter AV, Cvetic C, Gilbert CH, Prokhorova TA and Walter JC. 2002. MCM2-7 complexes bind chromatin in a distributed pattern

surrounding the origin recognition complex in Xenopus egg extracts. J Biol Chem 277:33049–33057.

Enemark EJ and Joshua-Tor L. 2006. Mechanism of DNA translocation in a replicative hexameric helicase. Nature 442:270–275.

Enemark EJ and Joshua-Tor L. 2008. On helicases and other motor proteins. Curr Opin Struct Biol 18:243–57.

Erzberger JP and Berger JM. 2006. Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu Rev Biophys Biomol Struct 35:93–114.

Erzberger JP, Mott ML and Berger JM. 2006. Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. Nat Struct Mol Biol 13:676–683.

Fletcher RJ, Bishop BE, Leon RP, Sclafani RA, Ogata CM and Chen XS. 2003. The structure and function of MCM from archaeal *M. thermoautotrophicum*. Nat Struct Biol 10:160–167.

Fletcher RJ, Shen J, Gomez-Llorente Y, Martin CS, Carazo JM and Chen XS. 2005. Double hexamer disruption and biochemical activities of *Methanobacterium thermoautotrophicum* MCM. J Biol Chem 280:42405–42410.

Forsburg SL. 2004. Eukaryotic MCM proteins: beyond replication initiation. Microbiol Mol Biol Rev 68:109–131, table of contents.

Gai D, Zhao R, Li D, Finkielstein CV and Chen XS. 2004. Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. Cell 119:47–60.

Gambus A, Jones RC, Sanchez-Diaz A, Kanemaki M, van Deursen F, Edmondson RD and Labib K. 2006. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nat Cell Biol 8:358–366.

Ge XQ, Jackson DA and Blow JJ. 2007. Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. Genes Dev. 21:3331–3341.

Gomez-Llorente Y, Fletcher RJ, Chen XS, Carazo JM and San Martin C. 2005. Polymorphism and double hexamer structure in the archaeal minichromosome maintenance (MCM) helicase from *Methanobacterium thermoautotrophicum*. J Biol Chem 280:40909–40915.

Gozuacik D, Chami M, Lagorce D, Faivre J, Murakami Y, Poch O, Biermann E, Knippers R, Brechot C and Paterlini-Brechot P. 2003. Identification and functional characterization of a new member of the human Mcm protein family: hMcm8. Nucleic Acids Res 31:570–579.

Grainge I, Scaife S and Wigley DB. 2003. Biochemical analysis of components of the pre-replication complex of Archaeoglobus fulgidus. Nucleic Acids Res 31:4888–4898.

Haering CH, Farcas AM, Arumugam P, Metson J and Nasmyth K. 2008. The cohesin ring concatenates sister DNA molecules. Nature 454:297–301.

Hardy CF, Dryga O, Seematter S, Pahl PM and Sclafani RA. 1997. mcm5/cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p. Proc Natl Acad Sci USA 94:3151–3155.

Hyrien O, Marheineke K and Goldar A. 2003. Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. BioEssays 25:116–125.

Ibarra A, Schwob E and Mendez J. 2008. Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. Proc Natl Acad Sci USA 105:8956–8961.

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

Iyer LM, Leipe DD, Koonin EV and Aravind L. 2004. Evolutionary history and higher order classification of AAA+ ATPases. J Struct Biol 146:11–31.

Jenkinson ER and Chong JP. 2006. Minichromosome maintenance helicase activity is controlled by N- and C-terminal motifs and requires the ATPase domain helix-2 insert. Proc Natl Acad Sci USA 103:7613–7618.

Jenkinson ER, Costa A, Leech AP, Patwardhan A, Onesti S and Chong JP. 2009. Mutations in subdomain B of the minichromosome maintenance (MCM) helicase affect DNA binding and modulate conformational transitions. J Biol Chem 284:5654–5661.

Johnson EM, Kinoshita Y and Daniel DC. 2003. A new member of the MCM protein family encoded by the human MCM8 gene, located contrapodal to GCD10 at chromosome band 20p12.3-13. Nucleic Acids Res 31:2915–2925.

Kaplan DL and O'Donnell M. 2004. Twin DNA pumps of a hexameric helicase provide power to simultaneously melt two duplexes. Mol Cell 15:453–465.

Kasiviswanathan R, Shin JH, Melamud E and Kelman Z. 2004. Biochemical characterization of the Methanothermobacter thermautotrophicus minichromosome maintenance (MCM) helicase N-terminal domains. J Biol Chem 279:28358–28366.

Kelman Z, Lee JK and Hurwitz J. 1999. The single minichromosome maintenance protein of *Methanobacterium thermoautotrophicum* DeltaH contains DNA helicase activity. Proc Natl Acad Sci USA 96:14783–14788.

