

Structural insights into DNA degradation by human mitochondrial nuclease MGME1

Chun Yang¹, Ruiqi Wu¹, Hehua Liu^{1,2}, Yiqing Chen¹, Yanqing Gao¹, Xi Chen¹, Yangyang Li¹, Jinbiao Ma², Jixi Li^{1,3,*} and Jianhua Gan^{1,*}

¹State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, Department of Physiology and Biophysics, School of Life Sciences, Fudan University, Shanghai 200433, China,

²State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, Department of Biochemistry, School of Life Sciences, Fudan University, Shanghai 200433, China and ³Department of Neurology, Huashan Hospital, Fudan University, Shanghai 200040, China

Received June 23, 2018; Revised September 06, 2018; Editorial Decision September 10, 2018; Accepted September 11, 2018

ABSTRACT

Mitochondrial nucleases play important roles in accurate maintenance and correct metabolism of mtDNA, the own genetic materials of mitochondria that are passed exclusively from mother to child. MGME1 is a highly conserved DNase that was discovered recently. Mutations in MGME1-coding gene lead to severe mitochondrial syndromes characterized by external ophthalmoplegia, emaciation, and respiratory failure in humans. Unlike many other nucleases that are distributed in multiple cellular organelles, human MGME1 is a mitochondria-specific nuclease; therefore, it can serve as an ideal target for treating related syndromes. Here, we report one HsMGME1-Mn²⁺ complex and three different HsMGME1-DNA complex structures. In combination with *in vitro* cleavage assays, our structures reveal the detailed molecular basis for substrate DNA binding and/or unwinding by HsMGME1. Besides the conserved two-cation-assisted catalytic mechanism, structural analysis of HsMGME1 and comparison with homologous proteins also clarified substrate binding and cleavage directionalities of the DNA double-strand break repair complexes RecBCD and AddAB.

INTRODUCTION

MGME1 (also known as DDK1) is a highly conserved, mitochondria-specific DNase (1). Like the mitochondrial nuclease DNA2, MGME1 belongs to the PD-(D/E)XK nuclease superfamily. However, in contrast to DNA2 (2–5) and many other mitochondrial DNases (6–13), MGME1 can cleave various types of DNA substrates with a preference

for single-stranded DNAs (ssDNA) (14–16). *In vitro* studies have shown that MGME1 can cleave ssDNAs in both 5'→3' and 3'→5' directions. As revealed by loss of function mutations (1), MGME1-deficient cells showed apparent accumulation of 7S DNA, an ssDNA that synthesized by DNA POL γ . 7S DNA can form displacement loop (D-loop) structure with the non-coding region (NCR) of mtDNAs (17,18). MGME1 has no ribonuclease activity; when a chimeric DNA-RNA substrate is provided, MGME1 will slide over the RNA fragment and cleave the DNA only. In addition to ssDNA and DNA-RNA chimeric oligonucleotides, MGME1 can also cleave double-stranded DNAs (dsDNAs) and DNAs with a 5'-flap or 3'-flap (16).

HsMGME1 does not interact with the commonly known factors involved in mtDNA transactions (such as the DNA-interacting proteins TFAM and mtSSB of the mitochondrial nucleoid), but recent studies (14,19,20) suggested that HsMGME1 could form a complex with PolgA, a catalytic subunit of the mtDNA polymerase. The interaction with PolgA suggests that MGME1 is a potential constituent of the mitochondrial replisome. Recent studies also suggested that HsMGME1, PolgA and TWINKLE (a mitochondrial DNA helicase) can work together in the double-strand degradation process of mtDNA (21,22). Like many known mitochondrial nucleases, the cleavage activity of MGME1 is important for maintenance of the mitochondrial genome. However, while other mitochondrial nucleases are mainly involved in the repair process of mtDNA, MGME1 participates in both replication and double-strand degradation processes of mtDNA. MGME1 is encoded by the gene *C20orf72* in humans. Loss-of-function mutations in the MGME1-coding gene can lead to severe multisystem mitochondrial disorders in humans with depletion and rearrangements of mtDNA (16). Patients with loss-of-function MGME1 mutations could develop progressive external ophthalmoplegia, skeletal muscle wasting/weakness,

*To whom correspondence should be addressed. Tel: +86 21 51630543; Email: ganjhh@fudan.edu.cn
Correspondence may also be addressed to Jixi Li. Tel: +86 21 51630539; Email: lijixi@fudan.edu.cn

respiratory distress, emaciation, severe dilated cardiomyopathy, mental retardation, microcephalus, and severe gastrointestinal symptoms (21,22).

Based on the structures of other PD-(D/E)XK nucleases, several MGME1 models have been proposed previously (1,16); however, due to the lack of high-resolution structures, especially the structures of MGME1 in complex with DNA, detailed information underlying substrate binding and cleavage of MGME1 remains elusive. Herein, we report crystallographic and *in vitro* catalytic studies of MGME1. In total, four MGME1 structures were solved: one is in a complex with Mn²⁺ and the other three are complexed with DNA substrates, with or without Ca²⁺. Structural analysis and comparison suggest that MGME1 follows a two-cation-assisted catalytic mechanism of substrate cleavage. In addition to DNA binding, our structures also revealed one unique structural pin; the pin is composed of two aromatic residues (Trp152 and Phe173) and is important for DNA duplex unwinding and flap DNA cleavage of MGME1. MGME1 shares both sequence and structural similarities with the RecB-type nucleases; the comparison also elucidated substrate binding and the catalytic mechanism of RecB and AddA proteins, which are important components of the DNA double-strand break repair complexes RecBCD (23–25) and AddAB (26,27), respectively.

MATERIALS AND METHODS

Plasmid construction

Full-length wild-type (WT) *HsMGME1* was obtained by PCR using a human cDNA library as a template. The PCR products were cleaved with enzymes BamHI and XhoI and cloned into a pET28-Sumo vector. The recombinant vector was then transferred into *Escherichia coli* BL21 (DE3) competent cells for protein expression. All the truncated or mutated *HsMGME1* were constructed by PCR or overlap PCR using WT *HsMGME1* as a template. Primers used in plasmid construction are listed in Supplementary Table S1. Sequences of plasmids were confirmed by DNA sequencing. All recombinant strains were protected using 30% glycerol and stored in a –80°C freezer prior to use.

Protein expression and purification

All *HsMGME1* proteins were expressed overnight in BL21 (DE3) cells with lysogeny broth (LB) medium supplemented with 50 µg/ml kanamycin at 37°C. When OD₆₀₀ reached 0.6, expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.2 mM for 18 h at 18°C. The cells were harvested by centrifugation and the pellets were stored at –80°C.

For overproduction of the Se-Met substituted WT *HsMGME1*, the revived recombinant strains from 20 ml overnight cultures were inoculated into 2 l of LB medium supplemented with kanamycin (50 µg/ml) and grown at 37°C. Cells were harvested by centrifugation when OD₆₀₀ was 0.4, then the pellets were washed and resuspended in M9 medium (47.7 mM Na₂HPO₄, 22 mM KH₂PO₄, 8.6 mM NaCl, and 28.2 mM NH₄Cl). The resulting cells were transferred into 2 l fresh M9 medium supplemented with 50 µg/ml kanamycin and 40 mg/l Se-Met (J&K). After

growing at 37°C for 1 h, protein expression was induced by adding IPTG at a final concentration of 0.2 mM. The induced cultures were then grown at 18°C for an additional 18 h and the cells were harvested by centrifugation.

To improve solubility, the N-terminal MTS sequence (residues 1–20) was deleted in all WT and mutant proteins of *HsMGME1*. All proteins were purified using the same protocol. The pellets were first resuspended in Ni binding buffer (Buffer A, 20 mM Tris pH 8.0, 500 mM NaCl and 25 mM imidazole pH 8.0) and lysed under high pressure with a JN-02C cell crusher. The homogenate was clarified by centrifugation and the supernatant was loaded onto a HisTrap™ HP column (GE Healthcare) equilibrated with Buffer A. The His-Sumo-*HsMGME1* fusion proteins were eluted from the column using elution buffer (20 mM Tris pH 8.0, 500 mM NaCl and 500 mM imidazole pH 8.0) with a gradient. Fractions containing the desired fusion proteins were pooled and dialyzed against Buffer S (20 mM Tris pH 8.0, 500 mM NaCl, and 25 mM imidazole pH 8.0) at 4°C for 3 h; Ulp1 protease was also added to the sample during the dialysis process. The sample was again loaded onto the HisTrap™ HP column and the flow-through containing the target proteins was collected and mixed with Tris buffer (20 mM Tris pH 8.0) to dilute the NaCl to a final concentration of 200 mM. The diluted sample was applied to a HiTrap Q HP column equilibrated with Q binding buffer (20 mM Tris pH 8.0, 200 mM NaCl). The flow-through containing the target proteins were concentrated and loaded onto a Hi 16/60 Superdex G200 column (GE Healthcare) equilibrated with a gel filtration buffer (20 mM Tris pH 8.0, 200 mM NaCl and 1 mM DTT). Protein purity was analyzed using an SDS-PAGE gel and the samples were concentrated and snap-frozen using liquid nitrogen and stored at –80°C until use.

Crystallization and data collection

The initial crystallization conditions of all *HsMGME1* crystals were identified using a Gryphon crystallization robot system (Arts Robin Instrument) with commercial kits at 18°C. The sitting drop vapor diffusion method was used during the initial screening, whereas all crystallization optimization processes were done using the hanging drop vapor diffusion method. Protein concentrations were 10 mg/ml for all samples, and the CaCl₂ concentration was 30 mM for the H180Q-DNA2-Ca²⁺ complex. The protein: DNA molecular ratio was 1:1.2 for all *HsMGME1*-DNA complexes. Prior to crystallization, all samples were incubated on ice for 30 min. The crystallization conditions were 0.35 M di-sodium tartrate, 21% PEG 3350 and 0.1 M MnCl₂, 0.1 M phosphoric-citric acid pH 4.2 and 20% PEG300 and 0.1 M CHES pH 9.5 and 20% PEG 8000 for *HsMGME1*-Mn²⁺, *HsMGME1*-DNA2, and H180Q-DNA2-Ca²⁺ complexes, respectively. Both native and Se-Met substituted *HsMGME1*-DNA3 complex crystals were grown in 0.1 M HEPES pH 7.5, 0.8 M NaH₂PO₄, and 0.8 M K₂HPO₄. The drop contained equal volumes of sample and crystallization solution for all crystals.

