Structure of a double hexamer of the *Pyrococcus furiosus* minichromosome maintenance protein N-terminal domain

Martin Meagher and Eric J. Enemark*

Department of Structural Biology, St Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105, USA. *Correspondence e-mail: eric.enemark@stjude.org

Received 24 March 2016

Accepted 27 May 2016

Edited by F. T. Tsai, Baylor College of Medicine, Houston, USA

Keywords: DNA replication; helicase; MCM; minichromosome maintenance; *Pyrococcus furiosus*.

PDB reference: PfMCM N-terminal domain double hexamer, 5iy0

Supporting information: this article has supporting information at journals.iucr.org/f

The crystal structure of the N-terminal domain of the *Pyrococcus furiosus* minichromosome maintenance (MCM) protein as a double hexamer is described. The MCM complex is a ring-shaped helicase that unwinds DNA at the replication fork of eukaryotes and archaea. Prior to replication initiation, the MCM complex assembles as an inactive double hexamer at specific sites of DNA. The presented structure is highly consistent with previous MCM double-hexamer structures and shows two MCM hexamers with a head-to-head interaction mediated by the N-terminal domain. Minor differences include a diminished head-to-head interaction and a slightly reduced inter-hexamer rotation.

1. Introduction

DNA replication is an essential process for organisms to duplicate their genetic material in preparation for cell division. Central to the process is a helicase enzyme that unwinds parental double-stranded DNA (dsDNA) to expose the individual strands for polymerases to use as templates in the synthesis of new DNA. In eukaryotes and archaea, the central engine of the helicase is a minichromosome maintenance (MCM) complex consisting of six proteins arranged in a ring (Deegan & Diffley, 2016). While the MCM complex of eukaryotes has six related but different proteins (Mcm2–7), the MCM complex of archaea often consists of a single protein in a homohexamer (Miller & Enemark, 2015).

In preparation for replication, the MCM complex is first loaded onto DNA at a replication origin as two hexamers arranged as a double hexamer with a head-to-head interaction mediated by the N-terminal domain. This double hexamer is not able to unwind duplex DNA. Instead, the double hexamer can passively slide over the DNA and fall off exposed ends, indicating that it topologically encircles double-stranded DNA without the ability to unwind it (Remus *et al.*, 2009; Evrin *et al.*, 2009). In eukaryotes, the activation of DNA unwinding in the MCM complex correlates with the conversion of the double hexamer to two single hexamers and also the association of the factor Cdc45 and the tetrameric GINS complex to form a larger CMG complex (Cdc45, Mcm2–7, GINS; Moyer *et al.*, 2006; Ilves *et al.*, 2010; Costa *et al.*, 2011). The helicase-inactive double hexamer therefore represents a key species within the precisely controlled pathway towards the activation of DNA unwinding by MCM.

Electron-microscopy images of the MCM complex of the archaeon *Methanobacterium thermautotrophicus* (MtMCM)

Table 1

Macromolecule-production information.

Source organism	<i>P. furiosus</i>
DNA source	pLE009.3†
Cloning vector	pSMT3‡
Expression vector	pSMT3‡
Expression host	<i>E. coli</i> BL21(DE3) RIPL
Complete amino-acid sequence of the construct produced§	SVDREEMIERFANFLREYTDGPNVYRGKITDL-LTITPKRSVAIDWMHLNSFDSELAHEVIENPE-EGISAAEDATQIVLRDFQREDVGKIHARFYN-LPETLMVKDIGAEHINKLIQVEGIVTRVGEIK-PFVSVAVFVCKDCGHEMIVPQKPYESLEKVKK-CEQCGSKNIELDVNKSSFVNFQSFRIQDRPET-LKGGEMPRFIDGILLDDIVDVALPGDRVIVTG-ILRVVLEKREKTPIFRKILEVNHIEPVSKET

† Froelich *et al.* (2014). ‡ Mossessova & Lima (2000). § An N-terminal His₆-SUMO tag was removed by specific cleavage by Ulp1 protease (Mossessova & Lima, 2000).

