2258–2270 Nucleic Acids Research, 2012, Vol. 40, No. 5 Published online 7 November 2011
doi:10.1093/nar/gkr999

Crystal structure of the NurA–dAMP–Mn^{2+} complex

Jina Chae[1], Young Chang Kim[2] and Yunje Cho[1,*]

[1]Department of Life Science, Pohang University of Science and Technology, Pohang 790–784, South Korea and [2]Biosciences Division, Structural Biology Center, Argonne National Laboratory, Argonne, IL60439, USA

Received August 11, 2011; Revised October 17, 2011; Accepted October 18, 2011

ABSTRACT

Generation of the 3′ overhang is a critical event during homologous recombination (HR) repair of DNA double strand breaks. A 5′–3′ nuclease, NurA, plays an important role in generating 3′ single-stranded DNA during archaeal HR, together with Mre11–Rad50 and HerA. We have determined the crystal structures of apo- and dAMP-Mn^{2+}-bound NurA from *Pyrococcus furiousus* (*Pf* NurA) to provide the basis for its cleavage mechanism. *Pf* NurA forms a pyramid-shaped dimer containing a large central channel on one side, which becomes narrower towards the peak of the pyramid. The structure contains a PIWI domain with high similarity to argonaute, endoV nuclease and RNase H. The two active sites, each of which contains Mn^{2+} ion(s) and dAMP, are at the corners of the elliptical channel near the flat face of the dimer. The 3′ OH group of the ribose ring is directed toward the channel entrance, explaining the 5′–3′ nuclease activity of *Pf* NurA. We provide a DNA binding and cleavage model for *Pf* NurA.

INTRODUCTION

DNA double-stranded breaks (DSB) are generated by genotoxic stresses such as reactive oxygen species, ionizing radiation, chemical agents and by normal cellular processes such as V(D)J recombination and DNA replication. If not properly repaired, DSBs cause chromosome losses or deletions, translocations and genomic instability. Cells primarily employ non-homologous end joining (NHEJ) and homologous recombination (HR) mechanisms to repair DSBs (1). While the NHEJ process involves processing and direct ligation of the DNA ends, the HR pathway uses a homologous template to repair DNA breaks.

HR processing is believed to be preferred route in bacteria (2). RecBCD, a helicase-nuclease in *Escherichia coli*, initiates HR repair by binding and degrading both ends of DNA as it translocates (3). When RecBCD binds to a *cis*-acting Chi sequence, its nuclease activity becomes attenuated and the polarity of the nuclease activity switches to 3′ strand nicking at Chi, which leads to 5′ DNA strand resection. The RecBCD-Chi interaction also facilitates loading of RecA recombinase onto 3′ single-stranded (ss) DNA, which catalyzes strand exchange. Unlike bacterial HR, no homologous RecBCD proteins have been identified in archaea and eukaryotes. Instead, the Mre11–Rad50 complex starts the archaeal and eukaryotic HR processes (4).

In eukaryotes, Mre11–Rad50–Nbs1 (human) or Mre11–Rad50-Xrs2 (budding yeast) in conjunction with Ctp1/CtIP (or Sae2) initially recognizes DNA DSBs and removes small nucleotides to form an early intermediate during HR (5–7). Exo1 nuclease, Sgs1 helicase and Dna2 nuclease bind to this intermediate and generate 3′ ssDNA (8,9). Exo1 is a member of the Rad2 family of structure-specific nucleases which possess 5′–3′ exonuclease and 5′ flap endonuclease activities *in vitro*, generating mononucleotide products (10). The 3′ ss tails are Rad51 recombinase substrates that initiate the homology search and strand invasion for recombination (11).

In archaea, the Mre11–Rad50 complex also senses DNA DSB ends, and processes DSB ends through 3′–5′ exonuclease and endonuclease activities to form a short 3′ overhang (12–14). However, NurA and HerA are believed to replace the role of Exo1 and Sgs1 to generate 3′ ssDNA, in which the RadA recombinase subsequently binds and catalyzes strand exchange (15). NurA, a highly conserved protein in archaea, exhibits a 5′–3′ ss- and dsDNA exonuclease and ssDNA endonuclease activities and functions together with HerA helicase to unwind dsDNA in both the 5′–3′ and 3′–5′ directions in an ATP-dependent manner (15–18). In all archaea genomes, the NurA and HerA genes are present in one operon along with genes encoding Mre11 and Rad50 (17). An analysis of transcriptional responses to induction by ultraviolet radiation revealed that the nurA, mre11, rad50 and herA genes are upregulated 2- to 3-fold in some strains of *Sulfolobus solfataricus* (19) and about 10-fold in other *Sulfolobus* strains (20), suggesting that the four proteins are associated and involved in DNA repair by recombination.