All crystals were cryoprotected using their mother liquor supplemented with 25% glycerol and snap-frozen in liquid nitrogen. X-ray diffraction data were collected on beamlines

BL17U1 and BL19U1 at the Shanghai Synchrotron Radiation Facility (SSRF) at cryogenic temperatures and maintained with a cryogenic system. Data processing was carried out using HKL2000 or HKL3000 programs (28). The data collection and processing statistics are summarized in Supplementary Table S2.

Structure determination and refinement

The Se-*HsMGME1*-DNA3 structure was solved using the single-wavelength anomalous diffraction (SAD) method (29) with the AutoSol program (30) embedded in the Phenix suite (31). The initial model that covered approximately 70% of protein residues in the asymmetric unit was built with the Autobuild program. The model was then refined against diffraction data of the native *HsMGME1*-DNA3 crystal with the Refmac5 program (32) of CCP4i, which revealed the detailed orientations of the missing residues. During refinement, 5% of randomly selected data was set aside for free R-factor cross validation calculations. The $2F_o - F_c$ and $F_o - F_c$ electron density maps were regularly calculated and used as guides for building the missing amino acids, DNA and solvent molecules with COOT (33). The *HsMGME1*-Mn²⁺, *HsMGME1*-DNA2 and H180Q-DNA2-Ca²⁺ complex structures were all solved using the molecular replacement method with the Phaser program of CCP4i suite using the *HsMGME1*-DNA3 structure as the search model. The final refinement of all structures was done with the phenix.refine program (34) of Phenix. The structural refinement statistics are summarized in Supplementary Table S2.

In vitro cleavage assays

6-Carboxy-fluorescein (FAM)-labelled substrates (Supplementary Table S3) were purchased from Generay Biotech Ltd. (Shanghai, China). To obtain double-stranded substrates, self-complementary oligonucleotides were dissolved in reaction buffer (20 mM Tris pH 8.0, 100 mM NaCl) and heated at 95°C for 15 min followed by slow cooling to room temperature.

The standard enzymatic assay was performed in a 20 μ l mixture containing 0.8 μ M of DNA substrates, 0.1 mg/ml bovine serum albumin (BSA), 100 mM NaCl, 5 mM MgCl₂, 20 mM Tris-HCl (pH 8.0), 1 mM DTT, and different concentrations of WT or mutant *HsMGME1*. Reaction mixtures were incubated at 37°C for the indicated time and terminated by adding an equal volume of formamide loading dye (95% formamide, 20 mM ethylenediaminetetraacetic acid, 0.025% SDS, 0.02% bromophenol blue, and 0.02% xylene cyanol in 1 \times TBE buffer). The mixtures were incubated at 95°C for 10 min and analyzed on denaturing 18% polyacrylamide 8 M urea gels. The gel was imaged using Typhoon FLA 9000. The intensities of the bands were quantified by ImageQuantTL and the data was compared in GraphPad Prism.

Crosslinking assays

Four-microliter samples consisting of 1 μ l 5 \times reaction buffer (150 mM HEPES-NaOH pH 7.6, 150 mM NaCl, 5

mM DTT and 0.5 mM EDTA), 2 μ l *HsMGME1* protein (50 μ M), and 1 μ l DNA3 (10 μ M) were mixed and incubated at room temperature for 30 min. One-microliter diluted DSS (disuccinimidyl suberate) was added to the reaction system at final concentrations of 50, 100 and 200 μ M. The samples were incubated at room temperature for an additional 20 min and quenched by the addition of 1.25 μ l SDS-loading buffer. Samples were heated at 95°C for 5 min, centrifuged, and loaded onto 10% SDS-PAGE gels. Gels were run at 200 V for 60 min in 1 \times SDS running buffer and stained with Coomassie brilliant blue.

RESULTS

Overall structure of *HsMGME1*

Human MGME1 (*HsMGME1*) protein is composed of 344 amino acids and belongs to the PD-(D/E)XK nuclease superfamily. In addition to the C-terminal nuclease domain, *HsMGME1* has one predicted mitochondria-targeting sequence (MTS) at the N-terminus (Figure 1A). Interestingly, besides the three common PD-(D/E)XK motifs (I, II and III), *HsMGME1* shares two additional characteristic motifs (Ia and IV) with RecB-type nucleases, such as *BsAddA* and *EcRecB*. Motif Ia contains one highly conserved histidine residue, corresponding to His180 in *HsMGME1*; motif IV is composed of five residues QhXXY, where h represents a hydrophobic residue and X represents any residue. *HsMGME1* shares very high (up to 83%) sequence similarity with MGME1 orthologs from other species, such as mouse, chicken, and zebrafish, but has very low similarity with other PD-(D/E)XK nucleases.

MGME1 plays an important role in mitochondrial genome maintenance, however, no MGME1 structure was available. To better understand the structure and function of MGME1, we carried out crystallographic studies using *HsMGME1*. In total, four *HsMGME1* structures were solved, including one apo- structure and three DNA-bound complex structures (Supplementary Table S2). The apo-*HsMGME1* structure belongs to the $P4_32_12$ space group and contains one *HsMGME1* molecule per asymmetric unit (Figure 1B); the monomeric state of *HsMGME1* is consistent with our gel filtration profile and a previous report (1). As revealed by all four structures, *HsMGME1* contains five β -strands (β 1- β 5) and six α -helices (α 1- α 6). The two long helices α 2 and α 3 are formed by the residues 143-160 and 162-187, respectively. α 2 and α 3 are assembled in a helical arch; the angle between the two helices is \sim 70°. In the apo-*HsMGME1* structure, residues 1-97, 113-117 and 190-198 are disordered (Supplementary Figure S1A); the β 3- α 5 connecting loop (aa 255-264) resides right under the helical arch (Figure 1B and Supplementary Figure S1B). The side chain of Phe261 forms a weak hydrophobic stacking interaction with the side chains of Trp152 and Lys153 of α 2 (Supplementary Figure S1B); the backbone O atoms of Lys257 and Pro258 form water-mediated hydrogen bond (H-bond) interactions with Tyr168 and Thr169 of α 3 (Supplementary Figure S1B), respectively. These interactions may help stabilize the conformation of both the helical arch and the β 3- α 5 connecting loop.

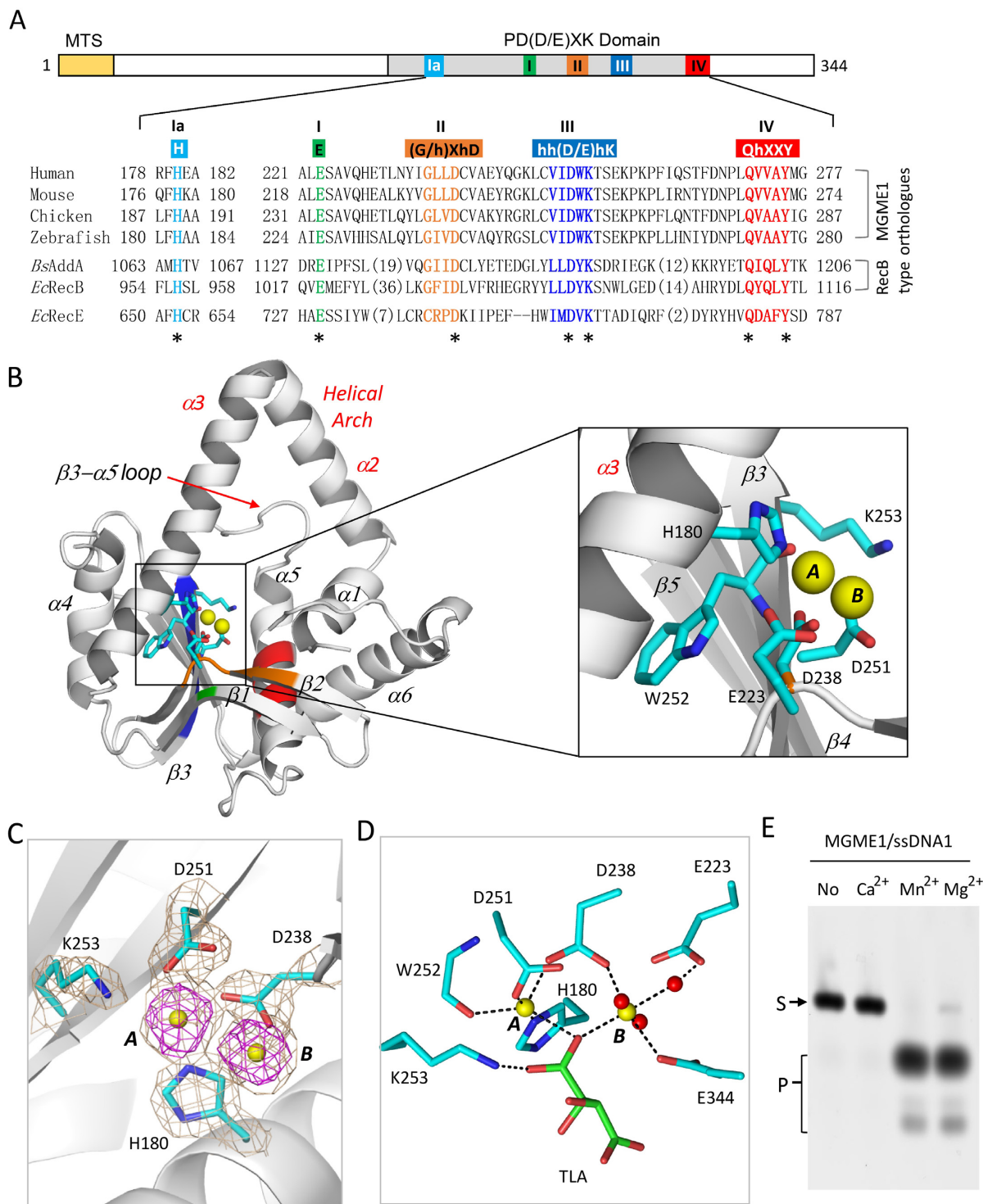


Figure 1. *HsMGME1* is a cation-dependent nuclease. (A) Domain architecture of *HsMGME1*. MTS indicates the mitochondrial targeting signal. Residues of the conserved motifs are highlighted in bold. (B) Overall structure of the *HsMGME1*-Mn²⁺ complex. *HsMGME1* is shown as cartoon in white. (C) $2F_o - F_c$ electron density of Mn²⁺ and surrounding residues. The wheaten and magenta maps are contoured at 1.5 and 5.0 σ level, respectively. (D) Detailed coordination of Mn²⁺. TLA stands for tartrate acid present in the crystallization condition. (E) *in vitro* cleavage assay showing the effect of cations. *HsMGME1*, ssDNA1, and cations (if present) concentrations are 0.1 μ M, 0.8 μ M and 5 mM, respectively. The conserved motifs Ia, I, II, III and IV are highlighted in identical colors in panels A and B. In (B–D), Mn²⁺ and coordinating water molecules are shown as yellow and red spheres, respectively. Active site residues are shown as sticks in atomic colors (C, cyan; N, blue; O, red).