revealed the dumbbell shape of the double hexamer (Chong *et al.*, 2000), and a crystal structure of the double-hexamer assembly of the N-terminal domain (*MtMCM-N*) revealed the atomic details of the inter-hexamer interactions within a double hexamer and also the architecture of the MCM N-terminal domain itself (Fletcher *et al.*, 2003). In this structure, the individual *MtMCM-N* subunit consists of three subdomains: a helical subdomain A, a zinc-binding subdomain B and an OB (oligonucleotide/oligosaccharide-binding)-fold (Murzin, 1993) subdomain C (Fletcher *et al.*, 2003). The fundamental subdomain architecture first identified for the crystal structure of *MtMCM-N* (Fletcher *et al.*, 2003) is conserved in archaeal and eukaryotic MCM N-terminal domain structures (Fletcher *et al.*, 2003; Liu *et al.*, 2008; Fu *et al.*, 2014; Li *et al.*, 2015), such as the MCM N-terminal domain of *Sulfolobus solfataricus* (*SsoMCM-N*). The two hexamers of the *MtMCM-N* double hexamer interact head-to-head through an interaction of the B subdomains, including an extensive interaction network involving Arg161 of each subunit (Fletcher *et al.*, 2003). The 3.8 Å resolution cryo-EM structure of the yeast *Mcm2-7* (*ScMcm2-7*) double hexamer purified from G1-arrested cells shows the defined orientation of the two heterohexamers and also extensive interactions mediating the association (Li *et al.*, 2015). Thus, the double hexamers of *MtMCM* and *ScMcm2-7* are held together by numerous interactions across the hexamer-hexamer interface.

Our laboratory previously presented the hexameric ring structure of the N-terminal domain of the MCM protein from *Pyrococcus furiosus* (*PfMCM-N*) bound to single-stranded DNA (ssDNA) via the OB-fold C subdomain (PDB entry 4pog; Froelich *et al.*, 2014). In this manuscript, we report the crystal structure that we determined when attempting to crystallize *PfMCM-N* bound to dsDNA. Although we did not detect electron density that would correspond to discretely bound dsDNA, the crystal structure revealed a head-to-head double-hexamer architecture that was highly similar to those of *MtMCM-N* and *ScMcm2-7*. This structure therefore provides an additional model of the species formed when MCM rings are first loaded at a replication origin. In contrast to other double-hexamer structures, we find much less extensive interactions across the hexamer-hexamer interface

Table 2

Crystallization.

Method	Vapor diffusion, hanging drop
Plate type	Hampton VDX 24-well plate
Temperature (K)	295
Protein concentration (mg ml ⁻¹)	10 (based on absorbance at 280 nm)
Buffer composition of protein solution	20 mM HEPES pH 7.6, 200 mM NaCl, 5 mM β-mercaptoethanol
Composition of reservoir solution	0.1 M HEPES pH 7.1, 0.1 M magnesium chloride, 10% PEG 4000
Volume and ratio of drop	1 µl:1 µl
Volume of reservoir (µl)	1000

in the double hexamer of *PfMCM-N* than in *MtMCM-N* or *ScMcm2-7*. Further details of the *PfMCM-N* double-hexamer crystal structure follow.

2. Materials and methods

2.1. Macromolecule production

The *PfMCM-N* protein was expressed in *Escherichia coli* as described previously with expression plasmid pLE009.3 (Froelich *et al.*, 2014). The expression plasmid generates an N-terminal His₆-SUMO fusion (Mossessova & Lima, 2000) of residues 1–256 of the minichromosome maintenance (MCM) protein of *P. furiosus* (Table 1). The protein was purified as described by Froelich *et al.* (2014), including removal of the SUMO tag by specific cleavage by Ulp1 protease (Mossessova & Lima, 2000).