Kinoshita Y, Johnson EM, Gordon RE, Negri-Bell H, Evans MT, Coolbaugh J, Rosario-Peralta Y, Samet J, Slusser E, Birkenbach MP, Daniel DC. 2008. Colocalization of MCM8 and MCM7 with proteins involved in distinct aspects of DNA replication. Microsc Res Tech 71:288–297.

Labib K, Tercero JA and Diffley JF. 2000. Uninterrupted MCM2-7 function required for DNA replication fork progression. Science 288:1643–1647.

Labib K, Kearsey SE and Diffley JF. 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 12:3658–3667.

Laskey R. 2005. The Croonian Lecture 2001 hunting the antisocial cancer cell: MCM proteins and their exploitation. Philos Trans R Soc Lond B Biol Sci 360:1119–1132.

Laskey RA and Madine MA. 2003. A rotary pumping model for helicase function of MCM proteins at a distance from replication forks. EMBO Rep 4:26–30.

Lee JK and Hurwitz J. 2001. Processive DNA helicase activity of the minichromosome maintenance proteins 4, 6, and 7 complex requires forked DNA structures. Proc Natl Acad Sci USA 98:54–59.

Lei M. 2005. The MCM complex: its role in DNA replication and implications for cancer therapy. Curr Cancer Drug Targets 5:365–380.

Lei M and Tye BK. 2001. Initiating DNA synthesis: from recruiting to activating the MCM complex. J Cell Sci 114:1447–1454.

Li D, Zhao R, Lilyestrom W, Gai D, Zhang R, DeCaprio JA, Fanning E, Jochimiak A, Szakonyi G and Chen XS. 2003. Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. Nature 423:512–518.

Liu J, Smith CL, DeRyckere D, DeAngelis K, Martin GS and Berger JM. 2000. Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. Mol Cell 6:637–648.

Liu W, Pucci B, Rossi M, Pisani FM and Ladenstein R. 2008. Structural analysis of the Sulfolobus solfataricus MCM protein N-terminal domain. Nucleic Acids Res 36:3235–3243.

Liu Y, Richards TA and Aves SJ. 2009. Ancient diversification of eukaryotic MCM DNA replication proteins. BMC Evol Biol 9:60.

Lutzmann M and Mechali M. 2008. MCM9 binds Cdt1 and is required for the assembly of prereplication complexes. Mol Cell 31:190–200.

Lutzmann M, Maiorano D and Mechali M. 2005. Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2-8 protein family. Gene 362:51–56.

Maiorano D, Cuvier O, Danis E and Mechali M. 2005. MCM8 is an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. Cell 120:315–328.

Maiorano D, Lutzmann M and Mechali M. 2006. MCM proteins and DNA replication. Curr Opin Cell Biol 18:130–136.

Matias PM, Gorynia S, Donner P and Carrondo MA. 2006. Crystal structure of the human AAA+ protein RuvBL1. J Biol Chem 281:38918–38929.

Matsubayashi H and Yamamoto MT. 2003. REC, a new member of the MCM-related protein family, is required for meiotic recombination in Drosophila. Genes Genet Syst 78:363–371.

McGeoch AT, Trakselis MA, Laskey RA and Bell SD. 2005. Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. Nat Struct Mol Biol 12:756–762.

Moreau MJ, McGeoch AT, Lowe AR, Itzhaki LS and Bell SD. 2007. ATPase site architecture and helicase mechanism of an archaeal MCM. Mol Cell 28:304–314.

Moyer SE, Lewis PW and Botchan MR. 2006. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. Proc Natl Acad Sci USA 103:10236–10241.

Pape T, Meka H, Chen S, Vicentini G, van Heel M and Onesti S. 2003. Hexameric ring structure of the full-length archaeal MCM protein complex. EMBO Rep 4:1079–1083.

Pucci B, De Felice M, Rossi M, Onesti S and Pisani FM. 2004. Amino acids of the Sulfolobus solfataricus mini-chromosome maintenance-like DNA helicase involved in DNA binding/remodeling. J Biol Chem 279:49222–49228.

Pucci B, De Felice M, Rocco M, Esposito F, De Falco M, Esposito L, Rossi M and Pisani FM. 2007. Modular organization of the *Sulfolobus solfataricus* mini-chromosome maintenance protein. J Biol Chem 282:12574–12582.

Rappas M, Schumacher J, Beuron F, Niwa H, Bordes P, Wigneshweraraj S, Keetch CA, Robinson CV, Buck M and Zhang X. 2005. Structural

insights into the activity of enhancer-binding proteins. Science 307:1972–1975.

Rappas M, Schumacher J, Niwa H, Buck M and Zhang X. 2006. Structural basis of the nucleotide driven conformational changes in the AAA+ domain of transcription activator PspF. J Mol Biol 357:481–492.