Mn²⁺ coordination of *Hs*MGME1

A previous study (1) showed that the catalytic activity of *Hs*MGME1 was dependent on a divalent cation, commonly Mg²⁺. Though Mg²⁺ was not present in either the crystallization sample or the crystallization buffer, the crystallization buffer of the apo-*Hs*MGME1 crystals contained 100 mM Mn²⁺. As depicted in Figure 1C and D, two well-defined Mn²⁺ ions were captured at the active site of the apo-*Hs*MGME1 structure, which was hereafter referred to as *Hs*MGME1-Mn²⁺. The A-site Mn²⁺ was five-coordinated. In addition to the O1 atom of the tartrate acid (TLA, which was present in the crystallization buffer) and the main chain O atom of Trp252, it was also coordinated by the side chain NE2, OD2 and OD1 atoms of His180, Asp238 and Asp251, corresponding to the highly conserved His, Asp, and Asp residues of motifs Ia, II and III, respectively. The B-site Mn²⁺ was six-coordinated. In addition to three water molecules and the side chain OE2 atom of Glu344 (the last residue of *Hs*MGME1), the B-site Mn²⁺ was also coordinated by TLA and Asp238, via their O1 and OD1 atoms, respectively. The average coordinating distances of the A-site and B-site Mn²⁺ ions were all ~2.15 Å.

As confirmed by both structures and catalytic assays, many Mg²⁺-dependent nucleases, including RNase III (35) and RNase H (36,37), are also active in the presence of Mn²⁺. To test whether Mn²⁺ can also support the cleavage activity of *Hs*MGME1, we performed an *in vitro* catalytic assay using wild type (WT) *Hs*MGME1 and one short 12-nt ssDNA (ssDNA1, 5'-TTCTTCTTCTTC-3', which was FAM-labelled at the 5'-end). As depicted in Figure 1E, *Hs*MGME1 is inactive in the absence of divalent cations, but it can efficiently cleave DNA1 in the presence of either Mg²⁺ or Mn²⁺, indicating that Mn²⁺ can mimic Mg²⁺ in substrate binding and cleavage of *Hs*MGME1. During the *in vitro* ssDNA cleavage assays (Supplementary Figure S2), we noticed that *Hs*MGME1 had some weak cleavage site preference; altering the compositions of the ssDNAs will lead to corresponding changes in product patterns.

ssDNA recognition of *Hs*MGME1

Among all substrates, *Hs*MGME1 has a strong preference for ssDNAs. Details of ssDNA recognition were revealed by the *Hs*MGME1-ssDNA2 complex (Figure 2A), which was determined at high resolution (1.75 Å). ssDNA2 (5'-AACAACAACAAC-3') is 12-nt in length and resides under the helical arch in the structure. As revealed by the structural superposition (Supplementary Figure S1C), binding of ssDNA2 has no obvious impact on the conformation of the helical arch, but it causes significant conformational changes to the β3-α5 connecting loop.

Due to structural disorder, the 5'-end A1A2 and 3'-end A11C12 were not observed, whereas the central seven nucleotides (from C3 to A9) and the phosphate group of A10 of ssDNA2 are all well defined in the *Hs*MGME1-ssDNA2 structure (Figure 2B). C3, A4, A5 and C6 of ssDNA2 are arranged in an A-form like conformation. C3 mainly forms a stacking interaction with A4. In addition to the stacking interactions between their nucleobases, A4, A5 and C6

also form extensive H-bond interactions with *Hs*MGME1. The phosphate group of A4 (Figure 2C) forms two water-mediated H-bonds with the main chain N and side chain OG atoms of Ser132. A5 forms four direct H-bonds (Figure 2C). The O1P atom of A5 forms one H-bond (2.9 Å) with the main chain N atom of Leu236 and one H-bond (2.6 Å) with the side chain OH atom of Tyr275. Via its main chain N atom and side chain OG1 atom, Thr134 forms two H-bonds with the O2P atom of A5; the average H-bond distance is 2.8 Å. Like A5, the phosphate group of C6 (Figure 2D) also forms four direct H-bonds with *Hs*MGME1, via the NE2 atom of His180, the OD1 atom of Asp238, the NZ atom of Lys253, and the NE2 atom of Gln271.

The backbone of ssDNA2 is severely bent in-between C6 and A7, resulting in an almost perpendicular orientation between the nucleobases of C6 and A7 (Figure 2B, right panel). The nucleobases of A7 and A8 stack with each other. Via their phosphate groups, A7 and A8 form extensive H-bond interactions with *Hs*MGME1 (Figure 2E). The O1P atom of A7 forms two H-bonds (2.8 and 2.7 Å) with Thr254, via the main chain N atom and side chain OG1 atom, respectively; the O2P atom of A7 forms one H-bond (2.8 Å) with the side chain NZ atom of Lys253. In addition to the direct H-bond (2.8 Å) between the O1P atom of A8 and the main chain N atom of Glu256, A8 also forms two water-mediated H-bonds with the main chain O atom of Thr254 and the side chain NZ atom of Lys259, respectively.

Interestingly, the backbone of ssDNA2 is also bent in-between A8 and C9 and the nucleobases of A8 and C9 are almost perpendicular to each other (Figure 2B and F). The side chain of Trp152 forms several hydrophobic interactions with the sugar pucker of A8 and the backbone C5* atom of C9. The nucleobase of C9 is sandwiched in between the side chains of Ile149 and Phe266, forming hydrophobic stacking interactions with them, especially Phe266. Both Ile149 and Trp152 reside at the middle of helix α2. The conformation of C9 was further stabilized by one H-bond (2.8 Å) between its O1P atom and the NZ atom of Lys257.

Active site comparison of *Hs*MGME1

As confirmed by our *in vitro* catalytic assay (Figure 1E), *Hs*MGME1 has no cleavage activity in the presence of Ca²⁺, a common inhibitor of Mg²⁺-dependent nucleases. Ca²⁺ has been extensively utilized in the structural and mechanistic studies of various nucleases (38). Surprisingly, though 10 mM CaCl₂ was included in the crystallization sample, no Ca²⁺ was bound at the active site of the *Hs*MGME1-ssDNA2 complex. Unlike many nucleases, which coordinate Mg²⁺ (or Mn²⁺) through the side chains of three or more negatively charged residues (Asp or Glu), *Hs*MGME1 coordinates Mn²⁺ through the side chains of His180, Asp238, and Asp251, and the main chain of Trp252 (Figure 1D). The NE2 atom of His180 forms one strong H-bond (2.7 Å) with the backbone O atom of Trp252 in the *Hs*MGME1-ssDNA2 structure; compared to the *Hs*MGME1-Mn²⁺ structure, the Trp252-Lys253 amide plane was tilted about 20° in the *Hs*MGME1-ssDNA2 structure (Supplementary Figure S1D).

We speculated that the direct His180-Trp252 interaction disrupts Ca²⁺ binding and leads to the associated con-

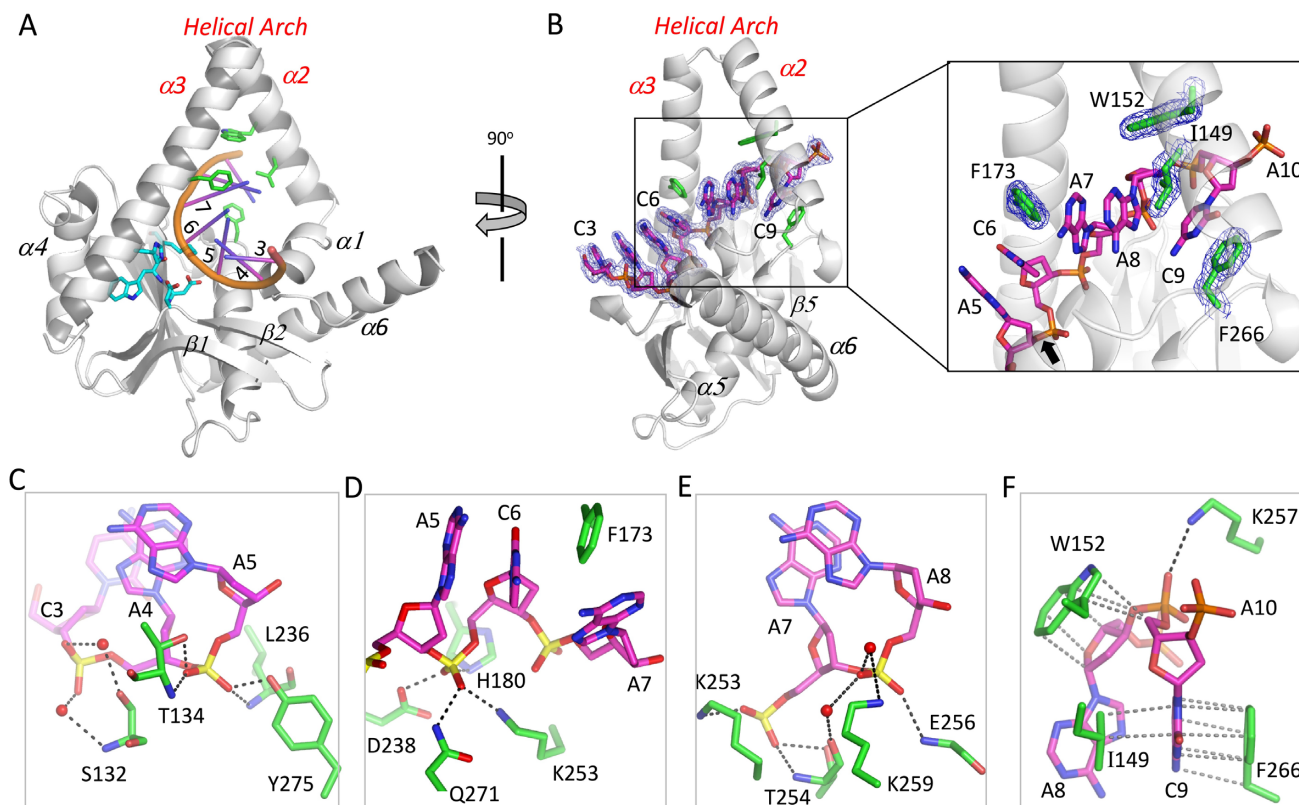


Figure 2. Single-stranded DNA recognition by *HsMGME1*. (A) and (B) Overall fold of the *HsMGME1*-ssDNA2 complex. Active site residues are shown as sticks in atomic colors (C, cyan; N, blue; O, red) in (A). In (B), $2F_o - F_c$ electron density maps are contoured at 1.5 σ level. Bending of the DNA backbones are shown in the right and the O-P bond that will be cleaved by *HsMGME1* is indicated by black arrow. (C–F) Detailed interactions between *HsMGME1* and the individual nucleotide of ssDNA2. H-bonds and hydrophobic interactions are indicated by black and gray dashed lines, respectively. Water molecules are shown as red spheres. ssDNA2 and ssDNA2-interacting residues are shown as sticks with their C-atoms colored in magenta and green in (C–F).

formational changes of Asp251 and Lys253 in the structures (Supplementary Figure S1D). To test our hypothesis, we constructed one *HsMGME1* mutant with His180 substituted by Gln180, which is similar to His180 in size but neutral in charge. Using this H180Q mutant, we successfully solved one *HsMGME1* ternary complex structure, H180Q-ssDNA2-Ca²⁺. As depicted in Figure 3A, one six-coordinated Ca²⁺ ion was bound at the active site of *HsMGME1*. The Ca²⁺ ion coordinates with the O atom of acetic acid (ACT), the side chain OD2 atom of Asp238, the side chain OD1 atom of Asp251, the main chain O atom of Trp252, as well as two water molecules. Beside Ca²⁺-coordination, water molecule W1 also forms one H-bond (2.6 Å) with the side chain NZ atom of Lys253.

