Published online 27 May 2009

Nucleic Acids Research, 2009, Vol. 37, No. 13 ***4441–4452***
doi:10.1093/nar/gkp376

A biochemically active MCM-like helicase in *Bacillus cereus*

Martin Samuels[1], Gaurav Gulati[1], Jae-Ho Shin[2], Rejoice Opara[1], Elizabeth McSweeney[1], Matt Sekedat[3], Stephen Long[4], Zvi Kelman[2] and David Jeruzalmi[1,*]

[1]Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, [2]Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, MD 20850, [3]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, 1230 York Avenue, NY 10065 and [4]Analytical Chemistry Division, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899-8391, USA

Received January 23, 2009; Revised April 13, 2009; Accepted April 26, 2009

ABSTRACT

The mini-chromosome maintenance (MCM) proteins serve as the replicative helicases in archaea and eukaryotes. Interestingly, an MCM homolog was identified, by BLAST analysis, within a phage integrated in the bacterium *Bacillus cereus* (*Bc*). *Bc*MCM is only related to the AAA region of MCM-helicases; the typical amino-terminus is missing and is replaced by a segment with weak homology to primases. We show that *Bc*MCM displays 3′→5′ helicase and ssDNA-stimulated ATPase activity, properties that arise from its conserved AAA domain. Isolated *Bc*MCM is a monomer in solution but likely forms the functional oligomer *in vivo*. We found that the *Bc*MCM amino-terminus can bind ssDNA and harbors a zinc atom, both hallmarks of the typical MCM amino-terminus. No *Bc*MCM-catalyzed primase activity could be detected. We propose that the divergent amino-terminus of *Bc*MCM is a paralog of the corresponding region of MCM-helicases. A divergent amino terminus makes *Bc*MCM a useful model for typical MCM-helicases since it accomplishes the same function using an apparently unrelated structure.

INTRODUCTION

The DNA-replication systems of eukaryotes, archaea and bacteria employ hexameric DNA helicases to unwind DNA ahead of advancing replication forks (1,2). These enzymes unwind duplex DNA by assembling on and translocating along single-strand (ss) DNA at the fork, while displacing the complementary strand. In eukaryotes and archaea, the mini-chromosome maintenance (MCM) ensemble serves as the replicative helicase. This view is supported by several observations including that: an MCM subcomplex exhibits ATP-dependent helicase activity *in vitro*; the complete MCM complex, along with several accessory proteins, also demonstrates ATP-dependent helicase activity; the MCM assembly travels with the replication fork; and that ATPase-inactivating mutations in MCM proteins prevent DNA unwinding (3–6). Moreover, these MCM genes have proven to be essential to viability in yeast (1).

MCM helicases are DNA-stimulated ATPases that unwind DNA from the 3′ to 5′ direction (1,2). In eukaryotes, the MCM ensemble is a large (Mw $\approx$ 600 kDa) hetero-oligomer composed of six homologous subunits (MCM2-7). Most archaeal genomes encode a single MCM protein, which assembles into an active homo-oligomer. The eukaryotic MCM2-7 ensemble displays ATPase activity, though helicase activity has only been observed under certain solution conditions *in vitro* (7–9). On the other hand, the archaeal MCM helicase assembly shows potent ATPase and helicase activity *in vitro*, allowing it to serve as a simpler and more tractable model to study the hetero-oligomeric eukaryotic MCM complex.

Sequence, biochemical, and structural analyses suggest that MCM helicases share a common core structural domain that extends for ~500–600 amino acids (1,2). The MCM domain can be sub-divided into two smaller structures: a conserved amino-terminal domain (MCM-N); and a more-highly conserved AAA module, with elements unique to MCM helicases (MCM-AAA). MCM-N binds DNA, is involved in oligomerization, and harbors a zinc-coordination module. Structural and biochemical analysis of the archaeal MCM-N region reveals a homohexameric ring that binds DNA within its central cavity (10,11). β-hairpins, tipped with positively charged

*To whom correspondence should be addressed. Tel: +1 617 496 9734; Fax: +1 617 496 9684; Email: dj@mcb.harvard.edu
Present address:
Jae-Ho Shin, School of Applied Biosciences, Kyungpook National University, Daegu 702-701, Republic of Korea.

residues, project from the walls into the MCM-N cavity and are shown to contact DNA (10). The zinc-coordination module, which is conserved in virtually every MCM helicase subunit [among the eukaryotic MCM subunits (1) and the archaeal orthologs], is essential for activity and appears to serve a structural role (10–13).

The MCM-AAA domain is located at the carboxy terminus and contains AAA-style (ATPases Associated with various cellular Activities) nucleotide-binding elements (e.g. P-loop/Walker A, Walker B, Sensor I, Arginine Finger and Sensor II) (Figure 1) (14). Electron microscopy of full-length MCM proteins shows that the MCM-AAA segment also forms a ring structure, likely by subunits interacting through composite nucleotide-binding sites at subunit interfaces (15). The MCM-AAA contains a second β-hairpin, which is inserted before the AAA Sensor 1 motif (referred to as the 'pre-sensor 1 β-hairpin') and is implicated in both contacting DNA and in nucleotide-dependent conformational changes (16). Additional elements, such as the helix-2-insert, help distinguish the MCM-AAA module from other proteins in the AAA ATPase superfamily and likely participate in the unwinding reaction (17). Additionally, archaeal MCMs have a conserved putative winged helix-turn-helix (wHTH) motif at the extreme C-terminus (Figure 1A) (2).

The complete MCM-helicase domain is, therefore, composed of two sub-structures (MCM-N and MCM-AAA). Oligomerization of MCM proteins generates an ensemble with a long tunnel through which DNA passes, guided, in part, by these two β-hairpins. Two recently described members of the MCM-helicase family, MCM8 and MCM9, retain the features seen in MCM2-7 and archaeal orthologs (18,19).

The structural basis for the unwinding activity by helicases in general, and especially MCM helicases, remains incompletely understood. Structural analysis by electron microscopy (EM) has shown eukaryotic and archaeal MCMs as ring-shaped hexamers, heptamers, or double-hexamers (15,20–22). The observed double-hexamers appear to be two single hexamers that are assembled head-to-head via contacts between MCM-N segments (10). These studies have also offered a glimpse of the richness of the structural changes of the MCM complex as it associates with various cofactors (15,22). Such EM studies, in combination with X-ray structures, enable modeling of the hexameric entity. These analyses now set the stage for a more complete understanding of the structural basis for DNA unwinding (23–27).

Below, we describe a phage MCM-homolog found within the genome of the bacterium *Bacillus cereus* (*Bc*), initially identified by computational methods (28,29), as an alternate model for studying DNA unwinding by MCM helicases. We show that this homolog (*Bc*MCM) possesses ATP-dependent helicase activity *in vitro*. Unwinding is exclusively in the $3'\rightarrow5'$ direction, similar to eukaryotic and archaeal MCM homologs. *Bc*MCM harbors an ATPase activity that is localized to its MCM-AAA domain and is stimulated by ssDNA. We show that the full-length protein is a monomer in solution, but higher-order oligomers are evident when the *Bc*MCM amino-terminal and MCM-AAA sub-domains are isolated from each other. Both of these isolated sub-domains can bind ssDNA. *Bc*MCM also contains an amino-terminal zinc-coordination site.

MATERIALS AND METHODS

Protein expression constructs

The gene for *B. cereus* MCM (*Bc*MCM) was amplified from genomic DNA (ATCC 14579) and inserted into the SphI and SacII sites of the plasmid pQE2 (Qiagen). DNA sequencing of the resulting plasmid clone revealed two sequence changes relative to the published sequence (nucleotides C663A and T2775A, resulting in the amino-acid substitutions S221R and D925E, respectively). These two differences may have arisen from sequencing different isolates or from sequencing errors. Fortuitously, the two changes seem relatively conservative. In addition to the full-length gene, we prepared several shorter constructs to isolate the predicted helicase and the primase activities (Figure 1A and Figure 2). Several such constructs were prepared owing to our inability to reliably establish domain boundaries. The C200, C403, C501 and N399 truncations, in addition to the full-length protein, have proven suitable for study (Figure 1B). Each construct contains an amino-terminal hexahistidine tag. We also prepared two full-length mutant proteins: one with a disrupted ATP-binding site (Walker A: K653A); and a second with a disrupted zinc-binding site (C261A) to serve as controls in biochemical studies. PCR primers for the constructs prepared for this study appear in Supplementary Table 1.