NurA physically interacts with HerA and functions cooperatively with the Mre11–Rad50 complex in *Pyrococcus*

*To whom correspondence should be addressed. Tel: +82 54 279 2288; Fax: +82 54 279 8111; Email: yunje@postech.ac.kr

furiosus (15). However, the NurA–HerA complex does not interact directly with the *Pf*Mre11–Rad50 complex. In contrast, Mre11 directly interacts and stimulates HerA helicase activity in *Sulfolobous tokodaii* and *Sulfolobous acidocaldarius* (21,22). NurA from *S. tokodaii* physically interacts with SSB to inhibit nuclease activity, implicating a possible regulatory role in Mre11–Rad50-mediated DSB repair (23).

Genetic analysis using *Thermococcus kodakaraensis* showed that cells lacking HerA or NurA are difficult to grow if intra-chromosomal recombination occurs to delete these genes, which illustrates the significance of these genes in cell proliferation (24). Secondary structure predictions of NurA suggest that NurA forms a novel pattern that does not bear any resemblance to the previously characterized folds found in nucleases or other proteins (25). The NurA group appears to be a rapidly diverging group, with a low level of sequence similarity between paralogous and even within orthologous groups (25). While Exo1, the eukaryotic functional homolog of NurA, has been extensively studied both *in vitro* and *in vivo*, not much information is available on NurA from the biochemical and structural aspects, despite its importance in archaeal HR.

To understand the functional and structural conservation of NurA and Exo1, and to provide the basis for 5′ to 3′ processing, we determined the structures of apo- and dAMP–Mn^{2+}-bound *Pf*NurA and characterized its biochemical properties. We show that *Pf*NurA alone cleaved the 5′-end of DNA substrates, whereas *Pf*NurA exhibited both 5′ and 3′ exonuclease activities in the presence of *Pf*HerA. *Pf*NurA possesses a PIWI domain with an RNase H fold, and forms a pyramid-shaped dimer through two interfaces and contains a wide channel on one face, which becomes narrower the other end. The surface near the channel contains two active sites at its corners, where Mn^{2+} ion(s) and a dAMP molecule bind to each active site. We show the active site architecture, which explains the basis for its nuclease activity. Finally, we provide a model for the NurA–DNA and HerA-NurA complexes.

MATERIALS AND METHODS

Protein expression and purification

All constructs were generated using a standard polymerase chain reaction (PCR)-based strategy. Genes encoding residues 1–451 of *Pf*NurA were inserted into the pET28a vector. *Escherichia coli* Rosetta (DE3) containing the vector was cultured in LB broth medium at 37°C. When the absorbance at 600 nm reached 0.6, the culture was induced with 1 mM isopropyl-β-D-thiogalactopyranoside for 4 h. The cells were harvested by centrifugation, and cell pellets containing the *Pf*NurA were resuspended in 25 mM Tris–HCl (pH 7.4), 300 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol and 1 mM phenylmethanesulfonylfluoride, and lysed by sonication. *Pf*NurA was initially purified by Ni–NTA affinity chromatography using a His-tag at the N-terminus of *Pf*NurA, and subsequently purified using cation exchange (HiTrap-SP) and gel-filtration chromatography (Superdex 200), concentrated to 10 mg/ml by ultrafiltration, and stored at −80°C. To facilitate structure determination, we made selenomethionine (Se-Met)-substituted *Pf*NurA by growing the *E. coli* strain B834 (DE3) harboring the pET28a-*Pf*NurA plasmid in M9 minimal medium. The Se-Met-*Pf*NurA was purified by the same method used to purify the native protein.

Crystallization and data collection

Crystals of the apo-*Pf*NurA were grown at 22°C by the hanging drop vapor diffusion method. The crystallization buffer contained 21–25% PEG400, 0.1 M Tris–HCl pH 8.5 and 5 mM dithiothreitol. Crystals were soaked in a solution containing crystallization buffer, 5 mM $MnCl_2$ and 10 mM dAMP for 8 days to obtain dAMP and Mn^{2+}-bound crystals before data collection. The Mn^{2+}-bound crystals were soaked in a crystallization buffer containing 100 mM $MnCl_2$ for 4 h before data collection. Diffraction data were collected at −170°C using crystals flash-frozen in crystallization buffer containing 25% (v/v) glycerol. Diffraction data were collected from apo- and Mn^{2+}-bound crystals at 0.9795 Å and 1.0000 Å, respectively, on beamline 4A at the Pohang Advanced Light Source. dAMP–Mn^{2+}-bound *Pf*NurA crystals were collected at 1.5418 Å using a home source. The apo-*Pf*NurA crystals form in space group $P2_12_12_1$ with $a = 65.0$ Å, $b = 114.2$ Å and $c = 122.3$ Å and have one dimer molecule in the asymmetric unit (Table S1). Data integration, scaling and merging were performed using the HKL2000 package (26).