The residue Gln180 does not coordinate Ca²⁺; compared with His180 in the *HsMGME1*-Mn²⁺ structure, the side chain of Gln180 is rotated clockwise around the CA-CB bond for about 150° in the H180Q-ssDNA2-Ca²⁺ complex (Figure 3B). Though they are not identical in coordination, the orientations of the Ca²⁺ and the A-site Mn²⁺ are very similar in the H180Q-ssDNA2-Ca²⁺ and *HsMGME1*-Mn²⁺ structures. Interestingly, the orientations of the Ca²⁺-coordinating water molecule W1 and the O11 atom of TLA are also very similar in the two structures; however, compared to *HsMGME1*-Mn²⁺, the NZ atom of Lys253 was shifted about 0.6 Å in the H180Q-ssDNA2-Ca²⁺ complex.

Binding of Ca²⁺ does not affect the folding of *HsMGME1*. As indicated by the very low root mean square deviation value (rmsd, 0.7 Å), the overall structure of H180Q-ssDNA2-Ca²⁺ is very similar to that of the *HsMGME1*-ssDNA2 complex (Supplementary Figure S1E). However, detailed conformations of the two structures are not identical at their active sites (Figure 3C). Instead of forming an H-bond interaction, the NZ atom of Lys253 in the *HsMGME1*-ssDNA2 structure almost overlaps with the Ca²⁺-coordinating water molecule W1 of the H180Q-ssDNA2-Ca²⁺ complex. Though one Ca²⁺ ion was bound at the A-binding site, the B-binding site was not occupied by any cations in the H180Q-ssDNA2-Ca²⁺ complex, which may have weakened the overall DNA coordinating ability of Ca²⁺ and led to the shifting of the A5 nucleotide.

Catalytic mechanism and *in vitro* DNA degradation by *HsMGME1*

The cleavage activity of *HsMGME1* is cation-dependent; though we did extensive screening, no structure with cations coordinated with both protein and DNA was obtained. To elucidate the catalytic mechanism of *HsMGME1*, we further analyzed our structures and produced one plausible model (Figure 3D), in which the Mn²⁺ ions and

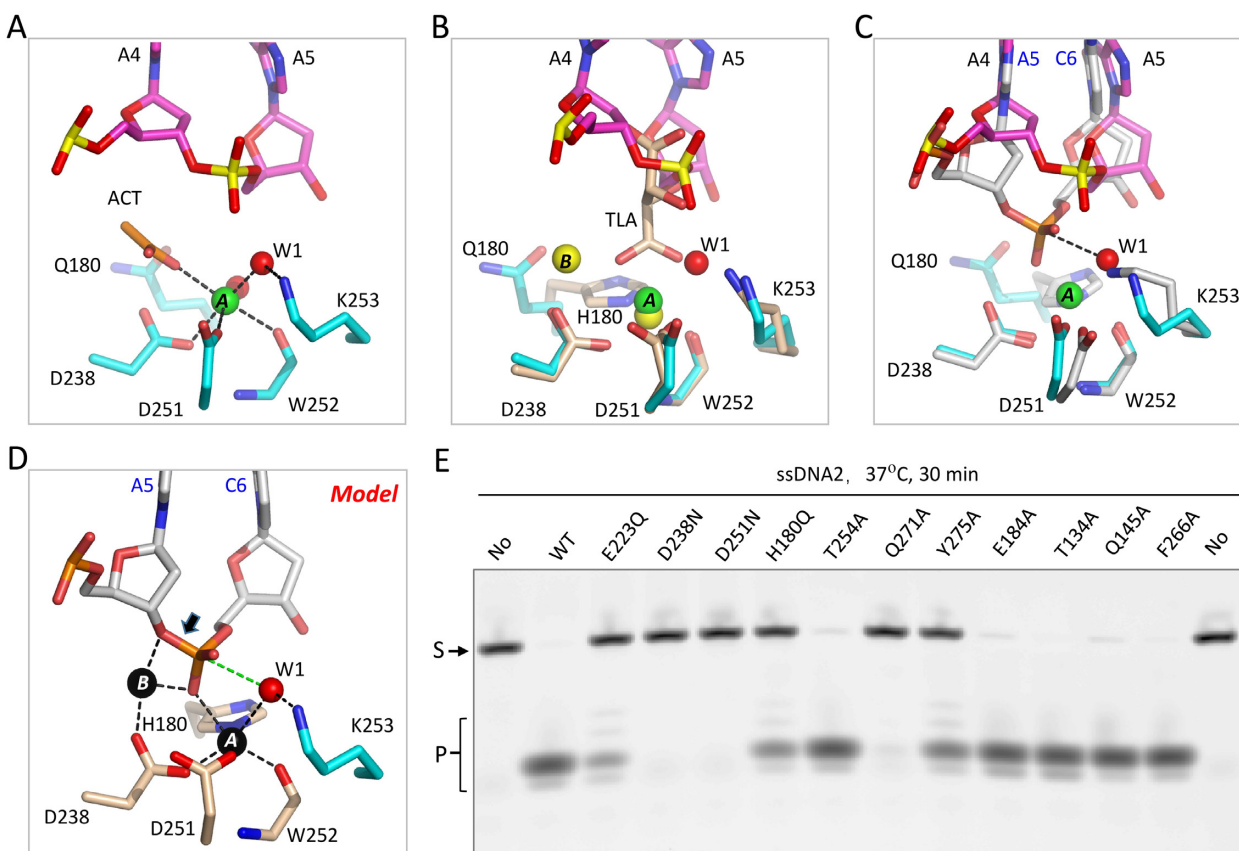


Figure 3. Catalytic site assembly of *HsMGME1*. (A) Catalytic site conformation of H180Q-ssDNA2-Ca²⁺. ACT stands for acetic acid present in the crystallization condition. Ca²⁺ and Ca²⁺-coordinating water molecules are shown as green and red spheres. DNA and catalytic site residues are shown as sticks with C-atoms colored in magenta and cyan, respectively. (B) Superposition of the catalytic site structures of H180Q-ssDNA2-Ca²⁺ and *HsMGME1*-Mn²⁺. In *HsMGME1*-Mn²⁺, Mn²⁺ ions are shown as yellow spheres, TLA and Mn²⁺-coordinating residues are shown as sticks with C-atoms colored in wheaten. (C) Superposition of the catalytic site structures of H180Q-ssDNA2-Ca²⁺ and *HsMGME1*-ssDNA2. C-atoms are colored in white in the *HsMGME1*-ssDNA2 complex. (D) Proposed model of the catalytic form *HsMGME1*, DNA, and cation ternary complex. The O-P bond that will be cleaved by *HsMGME1* is indicated by black arrow. (E) *in vitro* cleavage assay using ssDNA2 and WT or mutated *HsMGME1*. Substrates and products are labelled as S and P, respectively.

Mn²⁺-coordinating residues (His180, Asp238, Asp251 and Trp252), water molecule W1 and Lys253, and DNA positions were directly taken from the *HsMGME1*-Mn²⁺, H180Q-ssDNA2-Ca²⁺ and *HsMGME1*-ssDNA2 complexes, respectively. In the model, both A-site and B-site Mn²⁺ ions coordinate with the O1P atom of the cleavage site nucleotide. The A-site Mn²⁺ ion also coordinates with the water molecule W1, which is in-line with the phosphorus atom of cleavage site nucleotide and O3* atom of the leaving nucleotide; the distance between W1 and the phosphorus atom of the cleavage site nucleotide is 3.0 Å. In addition to O1P of C6, the B-site Mn²⁺ ion also coordinates with the O3* atom of the leaving nucleotide.

The complex structures of many divalent cation (Mg²⁺ or Mn²⁺)-dependent nucleases, such as C3PO (38), RNase III (39), and RNase H (36), have been previously reported. Though the overall structures are very different, the *HsMGME1* model and these nucleases share remarkable similarity in cation-coordination, suggesting that *HsMGME1* also follows a two-cation-assisted mechanism in substrate degradation. Via deprotonation, the A-site cation will activate the water molecule W1, which will at-

tack the phosphorus atom of the cleavage site nucleotide. In addition to assembly of the catalytic form complex, the B-site cation can also facilitate O-P bond breakage by neutralizing the negative charge developing on the leaving oxygen anion. Lys253 forms one H-bond and orients the nucleophilic water molecule W1 for in-line attack. As revealed by previous studies (38,39), mutation of Lys253 will significantly lower the catalytic activity of *HsMGME1*. Like *HsMGME1* and all other PD-(D/E)XK superfamily members, the active sites of *Saccharomyces cerevisiae* Rnt1P (40), *Kluyveromyces polysporus* DCL1 (41), and many other nucleases also contain one Lys residue, which may use a similar mechanism in regulating the activities of the enzymes.