Protein expression and purification

Full-length and truncated constructs of *Bc*MCM were expressed and purified using an identical sequence of steps. Proteins were expressed in *E coli* M15 cells in the following media: 90 mM potassium phosphate/pH 7.4; 2.4% yeast extract; 1.2% tryptone; 0.8% glycerol, supplemented with 50 μg/ml kanamycin and 100 μg/ml ampicillin. The culture was grown at 37°C until an OD_{600} of ~4.5 was reached. Protein expression was then induced at 22°C for full-length (or at 15°C for the shorter truncations) by addition of 1 mM isopropyl thiogalactopyranoside (IPTG). Sixteen hours later, the cells were harvested by centrifugation (4000 RPM for 30 min), resuspended in lysis buffer (50 mM sodium phosphate/pH 8.0, 1 M NaCl, 10 mM imidazole and 10% sucrose) and frozen at −80°C.

Freshly thawed cells were lysed by sonication and the soluble fraction was applied to a NiNTA column (Qiagen) pre-equilibrated in Buffer A (20 mM sodium phosphate/pH 8.0, 200 mM NaCl, 60 mM imidazole and 10% glycerol). The protein was eluted from the column using a linear gradient from Buffer A to Buffer B (20 mM sodium phosphate/pH 8.0, 200 mM NaCl, 500 mM imidazole and 10% glycerol). The fractions were analyzed using 12% SDS–PAGE, and those containing the protein were collected, pooled and then dialyzed against Buffer C (20 mM Tris–HCl/pH 7.5, 100 mM NaCl, 5 mM β-mercaptoethanol and 10% glycerol). The pool of *Bc*MCM was then applied to an SP-sepharose column pre-equilibrated in

Buffer C, and a linear gradient was run on the column from Buffer C to Buffer D (20 mM Tris–HCl/pH 7.5, 1 M NaCl, 5 mM β-mercaptoethanol and 10% glycerol). The eluted fractions containing the protein were then collected and pooled and dialyzed in Buffer C. The protein was next applied to a Q-sepharose column pre-equilibrated in Buffer C and eluted using a linear gradient from Buffer C to Buffer D. *Bc*MCM-containing fractions were pooled and concentrated by ultrafiltration (VivaSpin VivaScience; 10 kDa MWCO). The identity of all purified proteins was verified by mass-spectrometry.

Preparation of the Helicase DNA substrate

A single-strand of DNA was labeled with [γ^{32}P] ATP (Amersham, 10 mCi/ml). The labeling reaction consists of incubating 10 pmol of oligonucleotide (DF50; Supplementary Table 2) with 2.5 µl 10× T4 polynucleotide kinase reaction buffer (New England Biolabs (NEB)), 1 µl T4 polynucleotide kinase (NEB), and 5 µl [γ^{32}P]ATP in a total volume of 25 µl at 37°C for 75 min. The reaction was stopped by adding 1 µl of 0.5 M EDTA–NaOH/pH 8, and then purified by centrifuging the oligonucleotide through G50 desalting beads twice at 0.7RCF for 1 min each time.

A second, unlabeled oligonucleotide (DF61; Supplementary Table 2) was then hybridized to the first in a 3:1 ratio: 30 pmol of DF61 was added to the labeled oligo in 60 mM HEPES–NaOH/pH 7.5 and 50 mM NaCl before being brought to a boil for 3 min and then allowed to cool to room temperature. The resulting helicase substrate was then gel purified via a mini 8% polyacrylamide, 1× TBE native gel that was run at 80 V for 55 min. The radioactive band was then excised from the gel and extracted in 0.5 M ammonium acetate, 10 mM magnesium acetate and 1 mM EDTA; shaking at 400 RPM at room temperature overnight. After elution from the gel slice, the DNA was ethanol precipitated and finally resuspended in 20 mM Tris–HCl/pH 8.5. The final concentration of the substrate was determined by comparing the specific activity of this purified substrate to an earlier sample taken after the labeling reaction. In assays using substrates with biotin–streptavidin beads, 10 fmol of biotinylated DNA substrate was incubated with 1 pmol of streptavidin for 10 min at 30°C (30).

Helicase assay

Purified proteins were assayed for *in vitro* helicase activity by incubating them with the helicase DNA substrate. 10 fmol of the DNA substrate were incubated with enzyme in 20 mM Tris–HCl/pH 8.5, 10 mM $MgCl_2$, 2 mM DTT and 0.1 mg/ml BSA for 60 min at 37°C in a 15 µl reaction. The reaction was terminated by addition of 5 µl of 4× stop buffer (50% glycerol, 0.1 M EDTA, 0.1% Bromophenol Blue, 0.1% Xylene Cyanol). The results were then visualized on 8% polyacrylamide, 0.5× TBE native mini-gels that were run for 50 min at 90 V, and then exposed to a phosphor screen. The phosphor screen was read with a PhosphorImager (Molecular Dynamics), and visualized using the ImageQuant® software.

ATPase assay

The rate of ATP hydrolysis was measured using a steady-state, NADH-coupled spectrophotometric assay. The purified proteins were incubated in a reaction buffer consisting of 50 mM HEPES–KOH/pH 7.5, 150 mM potassium acetate and 8 mM magnesium acetate. Each assay was repeated eight times in a column on a UV-transparent, 96-well plate (Costar 3635). Rates were measured using a Molecular Devices, SpectraMax M5 plate reader and quantified using SoftMax Pro v5. Experiments that included ssDNA employed a 61-mer sequence (Supplementary Table 2). The ATP turnover rates were calculated from the equation: ATPase rate = $-(dA_{340}/dt) \times (1/K_{path}) \times (1/\text{mol of ATPase})$; where 'ATPase rate' is measured in terms of ATP turned over per minute; 'dA_{340}/dt' is expressed in OD/min; and 'K_{path}' is the molar absorption coefficient for NADH at a given optical pathlength (31).

Tryptophan-quenching assay

The ability to bind single-stranded, synthetic DNA oligonucleotides (Supplementary Table 2) was tested by measuring the quenching of *Bc*MCM's intrinsic tryptophan fluorescence. Binding constants were determined by keeping the protein concentration fixed (1 µM) and titrating DNA from 0 to 50 µM. The reactions were incubated in 20 mM Tris–HCl/pH 7.5, 125 mM NaCl and 10 mM $MgCl_2$. The reactions were conducted in 96-well, UV-transparent trays (CoStar 3635), repeated four times per tray, and the fluorescent quenching was read using the SoftMax Pro v5 software. Dissociation constants (K_d) were calculated using the equation: $dF/dF_{max} = [DNA]/([DNA] + K_d)$; where '$dF/dF_{max}$' represents the change in the protein's fluorescence relative to its maximal change in fluorescence. Calculations were performed using the programs Microsoft Excel and Kaleidagraph.

Primase assay

Full length *Bc*MCM was assayed for *in vitro* primase activity by providing a ssDNA substrate and observing whether the enzyme is competent to make an RNA primer that enables subsequent DNA synthesis. The enzymes were incubated with 7.5 nM of single-stranded ϕX174 plasmid (NEB); 100 µM each of ATP, CTP, GTP and UTP; 50 µM each of dCTP, dGTP and dTTP; 1.32 µM [α^{32}P] dATP (Amersham, 20 mCi/ml); and 1 µl Klenow in a buffer of 50 mM Tris–HCl/pH 7.5, 5 mM DTT, 10% glycerol and 0.1 mg/ml BSA. Reactions were carried out at 30°C for 10 min and terminated by spotting 1 µl of the 25 µl reaction onto DE81 filter paper (Whatman). The filters were then washed four times in a buffer of 0.3 M ammonium formate and 10 mM sodium pyrophosphate and left to dry overnight before counting their scintillation (32).

Multi-angle laser light scattering

Purified proteins were applied to a Shodex KW-804 column equilibrated in 25 mM Tris–HCl/pH 7.5 and 350 mM ammonium sulfate. Light scattering and refractive

index signals were measured using a Wyatt Optilab and Dawn EOS system. Scattering curves were processed using the provided Astra software package. The protein concentration for each experiment was 5 mg/ml.