2.2. Crystallization

Solutions of the *PfMCM-N* protein at a concentration of 10 mg ml⁻¹ (as determined from the absorbance at 280 nm wavelength using a calculated extinction coefficient of 11 460 M⁻¹ cm⁻¹) and a hybrid ssDNA/dsDNA substrate (HPLC-purified TTTTTTTTTTTTTTTTTTTTTTTTACTGCTATAGCAGTTT; the bases in bold form a 14-mer palindrome) were mixed in a 5:1 ratio and then directly screened for crystallization by hanging-drop vapor diffusion. Single crystals grew from drops consisting of 1:1 protein:well solution over the course of one week at 22°C. The well solution consisted of 0.1 M HEPES pH 7.1, 0.1 M magnesium chloride, 10% PEG 4000 (Table 2). Crystals of *PfMCM-N* can grow in many forms (including those described in Froelich *et al.*, 2014). We note that the crystal form reported here has only been obtained by co-crystallization of *PfMCM-N* with the ssDNA/dsDNA substrate and emphasize that this does not prove that DNA is required for growth of this specific crystal form.

2.3. Data collection and processing

Crystals were harvested with a fiber loop, cryoprotected by quickly passing them through a solution of ethylene glycol and well solution in a 1:2 ratio and flash-cooled in liquid nitrogen. Cooled crystals were robotically mounted on SER-CAT beamline 22-BM at the Advanced Photon Source (APS) of Argonne National Laboratory and evaluated for diffraction properties. Diffraction data were collected from a single crystal at a wavelength of 1.0 Å in 0.5° oscillations over a total

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	SER-CAT beamline 22-BM, APS
Wavelength (Å)	1.0
Temperature (K)	100
Detector	MAR Mosaic 225 mm CCD
Crystal-to-detector distance (mm)	324
Rotation range per image (°)	0.5
Total rotation range (°)	180
Exposure time per image (s)	9.4
Space group	C2
<i>a</i> , <i>b</i> , <i>c</i> (Å)	211.427, 122.610, 111.892
α , β , γ (°)	90, 92.35, 90
Mosaicity (°)	0.44–1.12
Resolution range (Å)	50–3.00 (3.11–3.00)
Total No. of reflections	215009 (19914)
No. of unique reflections	57090 (5666)
Completeness (%)	100.0 (100.0)
Multiplicity	3.8 (3.5)
$\langle I/\sigma(I) \rangle$	13.0 (2.8)
<i>R</i> _{meas} (%)	16.8 (55.7)
Overall <i>B</i> factor from Wilson plot (Å ²)	64

of 180° of crystal rotation. Data were integrated and scaled to 3.0 Å resolution with *HKL-2000* (Otwinowski & Minor, 1997) in a *C*-centered monoclinic lattice (Table 3).

2.4. Structure solution and refinement

The structure was solved by the molecular-replacement method with *Phaser* (McCoy *et al.*, 2007), which placed six crystallographically unique monomers of *PfMCM-N* (PDB entry 4pof; Froelich *et al.*, 2014) in a head-to-head double hexamer around a crystallographic twofold axis (Fig. 1). The

Table 4

Structure refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	50–3.00 (3.16–3.00)
Completeness (%)	99.39 (96.92)
σ Cutoff	0 (0)
No. of reflections, working set	54114 (7641)
No. of reflections, test set	2877 (393)
Final <i>R</i> _{cryst}	0.205 (0.315)
Final <i>R</i> _{free}	0.223 (0.354)
Cruickshank DPI	0.9446
No. of non-H atoms	
Protein	12186
Ion	6
Ligand	0
Water	0
Total	12192
R.m.s. deviations	
Bonds (Å)	0.009
Angles (°)	1.342
Average <i>B</i> factors (Å ²)	
Protein	52
Ion	72
Ligand	N/A
Water	N/A
Ramachandran plot	
Favored regions (%)	89.6
Additionally allowed regions (%)	10.0
Outliers (%)	0.4†

† Six copies of residue Asp209 (see text).

coordinates were refined with *REFMAC5* (Murshudov *et al.*, 2011) against data to 3.0 Å resolution (Table 4). A Ramachandran plot calculated by *PROCHECK* (Laskowski *et al.*, 1993) gave the following statistics: core, 1209 residues

Figure 1

PfMCM-N double-hexamer structure viewed perpendicular (*a*) and parallel (*b*) to the central channel. Each subunit is uniquely colored in each hexamer, with the zinc-binding domain at the interface between the two hexamers (Zn ions in gray). The ATPase domains would be located above the upper hexamer and below the lower hexamer in (*a*).