Structure determination and refinement

*Pf*NurA structure was determined by single-wavelength anomalous scattering dispersion (SAD) method using Se-Met-derivatized crystals. Fourteen selenium sites were identified in the asymmetric unit, and the apo *Pf*NurA structure was determined with the PHENIX program (27). After flattening the solvent, a high-quality electron density map with a resolution of 2.82 Å was obtained. Successive rounds of model building using COOT (28) and refinement using CNS (29) and PHENIX were performed to build the complete model. Restrained NCS was applied throughout the refinement. The final apo- model consists of 860 residues, seven glycerol molecules and 19 water molecules. The structure of Mn^{2+}-bound *Pf*NurA and dAMP–Mn^{2+}-bound *Pf*NurA were determined by difference Fourier methods using the apo–*Pf*NurA structure. The Mn^{2+}-bound *Pf*NurA consists of 845 residues, four Mn^{2+} ions, one glycerol molecule and 20 water molecules. The dAMP and Mn^{2+}-bound model consists of 865 residue, two Mn^{2+} ions, one dAMP molecule, one glycerol molecule and nine water molecules. The N-terminal His-tagged residues were not visible and were presumably disordered. The statistics are summarized in Supplementary Table S1. Missing residues in each structure are described in the Supplementary Data (Information on substrates, nuclease assays, analytical ultracentrifugation and mutant protein structural analysis is described in the Supplementary Data).

RESULTS

Nuclease activities of *Pf*NurA and the *Pf*HerA–NurA complex

Previous studies have reported that NurA from *Sulfolobous acidocaldarius* (*Sa*NurA) exhibits 5–3′ exonuclease activity which produced ~5 nt from the 5′-end of a substrate (16). *Pf*NurA also displays exonuclease activity (15). However, it is unclear if *Pf*NurA produces any specific cleavage products like *Sa*NurA. To understand if the products of the archaeal NurA nucleases are conserved, we further examined the nuclease activities of *Pf*NurA using a 5′[^{32}P]-labeled dsDNA oligonucleotide substrate with a 16-nt 3′-overhang (TP424/423) and a 50-nt duplex DNA containing five phosphorothioate bonds at the 3′-end of the top strand (TP580/124). *Pf*NurA (350 nM) in the presence of Mn^{2+} exhibited weak endonuclease activity on 20 nM of each substrate, and cleaved primarily ~10 nt from the 5′-end (Figure 1A, lanes 2 and 9). In the presence of a higher amount of *Pf*NurA (1750 nM), smaller amounts of 6- and 8-nt products were generated from the 5′-end of these substrates (Figure 1A, lanes 3 and 10).

Because NurA functions together with HerA, we examined the effect of *Pf*HerA on *Pf*NurA nuclease activity (Figure 1B) using various ratios of *Pf*HerA and *Pf*NurA of; 1:3, 1:5 and 1:10. *Pf*HerA (115 nM) alone did not exhibit any nuclease activity. We have used 350 nM of *Pf*NurA which generated a 10-nt product from the 5′-end. Interestingly, we found that adding various amounts of *Pf*HerA altered the *Pf*NurA products by cleaving ~1 or 2 nt from both ends of the TP424/423 substrate in an exonucleolytic manner (Figure 1B, lanes 4–6). A small amount of 10-nt product was formed in the presence of a reduced amount of *Pf*HerA (35 nM), which was no longer visible when *Pf*HerA was increased to 70 nM. *Pf*NurA also exhibited 5′-strand exonucleolytic activity towards the TP580/124 substrate in the presence of *Pf*HerA (Figure 1B, lanes 10–12). Thus, our data show that *Pf*NurA exhibits 5′-endonucleolytic activity alone and together with *Pf*HerA, and that *Pf*NurA has exonucleolytic activity on both 5′ and 3′ of a substrate.We presumed that *Pf*HerA unwinds the DNA substrates in both directions and allows them to be more susceptible to degradation, which resulted in the generation of altered products by *Pf*NurA in the presence of *Pf*HerA.

Overall structure of *Pf*NurA

Analytical ultracentrifuge and gel filtration analyses suggest that *Pf*NurA forms a dimer, consistent with the biochemical studies of *Sto*NurA (30) and *Pf*NurA (15) (Supplementary Figure S1). The structure of the full-length *Pf*NurA was determined at 2.82 Å resolution using single-wavelength anomalous dispersion method (Supplementary Table S1).