Zinc analysis by inductively coupled plasma mass spectrometry

Purified protein samples were prepared for zinc analysis by dialysis against 20 mM Tris–HCl/pH 7.5, 400 mM sodium chloride, 5% glycerol and 0.5 mM EDTA. Samples were flash frozen and submitted to the National Institute of Standards and Technology (NIST), Analytical Chemistry Division. Some trace solids, which were apparent upon thawing, were re-solubilized by treatment with high-purity nitric acid. The zinc content was measured using inductively coupled plasma–mass spectrometry (ICP–MS) (12). The zinc content measurements were calibrated against UvrA, whose crystal structure reveals three-bound zinc atoms per mole of protein (33).

RESULTS

*Bc*MCM is homologous to the MCM-AAA subdomain

Computational analysis of the *B. cereus* genome revealed a prophage harboring a gene with a segment (spanning *Bc*MCM residues 500–900) that is homologous to the MCM-AAA region of the complete MCM-helicase domain (28) (Figures 1 and 2A). The MCM-AAA portion

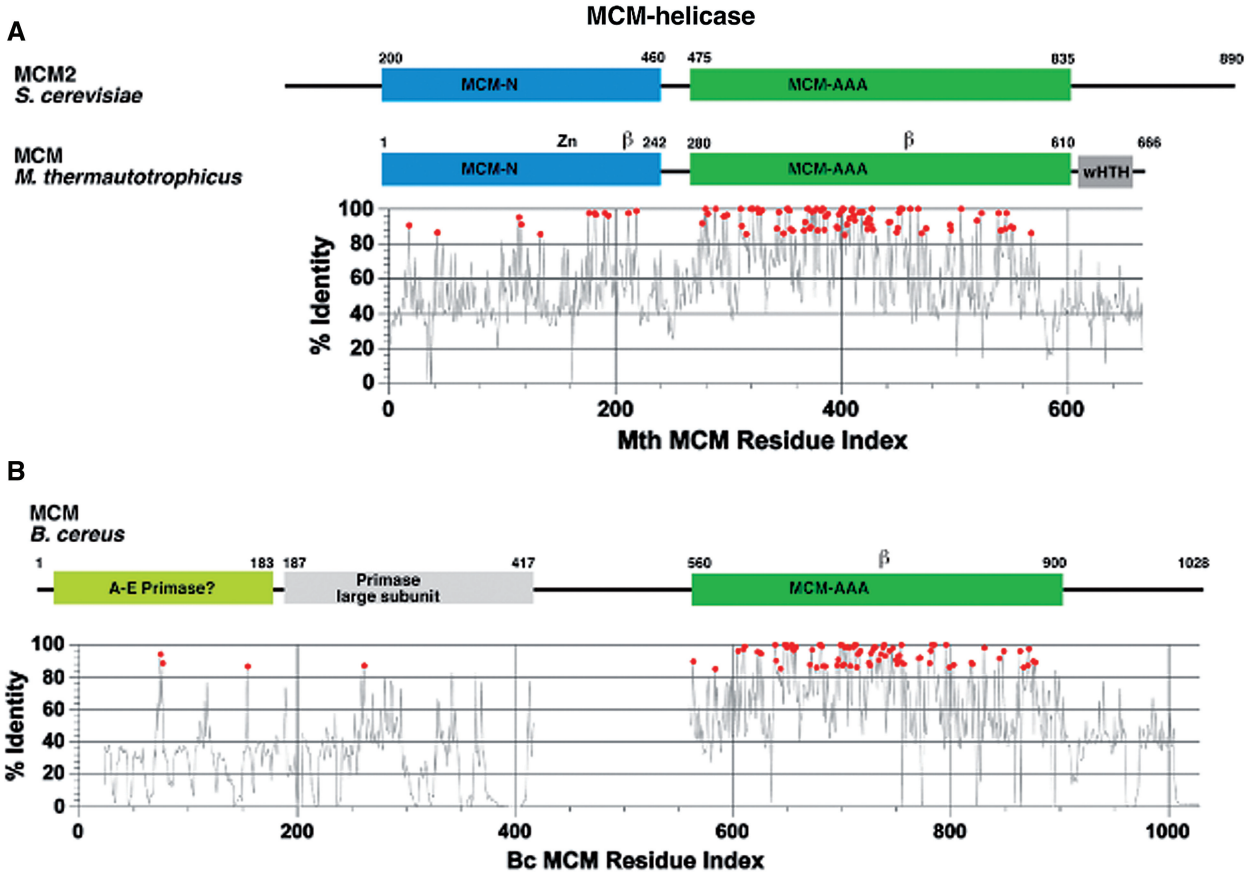

Figure 1. The *Bacillus cereus* MCM-helicase is only homologous to the MCM-AAA domain. (**A**) The MCM-helicase structure can be sub-divided into a conserved amino-terminal domain (MCM-N) and a highly conserved carboxy-terminal domain MCM-AAA. Archaeal orthologues contain a winged helix sub-domain that does not appear in its eukaryotic counterparts. DNA-binding β-hairpins and the zinc coordination site are indicated. Conservation of MCM-helicase proteins is depicted graphically by plotting the percent identity of a large group of homologues against the primary sequence of the MCM-helicase from *M. Thermautotrophicus*. Positions with a percent identity of 85% or higher are shown as magenta dots. Amino-acid sequences from this group, which includes orthologues from *S. cerevisiae* (MCM2-7), *S. pombe* (2–7), *H. sapiens* (2–9), *B. taurus* (2–7), *D. melanogaster* (2–7), *D. rerio* (2–7), *E. cuniculi* (2–7), *G. gallus* (2–6, 8), *M, musculus* (2–9), *P. troglodytes* (2–7), *R. norvegicus* (2–7, 8), *X. laevis* (9), *S. solfataricus, S. acidocaldarius*, *T. acidophilum*, *M. acetivorans*, *P. aerophilum*, *T. kodakarensis*, *M. mazei*, *A. pernix*, *A. fulgidus*, *M. Thermautotrophicus* and *Sulfolobus neozealandicus* (pTAU4 plasmid) were aligned with MUSCLE (47) using its default parameters. Percent sequence identity was calculated by averaging similarity scores at each position for all possible pairs of sequences [sum of pairs measure in (48), DJ, unpublished software]. Equivalence of non-identical residues in this calculation was established through a normalized BLOSUM62 amino-acid substitution matrix (49). (**B**) Sequence analysis of *Bc*MCM reveals regions related to the archaeo-eukaryotic family of primases (AEP, positions 73–182), the large subunit of eukaryotic primase (positions 187–417) and archaeal/eukaryotic MCM subunits (positions 560–900). Percent sequence identity, which represents three separate sequence analyses, are projected on the primary sequence of *Bc*MCM. Regions of the plot that correspond to archaeo-eukaryotic primases and the primase large subunits was calculated from Figure 2 and Figure 8 of (29). The portion of the plot that encompasses MCM-AAA was calculated using all the sequences from part A and the corresponding region of *Bc*MCM. Sequence identity was calculated as described in part A. Positions with a percent identity of 85% or higher are shown as magenta dots.

of eukaryotic and archaeal MCM proteins is highly conserved and contains the ATP binding and hydrolysis functions along with a number of structural elements unique to MCM-helicases. By contrast, residues 1–499 of *Bc*MCM align poorly with the MCM-N segment of the MCM-helicase fold, as confirmed in an extensive analysis of MCM helicase sequences (Figure 1B). Others have noted that the amino-terminus of *Bc*MCM shows weak homology to the archaeo-eukaryotic family of primases at residues 1–183 and to the large subunit of the eukaryotic primase at residues 183–417 (29) (Figure 2A). Iyer *et al.* (29) predicts that there is a zinc-coordination site within this latter segment (*Bc*MCM residues C261, C342, C364 and C369), serving as a key aspect of the region's homology to the primase large subunit. It is interesting to note that this site is located relative to the MCM-AAA region in a similar manner to the zinc site seen in archaeal and eukaryotic MCM helicases.