(89.6%); allowed, 135 (10.0%); generously allowed, 0 (0%); disallowed, 6 (0.4%). Each of the six disallowed residues is one of the copies of Asp209, which resides at the end of a tight structural turn. The Asp209 φ/ψ angles are just outside the allowed region of the plot ($\varphi/\psi = 60.5/-117.9, 60.3/-118.3, 61.0/-118.1, 61.5/-117.9, 61.5/-116.3$ and $60.1/-116.9^\circ$ for chains *A, B, C, D, E* and *F*, respectively), which are very close to the angles observed in the allowed region for the wild-type hexamer (PDB entry 4pof; $\varphi/\psi = 53.1/-121.6, 55.1/-124.4, 52.5/-123.1, 49.8/-122.9, 49.4/-123.9$ and $47.4/-126.1^\circ$ for chains *A, B, C, D, E* and *F*, respectively). Although the protein was co-crystallized with a DNA substrate, no significant difference electron density was observed after the *Pf*MCM-N portion of the crystal structure had been fully modeled, indicating that DNA does not significantly occupy well defined, discrete positions in the crystal structure.

3. Results and discussion

The individual subunits of the *Pf*MCM-N double hexamer have an overall structure and subdomain architecture that is equivalent to other MCM N-terminal domain crystal struc-

tures (Fletcher *et al.*, 2003; Liu *et al.*, 2008; Froelich *et al.*, 2014; Fu *et al.*, 2014). In particular, the structure does not notably differ from the single-hexamer structure of the same *Pf*MCM-N construct (PDB entry 4pof; Froelich *et al.*, 2014). The minimum radius of the channel (calculated from a model with side chains truncated to alanines as in Miller *et al.*, 2014) is 13.8 Å (residue Arg234), which is large enough to encircle B-form dsDNA. The minimum radius is close to that calculated for the *Pf*MCM-N single hexamer (13.4 Å), and is slightly larger than those calculated for *Mt*MCM (11.4 Å; Fletcher *et al.*, 2003) and *Sso*MCM (11.4 Å; Liu *et al.*, 2008) (Fig. 2). Each of these MCM N-terminal domain structures would be large enough at the minimum radius to accommodate B-form dsDNA if the side chains are truncated to alanine to mimic flexibility among the side chains. When full side chains are included for the structure of *Sso*MCM (Liu *et al.*, 2008) to consider the case of inflexible side chains, a minimum radius of 7.8 Å is calculated, which is likely to be too narrow to accommodate B-form DNA.

The overall architecture of the *Pf*MCM-N double hexamer is very similar to that of *Mt*MCM-N (Figs. 3*a* and 3*b*), with two hexamers that interact head-to-head through the zinc-binding

Figure 2

The architecture of one hexamer within the *Pf*MCM-N double hexamer is highly similar to other MCM hexamer structures. The six subunits are uniquely colored and the minimum channel radius value is reported for each MCM-N crystal structure with side-chain positions truncated to alanine. The structures shown are one hexamer of the *Pf*MCM-N double hexamer (*a*), one hexamer of the *Mt*MCM-N double hexamer (Fletcher *et al.*, 2003) (*b*), the *Pf*MCM-N single-hexamer structure (Froelich *et al.*, 2014) (*c*), the *Sso*MCM-N single-hexamer structure (Liu *et al.*, 2008) (*d*) and the N-terminal domain portion of one hexamer of the full-length *Sc*Mcm2–7 double-hexamer structure (Li *et al.*, 2015) (*e*).