Sakakibara N, Kasiviswanathan R, Melamud E, Han M, Schwarz FP and Kelman Z. 2008. Coupling of DNA binding and helicase activity is mediated by a conserved loop in the MCM protein. Nucleic Acids Res 36:1309–1320.

Sakakibara N, Kelman LM and Kelman Z. 2009. Unwinding the structure and function of the archaeal MCM helicase. Mol Microbiol 72:286–296.

Sakwe AM, Nguyen T, Athanasopoulos V, Shire K and Frappier L. 2007. Identification and characterization of a novel component of the human minichromosome maintenance complex. Mol Cell Biol 27:3044–3055.

Santocanale C and Diffley JF. 1996. ORC- and Cdc6-dependent complexes at active and inactive chromosomal relication origins in *Saccharomyces cerevisiae*. EMBO J 15:6671–6679.

Sato M, Gotow T, You Z, Komamura-Kohno Y, Uchiyama Y, Yabuta N, Nojima H and Ishimi Y. 2000. Electron microscopic observation and single-stranded DNA binding activity of the Mcm4,6,7 complex. J Mol Biol 300:421–431.

Schwacha A and Bell SP. 2001. Interactions between two catalytically distinct MCM subgroups are essential for coordinated ATP hydrolysis and DNA replication. Mol Cell 8:1093–1104.

Sclafani RA and Holzen TM. 2007. Cell Cycle Regulation of DNA Replication. Annu Rev Genet.

Shechter DF, Ying CY and Gautier J. 2000. The intrinsic DNA helicase activity of Methanobacterium thermoautotrophicum delta H minichromosome maintenance protein. J Biol Chem 275:15049–15059.

Shin JH, Jiang Y, Grabowski B, Hurwitz J and Kelman Z. 2003. Substrate requirements for duplex DNA translocation by the eukaryal and archaeal minichromosome maintenance helicases. J Biol Chem 278:49053–49062.

Shin JH, Heo GY and Kelman Z. 2009. The Methanothermobacter thermautotrophicus MCM helicase is active as a hexameric ring. J Biol Chem 284:540–546.

Singleton MR, Sawaya MR, Ellenberger T and Wigley DB. 2000. Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. Cell 101:589–600.

Singleton MR, Dillingham MS and Wigley DB. 2007. Structure and Mechanism of Helicases and Nucleic Acid Translocases. Annu Rev Biochem 76:23–50.

Speck C, Chen Z, Li H and Stillman B. 2005. ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. Nat Struct Mol Biol 12:965–971.

Stillman B. 2005. Origin recognition and the chromosome cycle. FEBS Lett 579:877–884.

Takahashi TS, Wigley DB and Walter JC. 2005. Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase. Trends Biochem Sci 30:437–444.

Volkening M and Hoffmann I. 2005. Involvement of human MCM8 in prereplication complex assembly by recruiting hcdc6 to chromatin. Mol Cell Biol 25:1560–1568.

Wigley DB. 2009. ORC proteins: marking the start. Curr Opin Struct Biol 19:72–78.

Woodward AM, Gohler T, Luciani MG, Oehlmann M, Ge X, Gartner A, Jackson DA and Blow JJ. 2006. Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. J Cell Biol 173:673–683.

Yabuta N, Kajimura N, Mayanagi K, Sato M, Gotow T, Uchiyama Y, Ishimi Y and Nojima H. 2003. Mammalian Mcm2/4/6/7 complex forms a toroidal structure. Genes Cells 8:413–421.

Yoshida K. 2005. Identification of a novel cell-cycle-induced MCM family protein MCM9. Biochem. Biophys Res Commun 331:669–674.

You Z and Masai H. 2005. DNA binding and helicase actions of mouse MCM4/6/7 helicase. Nucleic Acids Res 33:3033–3047.

You Z, Komamura Y and Ishimi Y. 1999. Biochemical analysis of the intrinsic Mcm4-Mcm6-mcm7 DNA helicase activity. Mol Cell Biol 19:8003–8015.

You Z, Ishimi Y, Mizuno T, Sugasawa K, Hanaoka F and Masai H. 2003. Thymine-rich single-stranded DNA activates Mcm4/6/7 helicase on Y-fork and bubble-like substrates. EMBO J 22:6148–6160.

Yu X, VanLoock MS, Poplawski A, Kelman Z, Xiang T, Tye BK and Egelman EH. 2002. The Methanobacterium thermoautotrophicum MCM protein can form heptameric rings. EMBO Rep 3:792–797.

Zhang X and Wigley DB. 2008. The 'glutamate switch' provides a link between ATPase activity and ligand binding in AAA+ proteins. Nat Struct Mol Biol 15:1223–1227.

Editor: Michael M. Cox