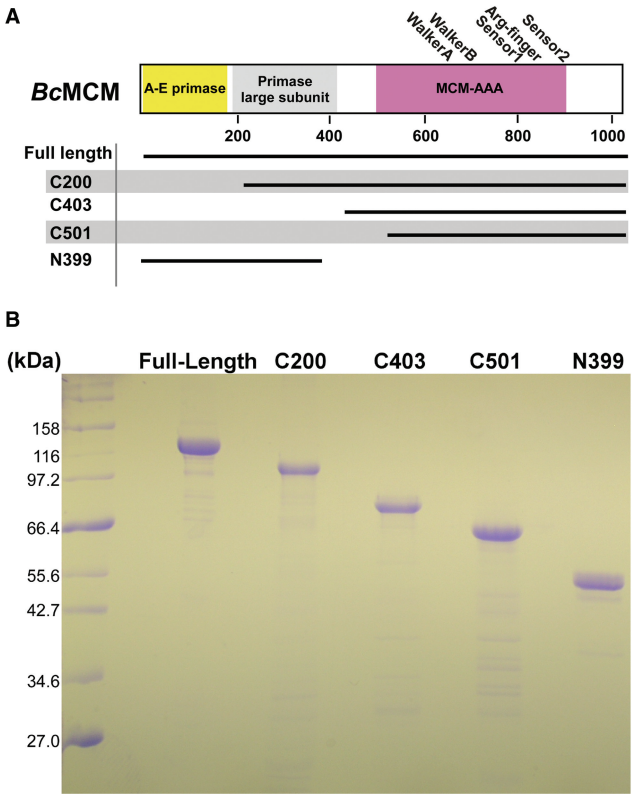

Figure 2. Constructions of *Bc*MCM-helicase used in this study. Three basic types of constructions of *Bc*MCM were prepared. These are (i) full-length; (ii) amino-terminal constructs that include the putative primase domains; and (iii) carboxy-terminal constructs that include the MCM-AAA domain. The isolated MCM and primase domain constructs were devised so that the two activities could be studied separately. Several versions were prepared owing to our inability to reliably establish domain boundaries. (**B**) Coomassie blue stained SDS–Page analysis of typical preparations of full-length and various shorter constructions of *Bc*MCM.

*Bc*MCM displays ATP-dependent DNA unwinding activity

*Bc*MCM was identified within a prophage in the genome of *B. cereus* using computational methods (28). To characterize its biochemical activity, we have used helicase assays to show that *Bc*MCM harbors ATP-dependent DNA unwinding activity. In Figure 3A, lane 4, we show that *Bc*MCM is an ATP-dependent DNA helicase, since it can convert the duplex substrate into ssDNA product. On the other hand, no unwinding was observed upon mutation of the ATP-binding site (Walker A: K635A), confirming that activity is due to *Bc*MCM (Figure 3B). ATP dependence is established in lanes 3 and 5–11, where no activity is seen with any other nucleotide (compare lane 4 to lanes 5–11). This experiment also rules out contamination by a bacterial nuclease as an explanation for the appearance of ssDNA in lane 4, since the substrate remains intact in all other lanes. Taken together, our data implies: (i) that *Bc*MCM is a helicase (when comparing the wild-type activity to that of the Walker A mutant); and (ii) that the helicase product band is not a result of nucleolytic activity given the absence of such a band when ATP is withheld from the reaction and when the Walker A mutant is employed.

*Bc*MCM unwinds DNA in the 3′→5′ direction

Many helicases unwind DNA in a specific direction (either 3′→5′ or 5′→3′). This polarity depends upon which DNA strand the helicase travels along as it displaces the

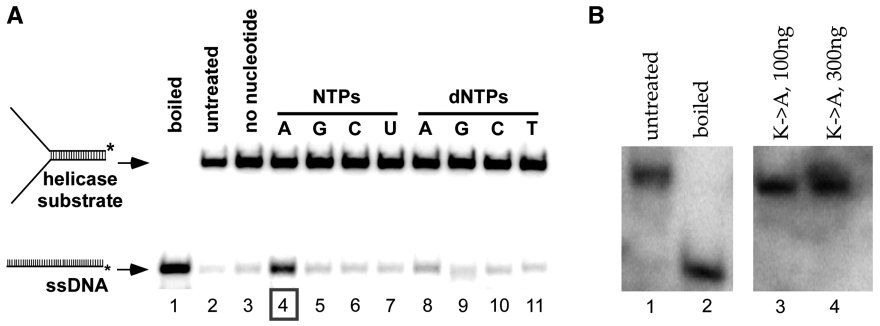

Figure 3. *Bc*MCM displays ATP dependent helicase activity. (**A**) Reaction mixture contained 2.7 pmol wild-type *Bc*MCM (as monomer) except for lanes 1 and 2. Lane 1, boiled substrate; lane 2, substrate only; lane 3, without NTPs or dNTPs; lane 4, with 0.67 mM ATP; lane 5, with 0.67 mM GTP; lane 6, with 0.67 mM CTP; lane 7, with 0.67 mM UTP; lane 8, with 0.67 mM dATP; lane 9, with 0.67 mM dGTP; lane 10, with 0.67 mM dCTP; lane 11, with 0.67 mM dTTP. Asterisk represents ^{32}P-label. (**B**) Reaction mixture contained 0.9 pmol and 2.7 pmol Walker A (K635A) *Bc*MCM in lanes 3 and 4, respectively, in the presence of 0.67 mM ATP. Lane 1, boiled substrate; lane 2, substrate only.

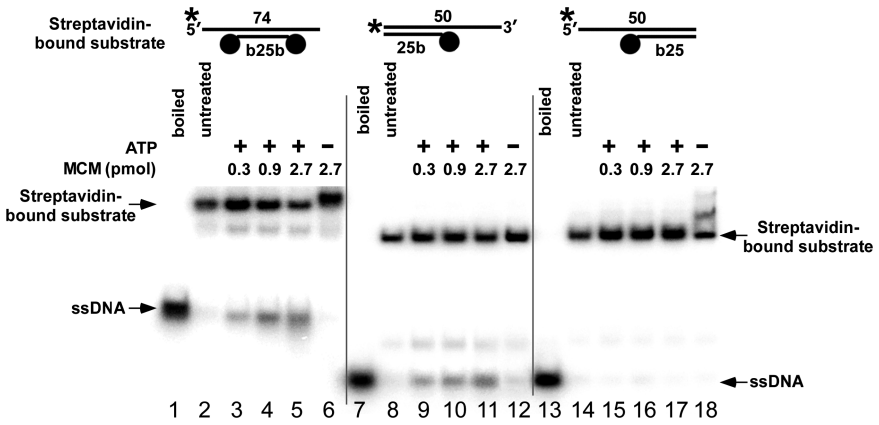

Figure 4. *Bc*MCM unwinds DNA with 3′→5′ polarity. Helicase assays were performed using the substrates shown in the top of figure. In each substrate the longer oligonucleotide is ^{32}P-labeled (asterisk) and streptavidin is marked with a closed circle. Lanes 1, 7 and 13: boiled substrate; lanes 2, 8 and 14: substrate only; lanes 3, 9 and 15: 0.3 pmol of *Bc*MCM as monomer; lanes 4, 10 and 16: 0.9 pmol of *Bc*MCM; lanes 5, 11 and 17: 2.7 pmol of *Bc*MCM; lanes 6, 12 and 18: 2.7 pmol of *Bc*MCM but without addition of ATP.

complementary strand. At the replication fork, a helicase that runs in the 3′→5′ will be traveling along the leading strand, whereas a 5′→3′ helicase will travel along the lagging strand. Experimentally, helicase polarity can be determined by observing activity on a series of substrates with a single ssDNA tail (either 3′ or 5′). In these experiments, the addition of a biotin-streptavidin group serves two roles: it blocks assembly and translocation on one strand as it also increases unwinding efficiency (30).

Three different substrates were used to determine the unwinding polarity of *Bc*MCM: (i) a molecule with an available 3′ tail; (ii) one with an available 5′ tail; and (iii) one with both 5′ and 3′ tails. Biotin-streptavidin groups (represented by black dots in Figure 4) restrict the helicase to the available strand and was succesfully used in the study of the directionality of other helicases (30). The pattern of activity of these substrates suggests that *Bc*MCM unwinds DNA from 3′→5′ as *Bc*MCM shows robust unwinding activity on substrates with only an accessible 3′ tail (Figure 4, lanes 9–11). No activity is seen on substrates with only an accessible 5′ tail (lanes 15–17). Naturally, a substrate with both 3′ and 5′ tails are unwound (lanes 1–6). This substrate establishes (in comparison to Figure 3) that altering the substrate by adding biotin-streptavidin does not, in and of itself, inhibit helicase activity. In lanes 6 and 18, we note a supershift of the substrate when it is incubated with enzyme in the absence of ATP. Other hexameric helicases, specifically *Escherichia coli* DnaB, have also been shown to bind DNA substrates more tightly when inactive (34–36). The lighter bands on this gel (e.g. lanes 3–5 below the substrate band) represent DNA molecules that have not been modified by streptavidin. This experiment establishes that the polarity of *Bc*MCM is 3′→5′, which is identical to that of other MCM helicases (1,2).