subdomain B. Although the zinc-coordination sphere of *Sso*MCM is different from those of other MCM proteins (Liu *et al.*, 2008), its overall subdomain B structure appears to be compatible with the double-hexamer interface present in the *Pf*MCM-N double hexamer. The channels of the two *Pf*MCM-N hexamers appear collinear and form one contiguous channel (Fig. 1*b*). The co-axial hexamer channels of the *Pf*MCM-N double hexamer contrasts the offset hexamers of *Sc*Mcm2–7 (Figs. 3*b* and 3*c*) that have been suggested to play a role in the initial melting of dsDNA at the onset of replication (Li *et al.*, 2015).

The *Pf*MCM-N and *Mt*MCM-N double hexamers differ slightly in the relative rotation of the constituent hexamers about the channel axis, as defined by the Zn-atom positions. When projected onto a plane perpendicular to the channel axis, the six Zn atoms of one *Pf*MCM-N hexamer are rotated 8.9° relative to the six Zn atoms of the other hexamer. In contrast, the two *Mt*MCM-N hexamers are rotated 19.5° with respect to each other based on analogous Zn-atom projections. As a result, the Zn atoms of neighboring hexamers are further apart in *Mt*MCM-N than in *Pf*MCM-N (Fig. 4). In both cases, the Zn atoms themselves occupy similar heights along the channel axis. The centroids of the six Zn atoms of each hexamer are separated by 6.4 Å for *Pf*MCM-N and by 6.8 Å for *Mt*MCM-N.

Within the *Pf*MCM-N double hexamer, the two hexamers interact only through modest van der Waals interactions and thus do not appear to be strongly tethered together. These

limited interactions contrast with the extensive hexamer–hexamer interactions observed for the double hexamers of *Mt*MCM-N (Fletcher *et al.*, 2003; Fig. 4*a*) and *Sc*Mcm2–7 (Li *et al.*, 2015). As a result, the double hexamer that we have captured here is not likely to be robustly maintained without additional stabilizing factors. An intriguing possibility is that the double-hexamer form of *Pf*MCM-N could be stabilized by dsDNA bound within the channel of each hexamer. Although we do not observe any electron density to ascribe to DNA in this crystal structure, the inclusion of dsDNA in the crystallization sample may have enhanced double-hexamer formation and thus facilitated the initiation of the growth of a head-to-head double-hexamer crystal form. In such a scenario, DNA could promote the initiation of growth of double-hexamer crystals, but DNA might not be needed for the continued growth of these crystals after crystallization has started.

Overall, the structure of the *Pf*MCM-N double hexamer reported here is highly consistent with the double-hexamer structures of *Mt*MCM-N (Fletcher *et al.*, 2003) and *Sc*Mcm2–7 (Li *et al.*, 2015). As such, this *Pf*MCM-N double-hexamer structure represents a plausible species along the pathway towards regulated DNA unwinding at the replication fork, analogous to the role of the *Sc*Mcm2–7 double hexamer in MCM loading prior to the onset of unwinding. The limited interactions across the two hexamers of the *Pf*MCM-N double hexamer suggests that robust double-hexamer interactions are not intrinsically required for the function of this species.

Figure 3

The overall architecture of the *Pf*MCM-N double hexamer is highly similar to other MCM double-hexamer structures. Three MCM double-hexamer structures were aligned based on the six OB-folds of one hexamer for the *Mt*MCM-N double hexamer (Fletcher *et al.*, 2003) (a), the *Pf*MCM-N double hexamer (b) and the full-length *Sc*Mcm2–7 double hexamer (Li *et al.*, 2015) (c).