To verify the catalytic mechanism of *HsMGME1*, we constructed several mutants with the cation-coordinating residues mutated, including H180Q, D238N, and D251N, and performed *in vitro* cleavage assays using 5' FAM-labeled ssDNA2. As depicted in Figure 3E, ssDNA2 cleavage activity of the H180Q mutant was significantly reduced compared with the WT *HsMGME1*. ssDNA2. More dramatic reduction was observed for the D238N and D251N mutants, suggesting that Asp238 and Asp251 are critical for

the catalytic activity of *HsMGME1*. Though it does not directly coordinate Mn^{2+} , the invariant Glu223 residue of motif I forms one H-bond (2.7 Å) with one water coordinating with the B-site Mn^{2+} (Figure 1D). Replacing Glu223 with Gln223 (for E223Q mutant) significantly lowered the catalytic ability of the protein; the ssDNA2 cleavage efficiency of E223Q mutant is comparable to that of H180Q mutant. Together, these observations suggest that His180 and Glu223 mainly affect the coordination of the cations, whereas Asp238 and Asp251 are involved in both cation-coordination and catalytic process.

In addition to the cation-coordinating residues, we also mutated many residues involved in DNA-binding. As depicted in Figure 3E, replacing Thr134 or Phe266 with an Ala residue has no obvious impact on the protein's ssDNA2 cleavage activity. However, compared to the WT protein, the ssDNA2 cleavage activities of Q145A, E184A, and T254A mutants are all slightly weakened. More dramatic reduction was observed for the Q271A and Y275A mutants. Tyr275 interacts with the O1P atom of the nucleotide before the cleavage site. Though it interacts with the O2P atom in the *HsMGME1*–ssDNA2 complex, Gln271 interacts with the O1P atom of the cleavage site nucleotide in the H180Q–ssDNA2– Ca^{2+} complex, which may facilitate catalytic complex assembly and contribute to its strong effect on the cleavage activity of *HsMGME1*.

Unique structural pin for duplex unwinding

HsMGME1 can cleave various types of DNAs. In addition to ssDNA2, we also successfully solved one structure of *HsMGME1* in complex with DNA3 (5'-GGATCCTTCTTCTTCTTC-3'). Via the GGATCC located at the 5'-end, DNA3 can form a 6-bp duplex with a 12-nt overhang on the 3'-ends of both strands. The *HsMGME1*–DNA3 crystal belongs to the C2 space group and contains four *HsMGME1* molecules per asymmetric unit (Figure 4A). Three of the *HsMGME1* molecules only recognize the 3'-overhang region, whereas the last one interacts with both duplex and overhang regions of DNA3. As indicated by the low rmsd values (~ 0.5 Å), the overall folds of the four *HsMGME1* molecules are very similar. Structural superposition revealed that folding of these proteins are also very similar to the one in the *HsMGME1*–DNA2 complex (Supplementary Figure S3A).

DNA duplexes normally adopt a B-form conformation, however, unlike the regular DNA duplexes, the DNA3 duplex is severely distorted. As depicted in Figure 4B, the nucleobase of G2' is almost perpendicular to that of A3'. Though the flanking G1' and A3' still form Watson–Crick-like pairing with C6 and T4 of the complementary strand, G2' does not form regular Watson–Crick pairing with C5. The nucleobases of G2' and C5 only forms one H-bond, which is between their O6 and N3 atoms, respectively. Phe173 is located in the middle of the $\alpha 3$ helix and its side chain stacks against the A3'–T4 pair of DNA3 (Figure 4B). Unlike a regular DNA duplex, the nucleobases of both A3 and T4 are slightly tilted toward the T4'–A3 base pair. Instead of stacking, the side chain of Phe173 directly points toward G2' and C5, which leads to almost 90° rotation for both nucleotides. G1' and C6 pair with each other; however,

compared to a regular Watson–Crick G–C pair, the average H-bond distance (3.15 Å) between G1' and C6 is ~ 0.3 Å longer. The weaker G1'–C6 pair and the irregular G2'–C5 pair all indicate that DNA3 is unwound to some extent.

To investigate the functional role of Phe173, we constructed one F173A mutant of *HsMGME1* and carried out an *in vitro* cleavage assay. As depicted in Figure 4C, replacing Phe173 with Ala173 clearly weakened the protein's degradation activity toward all tested duplex-containing DNAs (Supplementary Table S3), including double-stranded DNA4 (dsDNA4), 3'-overhang DNA4, and 5'-overhang DNA4. These observations indicated that Phe173 can facilitate the unwinding and degradation of duplex-containing DNAs by *HsMGME1*. Interestingly, mutation of Phe173 has no strong impact on single-stranded DNA (such as ssDNA1) cleavage by *HsMGME1*. Like the WT protein, F173A mutant can efficiently cleave ssDNA1; at low protein concentration (0.05 μ mol), the ssDNA1 cleavage activity of F173A is even slightly higher than that of WT *HsMGME1*.

Interestingly, though it is important for duplex-containing DNA degradation, Phe173 adopts a similar conformation in the *HsMGME1*–DNA3 and *HsMGME1*–ssDNA2 complexes (Supplementary Figure S3B). Like Phe173, Trp152 also adopts a similar conformation in the two structures (Supplementary Figure S3C). Trp152 is located at the single strand and duplex junction region of DNA3; its side chain is sandwiched by the sugar pucker of C6 on one side and by the side chains of Met156 and Phe165 on the other side. The strong hydrophobic interactions (as short as 3.55 Å) between Trp152 and the two hydrophobic residues fixed the conformation against the ribose. Compared to A8 of the *HsMGME1*–ssDNA2 complex, the nucleobase of C6 is slightly tilted toward Trp152 in *HsMGME1*–DNA3. Due to its close contact with Trp152, C6 could not undergo further movement, which may lead to the weak C6–G1' pairing. Like WT and F173A mutant of *HsMGME1*, the *in vitro* cleavage assays (Figure 4C) showed that the W152A mutant of *HsMGME1* can efficiently cleave ssDNA1; however, compared to Phe173, replacing Trp152 with an Ala residue caused more dramatic reduction on the protein's cleavage activities toward duplex-containing DNAs (especially dsDNA4 and 5'-overhang DNA4), suggesting that Trp152 is also important for the function of *HsMGME1*.

Functional characterization of the N- and C-terminal regions of *HsMGME1*

None of the crystals contain the N-terminal residues. Around 90 and 130 residues located at the N-terminus of *HsMGME1* were disordered in the complex structures of *HsMGME1*–DNA3 (Figure 4A) and H180Q–ssDNA2– Ca^{2+} (Figure 5A), representing the structures with longest and shortest visible sequences, respectively. Based on these two complex structures, we constructed two truncated *HsMGME1* proteins, WT-delN90 and WT-delN130, in which the N-terminal 90 and 130 residues, respectively, were deleted. As revealed by the *in vitro* cleavage assays (Figure 5B), deletion of the 90 N-terminal residues had no impact on ssDNA1 or 3'-flap DNA4 cleavage activities of

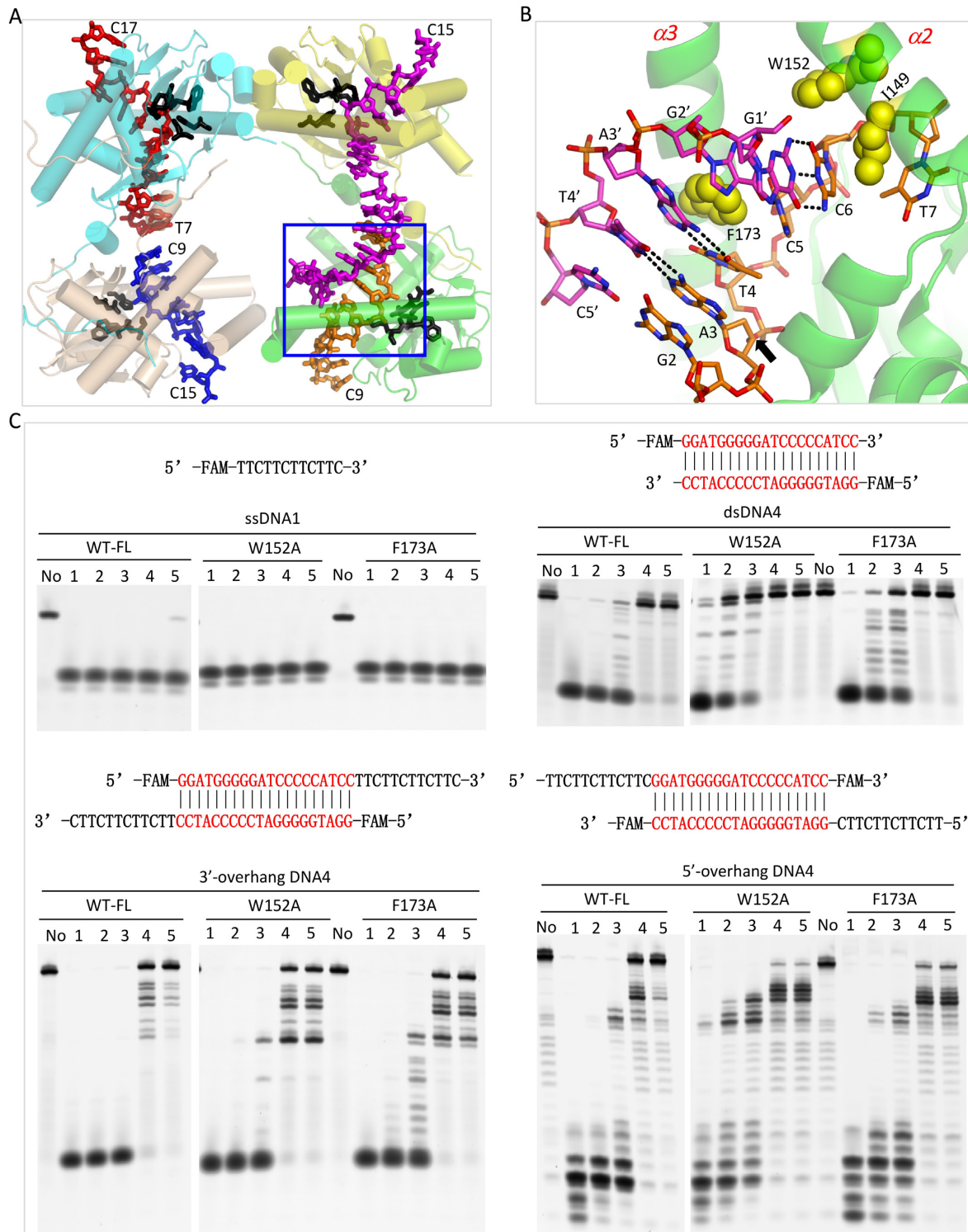


Figure 4. Functional characterization of the pin motif of *HsMGME1*. (A) Overall structure of the *HsMGME1*-DNA3 complex. *HsMGME1* and DNA molecules are shown as cartoon and sticks, respectively. The active site residues of all the four *HsMGME1* molecules are shown as black sticks. DNA duplex region is indicated by blue box. (B) Close-up view showing the detailed conformation of the duplex and pin residues, which are shown as sticks and spheres, respectively. The O-P bond that will be cleaved by *HsMGME1* is indicated by black arrow. (C) *in vitro* DNA cleavage assay using WT and pin region-mutated *HsMGME1*. ssDNA1 concentration is 0.8 μ M. WT or mutant *HsMGME1* concentrations are 0.8, 0.4, 0.2, 0.1 and 0.05 μ M in lanes 1–5, respectively. No protein is present in the lane labelled with No.