*Bc*MCM is an ATPase that is stimulated by DNA

DNA helicases (specifically, the bacterial replicative helicase DnaB) were initially identified as ssDNA-stimulated ATPases (37). We therefore sought to investigate *Bc*MCM's ATPase activity to more fully characterize its helicase activity. ATPase activity was measured using an NADH-coupled assay that allowed for steady-state kinetics measurements. Using this assay, we observed that *Bc*MCM exhibits a basal specific activity of 1.6 mol ATP/min/mol of enzyme in the absence of DNA (Figure 5). We note that ATP is hydrolyzed in AAA+ ATPases at subunit interfaces (14), suggesting that *Bc*MCM is oligomeric. Mutation of the ATP-binding site (Walker A: K653A) reduced activity to ~30% of wild-type (Figure 5A). This drop in ATPase activity corresponds to an abrogation of helicase activity observed in the same mutant (Figure 3B). A parallel analysis with our *Bc*MCM truncations (C200 and C501), which both harbor the MCM-AAA subdomain, revealed that the ATPase activity is indeed localized to the carboxy-terminal region. Despite observing ATPase activity in the C200 and C501 truncations, we were not able to reliably demonstrate helicase activity for these constructs (data not shown). No activity was seen with the isolated amino-terminal, putative primase domain (N399). Thus, *Bc*MCM is an active ATPase, and this activity is restricted to the MCM-AAA module.

We next examined whether ssDNA could stimulate this ATPase activity. Figure 5B shows that titration of a 61-nt long oligonucleotide molecule stimulated the ATPase of full-length *Bc*MCM by 1.5-fold when present at a DNA:protein ratio of 1:5. Additional stimulation (2.8-fold) could be obtained by using a 4:1 excess of DNA over protein (data not shown). The C200 truncation, which lacks the putative primase domain, was also modestly stimulated by ssDNA.

Both amino- and carboxy-termini of *Bc*MCM can bind ssDNA

Having measured the ability of *Bc*MCM to be stimulated by ssDNA, we next sought to measure the enzyme's binding affinity for ssDNA. This measurement employed tryptophan quenching and exploited our observation that addition of DNA quenches the intrinsic fluorescence of

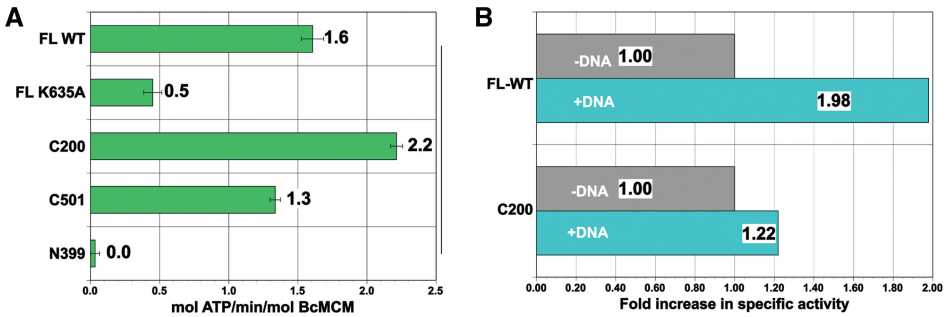

Figure 5. *Bc*MCM is an ATPase, which is stimulated by ssDNA. (**A**) Shown is a summary of ATP hydrolytic activity by various *Bc*MCM constructs in the absence of ssDNA. (**B**) Inclusion of ssDNA (61-mer) led to a 2-fold stimulation of activity by both the full-length, wild-type, enzyme and by a shorter construct (C200). Activity has been normalized to that of the isolated enzyme (values taken from part A). The *Bc*MCM concentration in these experiments is 1 μM.

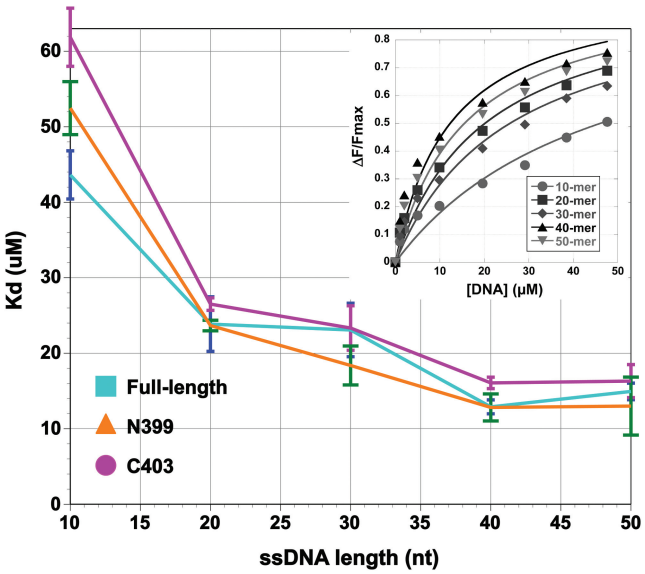

Figure 6. The amino and carboxy terminal regions of *Bc*MCM bind to ssDNA. ssDNA-binding affinities were determined by tryptophan quenching, as described under Materials and Methods section. Three proteins were used: the full-length protein; the N399 truncation that includes the enzyme's amino-terminus; and the C403 truncation, that includes the MCM-AAA subdomain. When added together, the N399 and C403 truncations reconstitute the full-length protein. For each protein, dissociation constant was determined for synthetic ssDNA oligonucleotides 10, 20, 30, 40 and 50 nt long. To determine the dissociation constant for each oligonucleotide, the DNA concentration was titrated from 0–50 μM in reactions containing 1 μM protein. Inset: representative data used to generate this plot. $\Delta F/\Delta F_{max}$ data are shown for the full length *Bc*MCM protein.

*Bc*MCM. Experiments that titrated ssDNA into protein solutions were analyzed according to the following equation: $(\Delta F/\Delta F_{max}) = [DNA]/(K_d + [DNA])$, where ΔF is the amount of fluorescence quenching at a given DNA concentration and ΔF_{max} is the maximum possible quenching. The dissociation constant (K_d) was determined by fitting measurements to the equation above using the Kaleidagraph program. Our fits yielded R^2 values of 0.94 or higher (Figure 6 inset).

We titrated ssDNA of various lengths (10, 20, 30, 40 or 50 nt) to allow calculation of K_d as a function of length (Figure 6). In general, full-length *Bc*MCM binds longer oligos more tightly. We measure a higher affinity to a 20-nt than to a 10-nt species, but the trend plateaus for chains longer than 30-nt. In our experiment, the highest affinity is recorded with a 40-nt molecule (K_d = 12.9 μM, SD = 0.9 μM). Both the N399 and C403 constructs also showed an interaction with ssDNA with similar length dependent properties. In general, the N399 construct (K_d = 13.0 μM, SD = 3.8, for 50-nt oligos) bound slightly tighter than C403 (K_d = 16.3 μM, SD = 2.2, for 50-nt oligos). We note that the full-length protein and the smaller constructs displayed similar dissociation constants for ssDNA. The use of tryptophan quenching to measure dissociation constants, however, makes it difficult to compare the data between the full-length protein and the shorter constructs since the number and exposure of tryptophan residues is different. Nevertheless, the observation that both the amino- and carboxy-terminal domains bind ssDNA suggests the presence of either two binding sites or one extended composite-binding surface in the full-length protein.