Figure 4

The *PfMCM-N* double hexamer has less extensive inter-hexamer interactions than the *MtMCM-N* double hexamer. (a) Magnified view of Fig. 3(a) showing that the two hexamers of *MtMCM-N* (Fletcher *et al.*, 2003) interact at the inter-hexamer interface through hydrogen bonds mediated by Arg161 (shown in magenta stick representation). Interactions across the hexamers are shown as dashed red lines, and interactions within the same hexamer are shown as dashed blue lines. (b) Magnified view of Fig. 3(b) showing that the *PfMCM-N* hexamers do not interact *via* hydrogen bonds across the inter-hexamer interface. (c) A close-up view of the inter-hexamer interface of *MtMCM-N* (Fletcher *et al.*, 2003) from inside the channel, highlighting the hydrogen-bonding network between subunits, especially that involving Arg161 (shown as magenta sticks). Interactions across the hexamers are shown as dashed red lines, and interactions within the same hexamer are shown as dashed blue lines. Subunits that do not interact with the green subunits are omitted for clarity. The distance between the two Zn atoms (gray) is 9.6 Å. (d) A close-up view of the inter-hexamer interface of *PfMCM-N* similarly oriented to that in (c). The hexamers do not interact *via* hydrogen bonds across the inter-hexamer interface. Two opposing subunits (yellow) and neighboring zinc-binding motifs are shown. Subunits that do not interact with the yellow subunits are omitted for clarity. The distance between the two Zn atoms (gray) is 7.5 Å. In all panels, O atoms are shown in red, N atoms in blue and hydrogen bonds (2.7–3.3 Å) as dashed lines.

Acknowledgements

This work was supported in part by ALSAC and grant R01GM098771 (to EJE) from NIGMS. We thank Dr Christopher D. Lima for the generous gifts of the SUMO vector and the Ulp1 protease plasmid (Mossesso & Lima, 2000). Data were collected on the Southeast Regional Collaborative Access Team (SER-CAT) 22-BM beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at <http://www.ser-cat.org/members.html>. We are grateful to the SER-CAT staff for experimental support. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38.

References

- Chong, J. P. J., Hayashi, M. K., Simon, M. N., Xu, R.-M. & Stillman, B. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 1530–1535.
- Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M. R. & Berger, J. M. (2011). *Nature Struct. Mol. Biol.* **18**, 471–477.
- Deegan, T. D. & Diffley, J. F. X. (2016). *Curr. Opin. Struct. Biol.* **37**, 145–151.
- Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B. & Speck, C. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 20240–20245.
- Fletcher, R. J., Bishop, B. E., Leon, R. P., Sclafani, R. A., Ogata, C. M. & Chen, X. S. (2003). *Nature Struct. Biol.* **10**, 160–167.
- Froelich, C. A., Kang, S., Epling, L. B., Bell, S. P. & Enemark, E. J. (2014). *Elife*, **3**, e01993.
- Fu, Y., Slaymaker, I. M., Wang, J., Wang, G. & Chen, X. S. (2014). *J. Mol. Biol.* **426**, 1512–1523.
- Ilves, I., Petojevic, T., Pesavento, J. J. & Botchan, M. R. (2010). *Mol. Cell*, **37**, 247–258.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Li, N., Zhai, Y., Zhang, Y., Li, W., Yang, M., Lei, J., Tye, B.-K. & Gao, N. (2015). *Nature (London)*, **524**, 186–191.
- Liu, W., Pucci, B., Rossi, M., Pisani, F. M. & Ladenstein, R. (2008). *Nucleic Acids Res.* **36**, 3235–3243.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Miller, J. M., Arachea, B. T., Epling, L. B. & Enemark, E. J. (2014). *Elife*, **3**, e03433.
- Miller, J. M. & Enemark, E. J. (2015). *Archaea*, **2015**, 305497.
- Mossesso, E. & Lima, C. D. (2000). *Mol. Cell*, **5**, 865–876.
- Moyer, S. E., Lewis, P. W. & Botchan, M. R. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 10236–10241.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D* **67**, 355–367.
- Murzin, A. G. (1993). *EMBO J.* **12**, 861–867.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Remus, D., Beuron, F., Tolun, G., Griffith, J. D., Morris, E. P. & Diffley, J. F. X. (2009). *Cell*, **139**, 719–730.