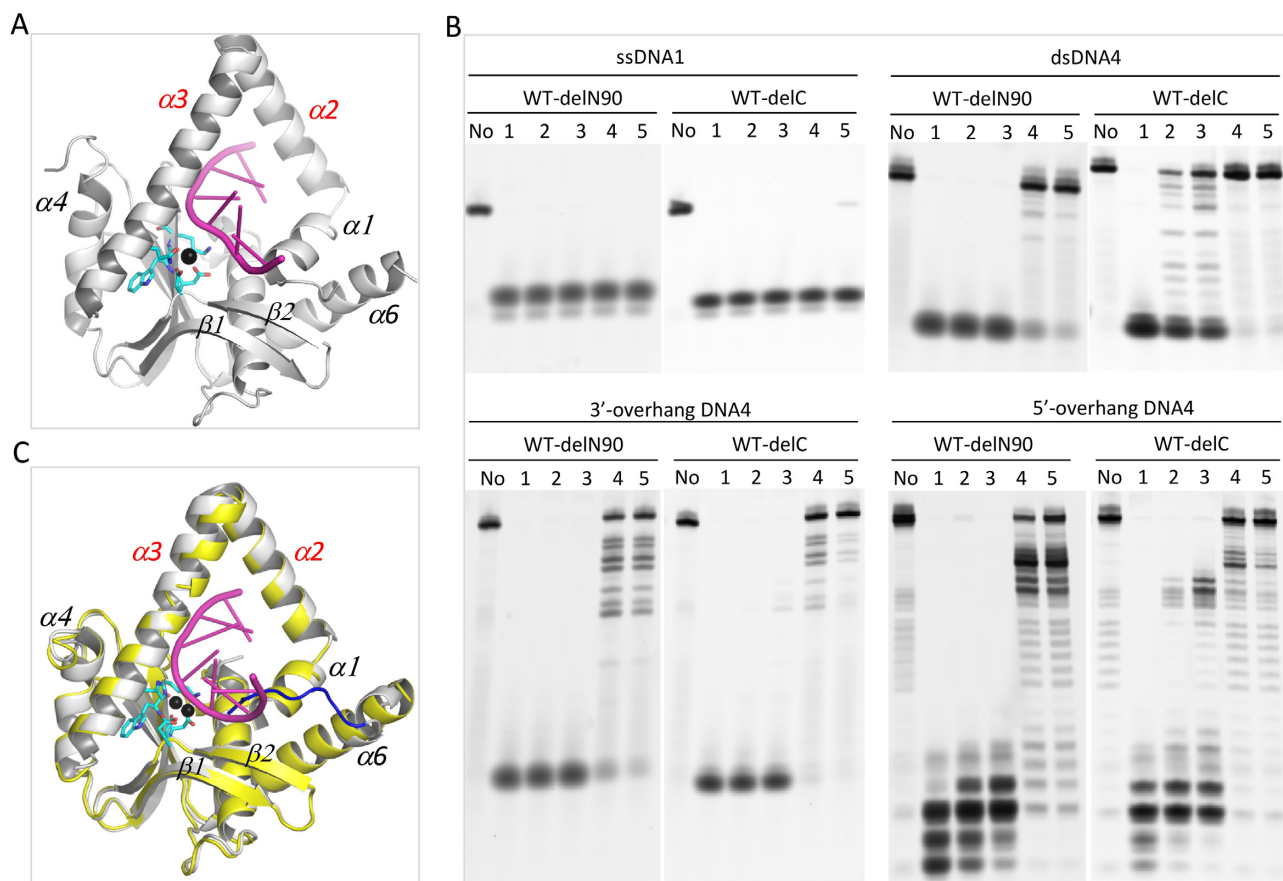


Figure 5. Characterization of the N- and C-terminus of *HsMGME1*. (A) Overall structure of the H180Q-ssDNA2-Ca²⁺ complex. H180Q and Ca²⁺ are shown as cartoon and black sphere, respectively. DNA is shown as cartoon in magenta. (B) *in vitro* DNA cleavage assay using N- or C-terminus truncated *HsMGME1*. (C) Superposition of the *HsMGME1*-Mn²⁺ and *HsMGME1*-ssDNA2 complexes, showing the conformational change of the C-terminal region. DNA and protein are colored in magenta and yellow in the *HsMGME1*-ssDNA2 complex. In the *HsMGME1*-Mn²⁺ complex, Mn²⁺ ions are shown as black spheres, *HsMGME1* is shown as cartoon in white, but the C-terminal region is colored in blue. DNA substrate concentrations are 0.8 μM. Truncated *HsMGME1* concentrations are 0.8, 0.4, 0.2, 0.1 and 0.05 μM in lanes 1–5, respectively. No protein is present in the lane labelled with No.

HsMGME1. Interestingly, compared to WT *HsMGME1*, WT-delN90 showed slightly higher cleavage activities toward dsDNA4 and 5'-flap DNA4 (Figure 5B).

Unlike WT-delN90 and WT *HsMGME1*, WT-delN130 is very unstable. As observed in the *HsMGME1*-ssDNA2 and *HsMGME1*-DNA3 complex structures, the 90–130 region is very flexible and can switch between two neighboring *HsMGME1* molecules. There are four *HsMGME1* molecules in the asymmetric unit of *HsMGME1*-DNA3 complex; arrangement of the four *HsMGME1* molecules (Figure 4A) is somewhat similar to the tetrameric structure of *EcRecE* (42), a nuclease that shares certain sequence similarity with *HsMGME1* (Figure 1A) and is responsible for dsDNA free end cleavage during double-stranded DNA break repair process. However, as revealed by our *in vitro* crosslinking assay using WT *HsMGME1* (Supplementary Figure S4), no clear bands corresponding to either *HsMGME1* dimer or tetramer were generated at lower concentrations (0–250 μM) of DSS (disuccinimidyl suberate). At higher concentrations (500–1000 μM) of DSS, some *HsMGME1* dimer and tetramer bands were observed, but the total percentage was very low. These observations suggest that *HsMGME1* mainly functions as a monomer and

the 90–130 region plays an important role in its stabilization.

In the *HsMGME1*-Mn²⁺ structure, the C-terminal residues 337–344 of *HsMGME1* are well ordered and fold directly back into the active center. Structural superposition (Figure 5C) showed that the C-terminal residues are partially overlapped with the DNA substrate; to avoid their potential conflict with DNA, the C-terminal region is shifted away from the active site and completely disordered in *HsMGME1*-ssDNA2 and all other *HsMGME1*-DNA complex structures. In *HsMGME1*-Mn²⁺, the most C-terminal residue Glu344 is involved in A-site Mn²⁺ coordination (Figure 1D); however, our *in vitro* cleavage assay (Figure 5B) showed that deletion of the C-terminal region (for WT-delC) has no impact on ssDNA1 or 3'-flap DNA4 cleavage, suggesting that Glu344-Mn²⁺ coordination is not critical for the protein's cleavage activity. Compared to the full-length *HsMGME1*, WT-delC has relatively weaker dsDNA4 or 5'-flap DNA4 cleavage activity, which is in contrast to WT-delN90. The detailed mechanism underlying the inhibitory or enhancing effects of the N- or C-terminal region of *HsMGME1* requires further investigation.

DISCUSSION

HsMGME1 is a recently discovered mitochondrial nuclease. Unlike many other mitochondrial nucleases, *HsMGME1* has very broad substrate tolerance with a preference for ssDNA. The detailed ssDNA binding mode of *HsMGME1* was revealed by our *HsMGME1*-DNA2 (Figure 2A) and H180Q-ssDNA2-Ca²⁺ (Figure 5A) complexes. As indicated by both structures, *HsMGME1* is an endonuclease that cleaves ssDNA in the middle. Both our *in vitro* assay (Supplementary Figure S4) and a previous study (1) showed that *HsMGME1* can cleave 5' P³²-labelled ssDNA and generate ladder-like products, indicating that *HsMGME1* is also an exonuclease.

HsMGME1 belongs to the PD-(D/E)XK superfamily; in addition to MGME1 orthologues from other species, it also shares sequence similarity with RecE-type and RecB-type nucleases (Figure 1A). Dali search results showed that the structure of MGME1 is most similar to that of *EcRecE* (PDB_ID: 3H4R) (42), followed by Cas4 nuclease SSO0001 from *Sulfolobus solfataricus* (PDB_ID: 4IC1) (43) and an uncharacterized exonuclease from *Eubacterium rectale* (PDB_ID: 3L0A). MGME1 also shares structural similarity with the structure of phage-related exonuclease (PDB_ID: 3K93) and the structure of K131A mutant of phage lambda exonuclease (PDB_ID: 3SM4) (44), which ranked fifth and sixth on the Dali search result list, respectively. Though the other four structures are all nuclease-alone structures, DNA substrate was present in the phage lambda exonuclease structure. Interestingly, conformations of the DNA and catalytic residues and coordination of the cations (Supplementary Figure S5A) are very similar in the phage lambda exonuclease structure and our substrate DNA binding model of *HsMGME1*, which further supports the catalytic mechanism we proposed.

Besides ssDNA, MGME1 can also cleave dsDNA and DNA with 5'-flap or 3'-flap. The detailed basis for 3'-overhang DNA binding by *HsMGME1* was revealed by the *HsMGME1*-DNA3 structure (Figure 4A). As depicted in Figure 4B, *HsMGME1* will break the O-P bond between the third and fourth nucleotides from the 3'-end of the duplex. Our *in vitro* cleavage assays (Figure 4C) showed that 5' FAM-labeled dsDNA4 can be cleaved by MGME1 (or MGME1 mutants) from the 3'-end and generate ladder-like products, suggesting that DNA duplex can bind to *HsMGME1* in a manner similar to the duplex region of DNA3. MGME1 can also bind 5'-overhang DNA in a similar way with the 5'-overhang simply flipped out. In addition to the catalytic mechanism, structural comparison also revealed that the region corresponding to the *HsMGME1* helical arch is very flexible. This region is completely disordered in the *EcRecE* structure; unlike MGME1, this region is significantly tilted in the phage lambda exonuclease structure, forming one cleft suitable for DNA duplex or 5'-overhang DNA binding (Supplementary Figure S5B). Similar to dsDNA4, ladder-like products were also observed in the presence of MGME1 and 5'-overhang DNA4 (which was FAM-labelled at the 3'-end). These observations suggested that *HsMGME1* can undergo large conformational changes, bind, and cleave 5'-overhang DNA from the 5'-end similar to the phage lambda exonuclease.