*Bc*MCM is a monomer in solution

MCM helicases are AAA+ ATPases and consequently their activity is expected to require oligomerization (14). Structural and biochemical studies of MCM helicases have yielded several different oligomerization states: the eukaryotic assembly is a heterohexamer (1,8), while the archaeal MCMs have been observed as hexamers, heptamers and/or double-hexamers (15,20,21). These oligomers are pre-formed and do not require substrates or co-factors to form. In order to understand the oligomeric state of *Bc*MCM, we measured its native molecular weight using size exclusion chromatography paired with static, multi-angle light scattering (38). These measurements reveal that the molecular weight of full-length *Bc*MCM is 127–138 kDa (Figure 7 and Table 1). Since the molecular weight calculated from the sequence is 117 kDa, we conclude that full-length *Bc*MCM is a monomer in solution (the spread in our data could reflect experimental errors or the presence of small amounts of oligomeric species). The oligomerization state of *Bc*MCM was unchanged when ATP and/or the helicase substrate were included

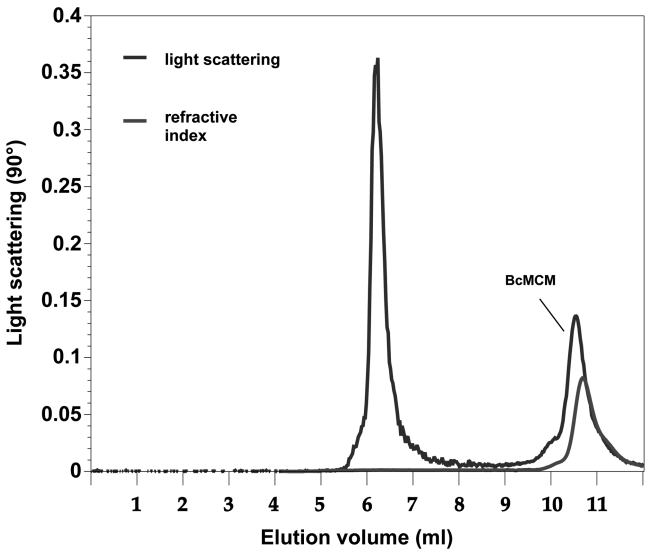

Figure 7. *Bc*MCM is a monomer in solution as revealed by multi-angle laser light scattering. Full length *Bc*MCM (5 mg/ml) in 50 mM Tris–HCl, pH = 7.5, 200 mM NaCl, 5% glycerol, 0.05 mM EDTA was applied to a Shodex KW-804 column equilibrated in 25 mM Tris–HCl, 350 mM KCl, 5 mM $Mg•(OAc)_2$. Light scattering and refractive index signals were measured using a Wyatt Optilab and Dawn EOS system. Scattering curves were processed using the provided Astra software package.

(data not shown). The first peak in Figure 7 represents a very small amount (negligible refractive index) of a large unknown aggregate or particulate.

One explanation for observing a monomer of *Bc*MCM instead of the expected hexamer is that our experimental conditions do not suitably mimic those found *in vivo*. To further pursue this question, we sought to determine whether domains of *Bc*MCM might form specific oligomers in isolation. Light scattering measurements with the C200 and C501 constructs (both contain the MCM-AAA domain) and the N399 construct (containing the putative primase domain) revealed that these constructs exist in mixtures of oligomeric states (Table 1). We speculate that the higher-order assemblies formed by discrete domains of *Bc*MCM may recapitulate the contacts in the active oligomeric species. These data imply that the full-length enzyme may operate as an oligomer when unwinding DNA but that oligomerization may be prevented in the context of the full-length protein, *in vitro*.

It is possible that the relatively high ionic strength of our measurements (350 mM ammonium sulfate) disrupts the oligomer. We used this condition with all our measurements to accommodate the poor behavior of the shorter constructs. To rule out disruption of the oligomer by salt, we performed light scattering measurements with full-length protein in 200 mM sodium chloride and observed only monomers. We note that the *Sulfolobus solfataricus* (*Sso*) archaeal MCM-helicases is oligomeric in 100 mM sodium chloride (39) and requires 1M sodium chloride to dissociate into monomers (27). The most probable hindrance to observing a higher-oligomeric form of *Bc*MCM, however, is its divergent amino-terminus, which in its conserved form provides a significant amount of the binding energy holding MCM hexamers together (10,11). The altered *Bc*MCM amino-terminus may still provide enough binding energy to enable oligomerization *in vivo*, as evident by our *in vitro* biochemical assays, but may not allow for the observation of oligomerization in solution.

Table 1. Molecular masses of full-length and shorter constructs of *Bc*MCM, measured by light scattering

Construct (kDa)	Measured molecular mass	Interpretation
Full-length (117)	127–138	Monomer
	258	Dimer
C200 (95.3)	191	Dimer
	286	Trimer
	388	Tetramer
C501 (60)	72–77	Monomer
N399 (45.8)	44	Monomer
	93	Dimer
	198	Tetramer

*Bc*MCM binds one zinc atom

An important component of the MCM-N domain is a zinc coordination site, which is conserved in virtually all archaeal and eukaryotic MCM helicases, and which disrupts MCM activity when mutated (10–13). Accordingly, inspection of the amino-acid sequence of *Bc*MCM revealed a number of cysteine residues upstream of the MCM-AAA segment that could form a zinc-coordination with a similar position to other MCM proteins. To help understand the relationship between the amino-terminal domain of *Bc*MCM and MCM-N, we sought to detect the presence of a coordinated zinc ion in *Bc*MCM using analytical methods. The zinc content of wild-type *Bc*MCM was measured using inductively-coupled plasma-mass spectrometry (ICP–MS, National Institute of Standards and Technology). While we sought to investigate whether this region was a component of the MCM-N domain, we note that this region has been suggested by Iyer *et al.* to be related to the large subunit of eukaryotic primases rather than the MCM family (29). In addition to testing the wild-type protein, we also mutated a prospective amino-terminal cysteine residue (C261A) in hopes of disrupting the putative zinc coordination site.

This analysis revealed that *Bc*MCM contains 0.11 zinc atoms per mole, while one mole of the C261A mutant contains only 0.03 atoms (Figure 8). These raw values were scaled using control measurements carried out with *Geobacillus stearothermophilus* UvrA (0.32 zinc atoms/mole), whose structure indicates coordination of three zinc ions per monomer (33), and hen egg white lysozyme (0.01 zinc atoms/mole), which is free of zinc. We conclude from our study that *Bc*MCM coordinates one zinc ion ($0.11 \times 3/0.32 = 1.03$), whereas the C261A mutant bound only 0.28 zinc ions per mol of protein ($= 0.03 \times 3/0.32$), indicating that cysteine 261 contributes to the coordination site. Scaling the data in this way is justified owing to the fact that low concentrations of zinc in the protein

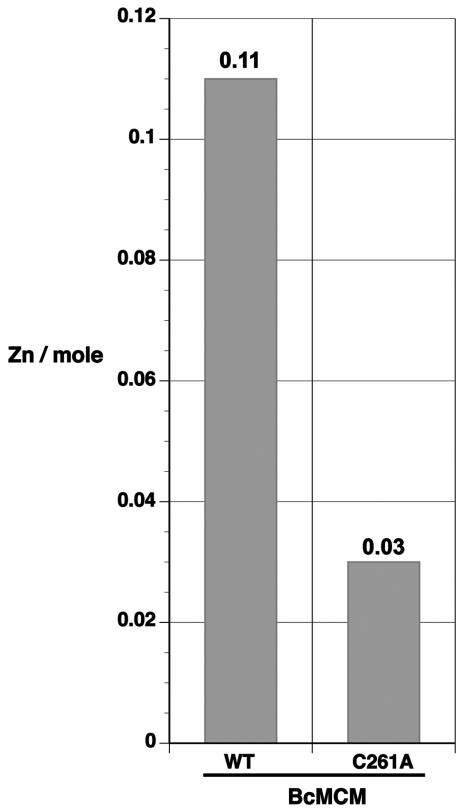

Figure 8. Full-length *Bc*MCM binds a zinc atom. Purified protein samples were prepared for zinc analysis by dialysis against 20 mM Tris–HCl/pH 7.5, 400 mM sodium chloride, 5% glycerol and 0.5 mM EDTA. Samples were subsequently analyzed for zinc using inductively coupled plasma-mass spectrometry (ICP–MS). Shown here are results for full-length, wild-type *Bc*MCM and full-length, C261A *Bc*MCM mutant.

samples are a challenge to measure by ICP–MS due to contamination and instrument background.

*Bc*MCM has no detectible primase activity

An intriguing prediction made by the informatic analysis of the *Bc*MCM amino-acid sequence is the presence of segments homologous to the archaeo-eukaryotic family of primases (residues 1–183) and to the large subunit of eukaryotic primase (residues 183–417) (Figure 2A). The assay that we used to evaluate this prediction calls for incubation of *Bc*MCM with ssDNA (ϕX174, NEB) and ribonucleotides. If a primer is synthesized, then the Klenow fragment is able to synthesize a DNA complement to the ϕX174plasmid using radiolabeled dNTPs. Using this assay and buffer conditions described for the *E. coli* DnaG primase, we did not observe any primase activity while using *Ec*DnaG itself as the positive control (data not shown) (32).