The crystal structure of *Pf*NurA revealed that the monomer forms an α/β fold that consists of 16 α-helices and 13 β-strands (Figure 2A and Supplementary Figure S2). Each subunit consists of two extended N-terminal helices and a compact body that assembles into a pyramid-shaped dimer with dimensions of 78 × 55 × 62 Å (Figure 2B and C). A large central channel is formed at the base of the pyramid and becomes narrower toward a peak that is formed by the two-helices (H7 and H8) from each subunit.

The overall structure of the *Pf*NurA monomer is highlighted by a twisted central eight-stranded sheet that is packed by several helices on both sides (Figure 2A). This part of the structure strongly resembles the PIWI domain with an RNase H fold, which is found in argonaute proteins (31–33). We divided the *Pf*NurA subunit into three domains (Figure 2A and B): the N-domain (magenta color, residues 3–43), the PIWI domain (aqua color, residues 44–188 and 287–441), and the M domain (orange colored middle domain, residues 189–286). In Figure 2C, the PIWI and N-domains form the base and the M-domain is on the top of the base of the pyramid to form a peak.

The N domain is composed of two perpendicularly located helices (H1 and H2), which play an important role in dimerization. The PIWI domain of *Pf*NurA, an α/β structure with an RNase H fold, is dominated by a central mixed β-sheet consisting of eight β-strands (11, 12, 13, 10, 4, 1, 2 and 3 order). The β-sheet can be divided into three-long strands (S1–S3) and five short strands (S4, S10–S13), with a deep cleft between them. The three long strands are flanked by three continuous helices (H14 to H16) on one face and helices H3 and H5 on the opposite face. The five short strands are packed by helix H13 on one side and helices H4 and H10 on another side (Figure 2A).

The PIWI domain is divided into two regions by the M domain; one part of the PIWI domain is formed with five strands (S1–S3, S4, S5) and three helices (H3–H5), which is followed by the M domain (orange color) with two helices (H6 and H9), a hairpin (S6 and S7), two strands (S8 and S9) and two additional helices (H7 and H8). Subsequently, another part of the PIWI domain with four strands (S10–S13), four helices (H10–H13) and three extended helices (H14–H16) that form the base of the pyramid and boundary of the channel follows.

Overall, the two *Pf*NurA subunits are similar with a r.m.s.d. value of 0.72 Å for all Cα atoms. However, the two subunits show significant differences in the region M domain region (residues 216–257) (Supplementary Figure S3). Different parts of the M domain contain similar secondary structures with three strands (S7–S9) and two helices (H7 and H8). However, they are separated by >20 Å and to approximately superimpose, one subunit has to be rotated by 42° and translated by 22 Å onto another *Pf*NurA.

*Pf*NurA active sites are located at the two corners of an elliptical hole, surrounded by helices H3 and H14 and strands S1 and S4. The active site is ~8 Å deep from the surface of the outer channel, and the two active sites in the *Pf*NurA dimer are ~29 Å apart. The dAMP-binding site is further away from the surface of the channel compared to that of the Mn^{2+} site (Figure 2B). The highly conserved C-terminal region of another subunit is directed toward the active site cleft of the *Pf*NurA subunit.