Unlike *EcRecE*, Cas4 nuclease SSO0001, the uncharacterized *E. rectale* exonuclease, phage-related exonuclease, and phage lambda exonuclease, which all function as oligomers, *HsMGME1* functions as a monomer. This unique feature sets *HsMGME1* apart from other nucleases and contributes to the broad substrate tolerance of *HsMGME1*. Perhaps due to its ability to cleave various substrates, *HsMGME1* participates in different biological processes of mtDNA, including replication and double-strand degradation. *HsMGME1* shares sequence similarity with RecB-type nucleases, including AddA and RecB (Figure 1A). AddA ranked fourth on the Dali search result list; it is the catalytic component of DNA double-strand break repair complex AddAB that trims DNA in the 3'→5' direction (27). Structural comparison (Figure 6A and B) could clarify the substrate binding and catalytic mechanism of AddAB. As revealed by the recent *BsAddAB*-DNA complex structure (PDB_ID: 4CEJ), the 3'-end of the DNA is located near the active site of AddA. However, conformations of the local DNAs are significantly different in the AddAB-DNA complex and our H180Q-ssDNA2-Ca²⁺ complex (Supplementary Figure S6A) or *HsMGME1*-ssDNA2 complex, owing to disordering of the helical arch and mutation of catalytic residue Asp1172. In addition to catalysis, these observations suggest that proper coordination of cations is also important for catalytic complex assembly.

RecB ranked seventh on the Dali search result list; it is the catalytic component of another DNA double-strand break repair complex, RecBCD. Upon binding of the chi sequence, RecB will trim the overhang regions of substrate DNAs in the 5'→3' direction. Several structures of RecBCD in complex with DNA have been reported previously (45), but no DNA was captured at or near the active site of RecB (Figure 6C). In one RecBCD-DNA complex structure (PDB_ID: 1W36), one Ca²⁺ coordinated with the catalytic site residues of RecB; coordination of Ca²⁺ is very similar to Mn²⁺ at the A-site of the *HsMGME1*-Mn²⁺ structure (Supplementary Figure S6B) and Ca²⁺ in the H180Q-ssDNA2-Ca²⁺ complex, suggesting that RecBCD and *HsMGME1* may follow a similar mechanism in cation binding and catalysis. Perhaps due to the absence of interacting DNA, RecB does not form a helical arch in any RecBCD structures; however, structural analysis (Figure 6D) suggests that RecB could undergo large conformational changes to create a helical arch similar to the one in the *HsMGME1*-DNA complex. The helical arch may play an important role in substrate binding and catalytic process.

HsMGME1 has one unique structural pin (Figure 4B and C) that can facilitate DNA unwinding and enhance the cleavage efficiency of duplex-containing DNA by *HsMGME1*. Instead of the homologous protein RecB, one pin motif was identified in RecC previously; though the detailed conformations of the pins are not identical in *HsMGME1* and RecC, they may play similar roles in DNA unwinding. Besides the structural pin, all the components of RecBCD and AddAB complexes also contain helicase domains; precise incorporation among these helicases are essential for substrate DNA unwinding and repair of the breaks. *HsMGME1* does not contain a helicase domain, but *HsMGME1*, PolgA, and TWINKLE (a mitochondrial

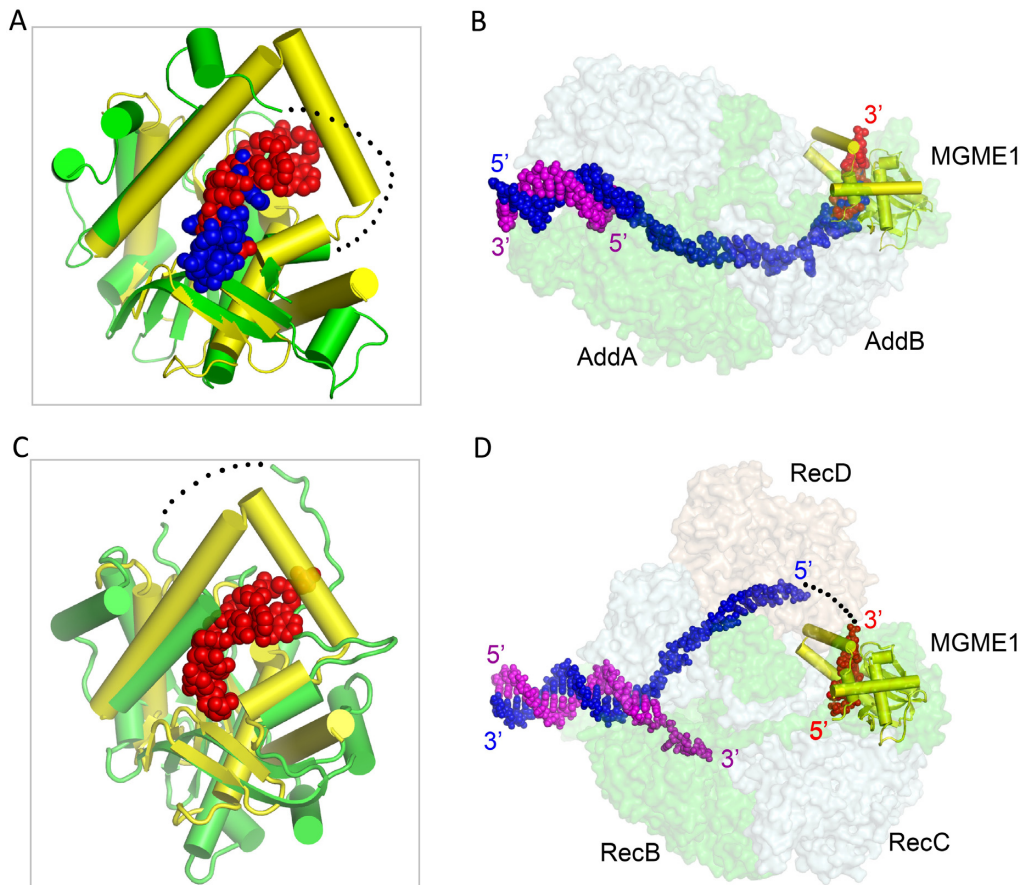


Figure 6. Structural comparison of *HsMGME1* and RecB-type nucleases. (A) Superposition of *HsMGME1*-ssDNA2 and the homologous domain from AddA within the AddAB-DNA complex (PDB.ID: 4CEJ). (B) Fitting of *HsMGME1*-ssDNA2 onto the AddAB-DNA complex. (C) Superposition of *HsMGME1*-ssDNA2 and the homologous domain from RecB within the RecBCD-DNA complex (PDB.ID: 1W36). (D) Fitting of *HsMGME1*-ssDNA2 onto the RecBCD-DNA complex. *HsMGME1* and ssDNA2 of the *HsMGME1*-ssDNA2 complex are shown as yellow cartoon and red spheres in all panels. The nuclease domains of AddA and RecB are shown as green cartoons in panels (A) and (C). In panels (B) and (D), RecBCD and AddAB protein complexes are shown as surface, but their substrate DNAs are shown as spheres in blue and magenta. Black dots indicated the residues disordered around the helical arch regions of RecB and AddA, and the possible pathway of substrate DNA 5'-end of RecBCD complex.

DNA helicase) can work together in the replication and double-strand degradation process of mtDNA (21,22). In the future, it is worth investigating whether TWINKLE can mimic the nuclease domains of RecBCD (or AddAB) and form stable complexes with *HsMGME1* and/or PolgA to ensure the rapid degradation of damaged mtDNAs.

7S DNA is the product of prematurely terminated replication of mtDNA, initiated at the NCR region; during the replication process, the NCR region and 7S DNA form a D-loop structure. Depending on the species, the size and sequence of the 5'-end of 7S DNA could be very different (18) and can be very rich in TMP or AMP (close to 50%). Both our results and previous *in vitro* cleavage assays all confirmed that *HsMGME1* can cleave ssDNAs in either the 5'→3' or 3'→5' direction. However, *in vivo* studies showed that loss of *HsMGME1* mainly led to the accumulation of 7S DNA with elongated 5'-ends in mitochondria. The detailed structural basis for the cleavage site preference and *in vivo* 5'-end preference of 7S DNA degradation of *HsMGME1* are still not clear. The N- and C-terminal region of *HsMGME1* are very flexible and have certain enhancing or inhibitory impacts on *in vitro* DNA degradation

by *HsMGME1* (Figure 5B). As revealed by structural comparison, the β3-α5 loop could also undergo large conformational changes (Supplementary Figure S1C). It is worth investigating whether these flexible regions of *HsMGME1* is involved in PolgA and TWINKLE interactions and contributes to the *in vivo* 7S DNA degradation directionality of *HsMGME1*.

In summary, we solved four high resolution structures of *HsMGME1* in this study. Our structures revealed the detailed basis for various DNA binding by *HsMGME1*. In combination with *in vitro* cleavage assays and structural analysis, our structures also reveal the possible catalytic mechanism for *HsMGME1* and homologous proteins, including AddA and RecB of the DNA double-strand break repair complex AddAB and RecBCD. Unlike many other mitochondrial nucleases present in both mitochondria and other cellular organelles, MGME1 is only found in mitochondria. Though how *HsMGME1* mutations lead to the associated mitochondrial diseases needs to be further investigated, our *HsMGME1* structures provided one potential target in regulating DNA replication and double-strand degradation processes specific to mitochondria.

DATA AVAILABILITY

Structure factors and coordinates have been deposited in the Protein Data Bank under accession codes 5ZYW, 5ZYU, 5ZYV and 5ZYT for the *Hs*MGME1–Mn²⁺, *Hs*MGME1–ssDNA2, H180Q–ssDNA2–Ca²⁺ and *Hs*MGME1–DNA3 complexes, respectively.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the BL17U1 and BL19U1 beamlines staff at the Shanghai Synchrotron Radiation Facility for assistance during data collection.

FUNDING

Key Research and Development Project of China [2016YFA0500600, 2015CB943300]; National Natural Science Foundation of China [31870721, 31470724]. Funding for open access charge: Key Research and Development Project of China [2016YFA0500600].

Conflict of interest statement. None declared.