DISCUSSION

The main goal of this study is to develop the phage *Bc*MCM-helicase as a tractable model for the more elaborate heterohexameric helicases found in eukaryotes. The study of MCM helicases is important given their critical roles in both initiating eukaryotic DNA replication and in fork progression after replication begins. Here, we describe the initial biochemical analysis of an unusual MCM helicase found in a phage genome integrated within that of the bacterium *B. cereus*. Phage proteins such as the DNA, RNA polymerases and helicases have long proven to be excellent models to study core functions of more complicated eukaryotic entities and we expect that *Bc*MCM may play a similar role (40,41).

The *Bc*MCM-helicase is interesting for two reasons. First, MCM-helicases are typical to eukaryotes and archaea, but have not been previously described in bacteria (although a homolog has now also been found in *Exiguibacterium* species AT1b, NCBI accession #ZP 02990802). Second, only the carboxy-terminal AAA domain of *Bc*MCM (the MCM-AAA domain) is related to eukaryotic and archaeal MCM proteins, whereas the amino-terminal portion of *Bc*MCM is divergent in sequence. The typical amino-terminus of the MCM family (MCM-N) appears to be missing from *Bc*MCM. This is especially puzzling given the importance of MCM-N to the unwinding reaction (13,42,43). Rather than the conserved MCM-N domain, *Bc*MCM contains a stretch of 500 amino acids that exhibits weak similarity to the archaeo-eukaryotic primase family and the large subunit of eukaryotic primase subunit (Figures 1 and 2) (28,29).

In this study, we have shown that *Bc*MCM unwinds DNA in an ATP-dependent manner with $3' \rightarrow 5'$ directionality. These properties resemble those of conventional MCM-helicases despite the lack of a standard MCM amino-terminus. From our experiments, we conclude that *Bc*MCM helicase activity arises from the cooperation between its conserved MCM-AAA domain and a paralog of the MCM-N domain. We speculate that the divergent amino-terminus may allow *Bc*MCM to serve as an alternate model for MCM function, since study of different structures (i.e. the MCM-N versus the *Bc*MCM amino-terminus) that perform the same function often proves illuminating.

Previous studies have established that the MCM-N domain regulates the DNA unwinding activity and influences the ATPase activity (13,42,43). Our results suggest that an intact amino-terminus of *Bc*MCM is required for unwinding activity. By contrast, the MCM-N domain of the archaeal MCMs can stimulate unwinding activity, but is not in and of itself necessary. Our work with *Bc*MCM truncations indicates that the amino-terminus also diminishes the intrinsic ATPase activity of the MCM-AAA (compare the activity of the C200 construct to that of the full-length protein in Figure 5). These results are consistent with those with the *Aeropyrum pernix* and the *Archaeoglobus fulgidus* MCM proteins, in which amino-terminal truncations also demonstrated increased ATPase activity (13,42).

Our studies suggest that *Bc*MCM binds ssDNA with both the MCM-AAA domain and the amino-terminus (Figure 6). Within error, our measurements suggest that both isolated segments display a length-dependent affinity to ssDNA. The affinity displayed by the amino-terminus

and MCM-AAA domains is virtually identical to that of full-length. The tryptophan-binding assay makes it complicated to directly compare these results due to the differing number of tryptophan residues and their degrees of exposure. Nevertheless, our data argues for DNA-binding sites within the amino-terminal and MCM-AAA segments and these may be organized into one composite-binding site in the full-length chain. It should also be noted that the MCM-N and MCM-AAA domains of the *Sso*MCM protein bind ssDNA with similar affinities (16).

The presence of a zinc-coordination module in the MCM-N region is another hallmark of MCM helicase subunits. Several studies have implicated this structural feature in contacts to DNA (12,13), though the molecular details of these contacts are not well understood. We have measured the zinc content of full-length, wild-type *Bc*MCM and obtained evidence for the presence of a zinc atom as well as identifying cysteine 261 as one of the coordinating residues (Figure 8). The finding that *Bc*MCM harbors a zinc ion in its divergent amino-terminus lends support to the idea that this element substitutes for the missing MCM-N region.

Instead of the expected MCM-N domain found in MCM helicases, the amino-terminus of *Bc*MCM displays weak similarities to eukaryotic/archaeal primases (29). Efforts to measure primase activity using a standard assay did not reveal primer synthesis. While we may have failed to detect primase activity due to sub-optimal buffer conditions or the indirect nature of the primase assay, an alternate view suggests that the sequence in question is no longer competent to synthesize RNA primers. While sequence analysis shows the presence of two aspartate residues (D74 and D76), which would be expected to lie in the primase active site, sequence divergence throughout the rest of the region does not allow us to discern whether other residues required by a primase are truly present (Figure 1).

Taken together, our experiments suggest that the crucial role played by the MCM-N domain of conventional MCM-helicases in ATP hydrolysis, DNA binding and unwinding is being carried out by an alternate structure formed at the amino-terminus of *Bc*MCM. This conclusion relies on three observations: the ability of the amino-terminus to modulate both the helicase and ATPase activity of the MCM-AAA domain; its role in binding DNA; and its ability to coordinate zinc.

Efforts to establish the oligomeric state of the active entity revealed that *Bc*MCM is present in solution predominantly as a monomer. These results are unchanged when ATP and magnesium are included in the analysis. The finding that *Bc*MCM is a monomer in solution is unexpected and distinguishes it from other MCM helicases, which are active as oligomers (usually hexamers), and do not require co-factors or substrates to form. We suspect that our inability to detect the active hexameric form is due to experimental conditions that do not suitably mimic those found *in vivo*.

We fully expect that the active *Bc*MCM helicase is an oligomer, most likely a hexamer, which assembles under the right conditions. We rest this conclusion on three observations. First, *Bc*MCM exhibits the sequence motifs that are hallmarks of the AAA+ family of ATPases. The active AAA+ ATPase is always an oligomer since nucleotides are bound and hydrolyzed at subunit interfaces (14). Indeed, isolated subunits of the yeast heterohexameric MCM2-7 do not hydrolyze nucleotide until provided with the appropriate neighboring subunit (7). Moreover, the ssDNA-ATPase stimulation is also indicative of oligomerization. Recent crystallographic studies of closed ring, hexameric helicases (similar in form to the MCM helicases) indicate that DNA stimulates ATPase activity by properly orienting neighboring ATPase domains within the hexamer to optimize nucleotide hydrolysis (25,44). Second, all characterized MCM-helicases unwind DNA with identical polarity to *Bc*MCM ($3' \rightarrow 5'$) and do so as oligomers (usually as hexamers). Lastly, oligomerization is strongly suggested by light scattering analysis of the shorter constructions, which could be recapitulating the subunit interactions present in the active helicase (Table 1).

The idea that an active ring helicase can be assembled from monomers has precedent. One example is the SV40 large T antigen helicase, which can only form the active oligomer by assembling on DNA from monomers (45). The bacterial helicase DnaC serves as another example (46).

The question of the biological function, if any, of the *Bc*MCM helicase in the physiology of *B. cereus*—with its clear evolutionary implications—is beyond the scope of this study. We note that the *B. cereus* genome harbors a complete set of the typical bacterial-replication initiation and replisome proteins (DnaA, DnaB, DnaI, etc.) and thus we do not discern any obvious reason why the presence of additional genes from an inserted phage would provide a selective advantage.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank past and present members of the Jeruzalmi laboratory for useful discussions, Stanley Lo and Nozomi Sakakibara for assistance with some of the experiments. We thank Mike O'Donnell, Matt Michael, Brian Chait and Danaya Pakotiprapha for providing reagents and advice. Some of the mutants in this study were prepared by the Harvard College Class, LS100R. We are grateful to Nicole Francis for help with experiments and with comments on the manuscript.

FUNDING

National Institutes of Health [grant number GM084162 to D.J.]; National Science Foundation (grant number MCB-0815646 to Z.K.).

Conflict of interest statement. None declared.