REFERENCES

- Szczesny,R.J., Hejnowicz,M.S., Steczkiewicz,K., Muszewska,A., Borowski,L.S., Ginalski,K. and Dziembowski,A. (2013) Identification of a novel human mitochondrial endo-/exonuclease Ddk1/c20orf72 necessary for maintenance of proper 7S DNA levels. *Nucleic Acids Res.*, **41**, 3144–3161.
- Zheng,L., Zhou,M.A., Guo,Z.G., Lu,H.M., Qian,L.M., Dai,H.F., Qiu,J.Z., Yakubovskaya,E., Bogenhagen,D.F., Dempe,B. *et al.* (2008) Human DNA2 Is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates. *Mol. Cell*, **32**, 325–336.
- Zhu,Z., Chung,W.H., Shim,E.Y., Lee,S.E. and Ira,G. (2008) Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. *Cell*, **134**, 981–994.
- Stewart,J.A., Campbell,J.L. and Bambara,R.A. (2010) Dna2 is a structure-specific nuclease, with affinity for 5'-flap intermediates. *Nucleic Acids Res.*, **38**, 920–930.
- Lai,M.S. and Foiani,M. (2012) Dna2 offers support for stalled forks. *Cell*, **149**, 1181–1183.
- Tsuchimoto,D., Sakai,Y., Sakumi,K., Nishioka,K., Sasaki,M., Fujiwara,T. and Nakabeppu,Y. (2001) Human APE2 protein is mostly localized in the nuclei and to some extent in the mitochondria, while nuclear APE2 is partly associated with proliferating cell nuclear antigen. *Nucleic Acids Res.*, **29**, 2349–2360.
- Tigchelaar,W., Yu,H.J., de Jong,A.M., van Gilst,W.H., van der Harst,P., Westenbrink,B.D., de Boer,R.A. and Sillje,H.H.W. (2015) Loss of mitochondrial exo/endonuclease EXOG affects mitochondrial respiration and induces ROS-mediated cardiomyocyte hypertrophy. *Am. J. Physiol.-Cell Ph.*, **308**, C155–C163.
- Szymanski,M.R., Yu,W.S., Gmyrek,A.M., White,M.A., Molineux,I.J., Lee,J.C. and Yin,Y.W. (2017) A domain in human EXOG converts apoptotic endonuclease to DNA-repair exonuclease. *Nat. Commun.*, **8**, 14959.
- Stuart,J.A., Hashiguchi,K., Wilson,D.M., Copeland,W.C., Souza-Pinto,N.C. and Bohr,V.A. (2004) DNA base excision repair activities and pathway function in mitochondrial and cellular lysates from cells lacking mitochondrial DNA. *Nucleic Acids Res.*, **32**, 2181–2192.
- Mol,C.D., Izumi,T., Mitra,S. and Tainer,J.A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 DNA repair and coordination. *Nature*, **403**, 451–456.
- Duguay,B.A. and Smiley,J.R. (2013) Mitochondrial nucleases ENDOG and EXOG participate in mitochondrial DNA depletion initiated by Herpes Simplex Virus 1 UL12.5. *J. Virol.*, **87**, 11787–11797.
- Cymerman,I.A., Chung,I., Beckmann,B.M., Bujnicki,J.M. and Meiss,G. (2008) EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Res.*, **36**, 1369–1379.
- Burkovics,P., Szukacsov,V., Unk,I. and Haracska,L. (2006) Human Ape2 protein has a 3'-5' exonuclease activity that acts preferentially on mismatched base pairs. *Nucleic Acids Res.*, **34**, 2508–2515.
- Nicholls,T.J., Zsurka,G., Peeva,V., Scholer,S., Szczesny,R.J., Cysewski,D., Reyes,A., Kornblum,C., Sciacco,M., Moggio,M. *et al.* (2014) Linear mtDNA fragments and unusual mtDNA rearrangements associated with pathological deficiency of MGME1 exonuclease. *Hum. Mol. Genet.*, **23**, 6147–6162.
- Kornblum,C., Nicholls,T., Haack,T.B., Schoeler,S., Peeva,V., Danhauser,K., Hallmann,K., Zsurka,G., Rorbach,J., Iuso,A. *et al.* (2013) Loss of function of MGME1, a novel player in mitochondrial DNA replication, causes a distinct autosomal recessive mitochondrial disorder. *Neuromusc. Disord.*, **23**, 852–852.
- Kornblum,C., Nicholls,T.J., Haack,T.B., Scholer,S., Peeva,V., Danhauser,K., Hallmann,K., Zsurka,G., Rorbach,J., Iuso,A. *et al.* (2013) Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease. *Nat. Genet.*, **45**, 214–219.
- Nicholls,T.J. and Minczuk,M. (2014) In D-loop: 40 years of mitochondrial 7S DNA. *Exp. Gerontol.*, **56**, 175–181.
- Brown,W.M., Shine,J. and Goodman,H.M. (1978) Human mitochondrial DNA: analysis of 7S DNA from the origin of replication. *Proc. Natl. Acad. Sci. U.S.A.*, **75**, 735–739.
- Uhler,J.P., Thorn,C., Nicholls,T.J., Matic,S., Milenkovic,D., Gustafsson,C.M. and Falkenberg,M. (2016) MGME1 processes flaps into ligatable nicks in concert with DNA polymerase gamma during mtDNA replication. *Nucleic Acids Res.*, **44**, 5861–5871.
- Macao,B., Uhler,J.P., Siibak,T., Zhu,X., Shi,Y., Sheng,W., Olsson,M., Stewart,J.B., Gustafsson,C.M. and Falkenberg,M. (2015) The exonuclease activity of DNA polymerase gamma is required for ligation during mitochondrial DNA replication. *Nat. Commun.*, **6**, 7303.
- Peeva,V., Blei,D., Trombly,G., Corsi,S., Szukacsov,M.J., Rebelo-Guiomar,P., Gammage,P.A., Kudin,A.P., Becker,C., Altmuller,J. *et al.* (2018) Linear mitochondrial DNA is rapidly degraded by components of the replication machinery. *Nat. Commun.*, **9**, 1727.
- Matic,S., Jiang,M., Nicholls,T.J., Uhler,J.P., Dirksen-Schwandenland,C., Polosa,P.L., Simard,M.L., Li,X.P., Atanassov,I., Rackham,O. *et al.* (2018) Mice lacking the mitochondrial exonuclease MGME1 accumulate mtDNA deletions without developing progeria. *Nat. Commun.*, **9**, 1202.
- Wigley,D.B. (2013) Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. *Nat. Rev. Microbiol.*, **11**, 9–13.
- Wigley,D.B. (2007) RecBCD: the supercar of DNA repair. *Cell*, **131**, 651–653.
- Spies,M., Amitani,I., Baskin,R.J. and Kowalczykowski,S.C. (2007) RecBCD enzyme switches lead motor subunits in response to chi recognition. *Cell*, **131**, 694–705.
- Meima,R., Haijema,B.J., Venema,G. and Bron,S. (1995) Overproduction of the ATP-dependent nuclease AddAB improves the structural stability of a model plasmid system in *Bacillus subtilis*. *Mol. Gen. Genet.*, **248**, 391–398.
- Krajewski,W.W., Fu,X., Wilkinson,M., Cronin,N.B., Dillingham,M.S. and Wigley,D.B. (2014) Structural basis for translocation by AddAB helicase-nuclease and its arrest at chi sites. *Nature*, **508**, 416–419.
- Minor,W., Cymborowski,M., Otwinowski,Z. and Chruszcz,M. (2006) HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. D. Biol. Crystallogr.*, **62**, 859–866.
- Giacovazzo,C. and Siliqi,D. (2004) Phasing via SAD/MAD data: the method of the joint probability distribution functions. *Acta Crystallogr. D. Biol. Crystallogr.*, **60**, 73–82.
- Terwilliger,T.C., Adams,P.D., Read,R.J., McCoy,A.J., Moriarty,N.W., Grosse-Kunstleve,R.W., Afonine,P.V., Zwart,P.H.

- and Hung, L.W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D. Biol. Crystallogr.*, **65**, 582–601.
31. Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K. and Terwilliger, T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D. Biol. Crystallogr.*, **58**, 1948–1954.
 32. Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F. and Vagin, A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.*, **67**, 355–367.
 33. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.*, **60**, 2126–2132.
 34. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.*, **68**, 352–367.
 35. Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S. and Ji, X. (2006) Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell*, **124**, 355–366.
 36. Nowotny, M., Gaidamakov, S.A., Crouch, R.J. and Yang, W. (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell*, **121**, 1005–1016.
 37. Goedken, E.R. and Marqusee, S. (2001) Co-crystal of Escherichia coli RNase HI with Mn²⁺ ions reveals two divalent metals bound in the active site. *J. Biol. Chem.*, **276**, 7266–7271.
 38. Zhang, J., Liu, H., Yao, Q., Yu, X., Chen, Y., Cui, R., Wu, B., Zheng, L., Zuo, J., Huang, Z. *et al.* (2016) Structural basis for single-stranded RNA recognition and cleavage by C3PO. *Nucleic Acids Res.*, **44**, 9494–9504.
 39. Gan, J., Shaw, G., Tropea, J.E., Waugh, D.S., Court, D.L. and Ji, X. (2008) A stepwise model for double-stranded RNA processing by ribonuclease III. *Mol. Microbiol.*, **67**, 143–154.
 40. Liang, Y.H., Lavoie, M., Comeau, M.A., Abou Elela, S. and Ji, X. (2014) Structure of a eukaryotic RNase III postcleavage complex reveals a double-ruler mechanism for substrate selection. *Mol. Cell*, **54**, 431–444.
 41. Weinberg, D.E., Nakanishi, K., Patel, D.J. and Bartel, D.P. (2011) The inside-out mechanism of Dicers from budding yeasts. *Cell*, **146**, 262–276.
 42. Zhang, J., Xing, X., Herr, A.B. and Bell, C.E. (2009) Crystal structure of E. coli RecE protein reveals a toroidal tetramer for processing double-stranded DNA breaks. *Structure*, **17**, 690–702.
 43. Lemak, S., Beloglazova, N., Nocek, B., Skarina, T., Flick, R., Brown, G., Popovic, A., Joachimiak, A., Savchenko, A. and Yakunin, A.F. (2013) Toroidal structure and DNA cleavage by the CRISPR-associated [4Fe-4S] cluster containing Cas4 nuclease SSO0001 from *Sulfolobus solfataricus*. *J. Am. Chem. Soc.*, **135**, 17476–17487.
 44. Zhang, J., McCabe, K.A. and Bell, C.E. (2011) Crystal structures of lambda exonuclease in complex with DNA suggest an electrostatic ratchet mechanism for processivity. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 11872–11877.
 45. Singleton, M.R., Dillingham, M.S., Gaudier, M., Kowalczykowski, S.C. and Wigley, D.B. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature*, **432**, 187–193.