REFERENCES

1. Forsburg,S.L. (2004) Eukaryotic MCM proteins: beyond replication initiation. *Microbiol. Mol. Biol. Rev.*, **68**, 109–131.
2. Sakakibara,N., Kelman,L.M. and Kelman,Z. (2009) Unwinding the structure and function of the archaeal MCM helicase. *Mol. Microbiol.*, **72**, 286–296.
3. Ying,C.Y. and Gautier,J. (2005) The ATPase activity of MCM2-7 is dispensable for pre-RC assembly but is required for DNA unwinding. *EMBO J.*, **24**, 4334–4344.
4. Moyer,S.E., Lewis,P.W. and Botchan,M.R. (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.
5. Pacek,M., Tutter,A.V., Kubota,Y., Takisawa,H. and Walter,J.C. (2006) Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol. Cell*, **21**, 581–587.
6. Ishimi,Y. (1997) A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J. Biol. Chem.*, **272**, 24508–24513.
7. Davey,M.J., 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.
8. Kanter,D.M., Bruck,I. and Kaplan,D.L. (2008) Mcm subunits can assemble into two different active unwinding complexes. *J. Biol. Chem.*, **283**, 31172–31182.
9. Bochman,M.L. and Schwacha,A. (2008) The Mcm2-7 complex has in vitro helicase activity. *Mol. Cell*, **31**, 287–293.
10. Fletcher,R.J., Bishop,B.E., Leon,R.P., Sclafani,R.A., Ogata,C.M. and Chen,X.S. (2003) The structure and function of MCM from archaeal M. Thermoautotrophicum. *Nat. Struct. Biol.*, **10**, 160–167.
11. Lium,W., Pucci,B., Rossi,M., Pisani,F.M. and Ladenstein,R. (2008) Structural analysis of the Sulfolobus solfataricus MCM protein N-terminal domain. *Nucleic Acids Res.*, **36**, 3235–3243. [PMCID: 2425480].
12. Poplawski,A., Grabowski,B., Long,S.E. and Kelman,Z. (2001) The zinc finger domain of the archaeal minichromosome maintenance protein is required for helicase activity. *J. Biol. Chem.*, **276**, 49371–49377.
13. Atanassova,N. and Grainge,I. (2008) Biochemical characterization of the minichromosome maintenance (MCM) protein of the crenarchaeote Aeropyrum pernix and its interactions with the origin recognition complex (ORC) proteins. *Biochemistry.*, **47**, 13362–13370.
14. Duderstadt,K.E. and Berger,J.M. (2008) AAA+ ATPases in the initiation of DNA replication. *Crit. Rev. Biochem. Mol. Biol.*, **43**, 163–187.
15. Costa,A., Pape,T., van Heel,M., Brick,P., Patwardhan,A. and Onesti,S. (2006) Structural basis of the Methanothermobacter thermautotrophicus MCM helicase activity. *Nucleic Acids Res.*, **34**, 5829–5838.
16. McGeoch,A.T., Trakselis,M.A., Laskey,R.A. and Bell,S.D. (2005) Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nat. Struct. Mol. Biol.*
17. Iyer,L.M., Leipe,D.D., Koonin,E.V. and Aravind,L. (2004) Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.*, **146**, 11–31.
18. 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.
19. 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.
20. Yu,X., VanLoock,M.S., Poplawski,A., Kelman,Z., Xiang,T., Tye,B.K. and Egelman,E.H. (2002) The Methanobacterium thermoautotrophicum MCM protein can form heptameric rings. *EMBO Rep.*, **3**, 792–797.
21. 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.
22. Costa,A., Pape,T., van Heel,M., Brick,P., Patwardhan,A. and Onesti,S. (2006) Structural studies of the archaeal MCM complex in different functional states. *J. Struct. Biol.*, **156**, 210–219.
23. Wang,G., Klein,M.G., Tokonzaba,E., Zhang,Y., Holden,L.G. and Chen,X.S. (2008) The structure of a DnaB-family replicative helicase and its interactions with primase. *Nat. Struct. Mol. Biol.*, **15**, 94–100.
24. Bailey,S., Eliason,W.K. and Steitz,T.A. (2007) Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science.*, **318**, 459–463.
25. Enemark,E.J. and Joshua-Tor,L. (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature.*, **442**, 270–275.
26. Bae,B., Chen,Y., Costa,A., Onesti,S., Brunzelle,J., Lin,Y., Cann,I. and Nair,S. (2009) Insights into the architecture of the replicative helicase from the structure of an archaeal MCM homolog. *Structure*, **17**, 211–22.
27. Brewster,A.S., Wang,G., Yu,X., Greenleaf,W.B., Carazo,J.M., Tjajadia,M., Klein,M. and Chen,X. (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.
28. McGeoch,A.T. and Bell,S.D. (2005) Eukaryotic/archaeal primase and MCM proteins encoded in a bacteriophage genome. *Cell*, **120**, 167–168.
29. Iyer,L.M., Koonin,E.V., Leipe,D.D. and Aravind,L. (2005) Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. *Nucleic Acids Res.*, **33**, 3875–3896.
30. Shin,J.H. and Kelman,Z. (2006) DNA unwinding assay using streptavidin-bound oligonucleotides. *BMC Mol. Biol.*, **7**, 43.
31. Kiianitsa,K., Solinger,J.A. and Heyer,W.D. (2003) NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Anal. Biochem.*, **321**, 266–271.
32. Rowen,L. and Kornberg,A. (1978) Primase, the dnaG protein of Escherichia coli. An enzyme which starts DNA chains. *J. Biol. Chem.*, **253**, 758–764.
33. Pakotiprapha,D., Inuzuka,Y., Bowman,B.R., Moolenaar,G.F., Goosen,N., Jeruzalmi,D. and Verdine,G.L. (2008) Crystal structure of Bacillus stearothermophilus UvrA provides insight into ATP-modulated dimerization, UvrB interaction, and DNA binding. *Mol. Cell*, **29**, 122–133.
34. Galletto,R., Jezewska,M.J. and Bujalowski,W. (2003) Interactions of the Escherichia coli DnaB helicase hexamer with the replication factor the DnaC protein. Effect of nucleotide cofactors and the ssDNA on protein-protein interactions and the topology of the complex. *J. Mol. Biol.*, **329**, 441–465.
35. Learn,B.A., Um,S.J., Huang,L. and McMacken,R. (1997) Cryptic single-stranded-DNA binding activities of the phage lambda P and Escherichia coli DnaC replication initiation proteins facilitate the transfer of E. coli DnaB helicase onto DNA. *Proc. Natl Acad. Sci. USA*, **94**, 1154–1159.
36. Arai,K. and Kornberg,A. (1981) Mechanism of dnaB protein action. III. Allosteric role of ATP in the alteration of DNA structure by dnaB protein in priming replication. *J. Biol. Chem.*, **256**, 5260–5266.
37. Arai,K. and Kornberg,A. (1981) Mechanism of dnaB protein action. II. ATP hydrolysis by dnaB protein dependent on single- or double-stranded DNA. *J. Biol. Chem.*, **256**, 5253–5259.
38. Wyatt,P. (1993) Light scattering and the absolute characterization of macromolecules. *Analytica Chimica Acta*, **272**, 1–40.
39. Pucci,B., De Felice,M., Rocco,M., Esposito,F., De Falco,M., Esposito,L., Rossi,M. and Pisani,F.M. (2007) Modular organization of the Sulfolobus solfataricus mini-chromosome maintenance protein. *J. Biol. Chem.*, **282**, 12574–12582.
40. Jeruzalmi,D., O'Donnell,M. and Kuriyan,J. (2002) Clamp loaders and sliding clamps. *Curr. Opin. Struct. Biol.*, **12**, 217–224.
41. Steitz,T.A. and Yin,Y.W. (2004) Accuracy, lesion bypass, strand displacement and translocation by DNA polymerases. *Philos Trans R Soc Lond, B, Biol. Sci.*, **359**, 17–23.
42. Grainge,I., Scaife,S. and Wigley,D.B. (2003) Biochemical analysis of components of the pre-replication complex of Archaeoglobus fulgidus. *Nucleic Acids Res.*, **31**, 4888–4898. PMCID: 169903.
43. Barry,E.R., McGeoch,A.T., Kelman,Z. and Bell,S.D. (2007) Archaeal MCM has separable processivity, substrate choice and helicase domains. *Nucleic Acids Res.*, **35**, 988–998.

44. Singleton,M.R., Sawaya,M.R., Ellenberger,T. and Wigley,D.B. (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, **101**, 589–600.
45. Dean,F.B., Borowiec,J.A., Eki,T. and Hurwitz,J. (1992) The simian virus 40 T antigen double hexamer assembles around the DNA at the replication origin. *J. Biol. Chem.*, **267**, 14129–14137.
46. Davey,M.J. and O'Donnell,M. (2003) Replicative helicase loaders: ring breakers and ring makers. *Curr. Biol.*, **13**, R594–R596.
47. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797. PMCID: 390337.
48. Pei,J. and Grishin,N.V. (2001) AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics.*, **17**, 700–712.
49. Henikoff,S. and Henikoff,J.G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl Acad. Sci. USA*, **89**, 10915–10919.