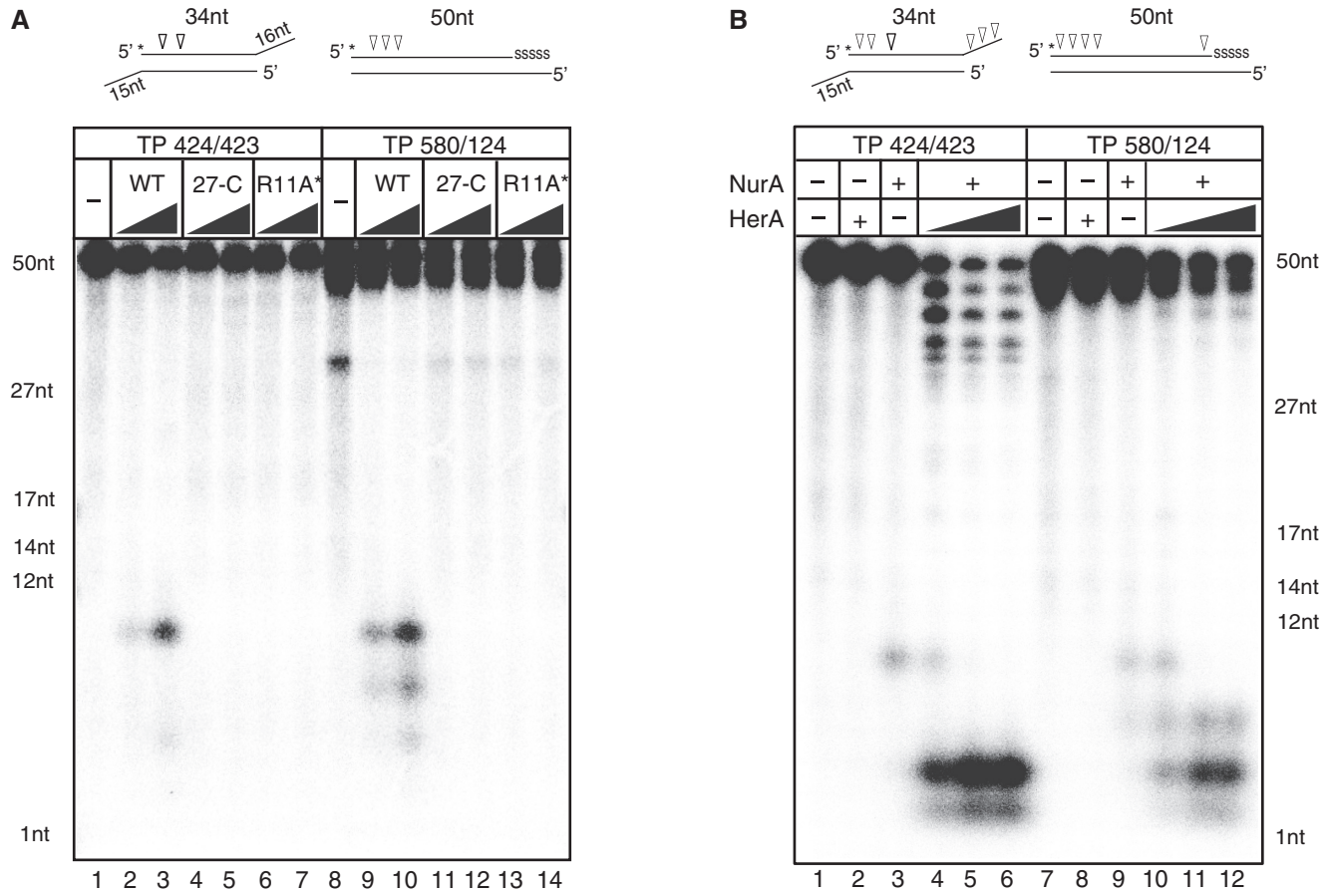

Figure 1. Biochemical properties of *Pf*NurA and *Pf*HerA (**A**) Comparison of nuclease activities of *Pf*NurA and dimeric interface mutants. TP 424/423 (lanes 1–7) or TP 580/124 (lanes 8–14) DNA substrates (20 nM) were incubated for 120 min at 65°C with increasing amounts of *Pf*NurA (350 and 1750 nM) and 5 mM $MnCl_2$. 27-C (Δ1–26) and R11A/I12E/S60Y (R11A*) are the *Pf*NurA dimeric interface mutant proteins. Reaction mixtures were analyzed on 15% denaturing polyacrylamide gels containing 7 M urea in TBE buffer. Schematic diagrams of each substrate with a ^{32}P-labeled 5′-end (asterisk) and cleavage sites (arrows) are shown on the top. Five nucleotides in one 3′-end are connected through phosphorothioate bonds, which are shown as 'SSSSS'. ssDNA markers are indicated. (**B**) Nuclease activities of *Pf*NurA in the presence of *Pf*HerA. DNA substrates (20 nM) were incubated for 120 min at 65°C with 350 nM of *Pf*NurA, increasing amounts of *Pf*HerA (35, 70 and 115 nM), 5 mM $MnCl_2$ and 1 mM ATP. Lanes 2 and 8 contain 115 nM *Pf*HerA.

Dimerization of *Pf*NurA creates a large central channel

The base of the pyramid (front view of the dimer in Figure 2B and D) is formed with an elliptical channel with three helices (H14–H16) from each subunit, which is followed by strands S1, S4 and S10, and helices H3, H4 and H5 in the next layer. As the channel directed toward the peak, the width of the channel becomes narrower and helices H7 and H8 form the last layer (Figure 2E). Although the channel that has a positive electrostatic potential becomes narrower, it is large enough to hold ssDNA (Figure 2D). In particular, several conserved residues including Arg54, Arg58 and Arg135 are aligned along the channel.

Dimeric interface of *Pf*NurA

*Pf*NurA forms a tightly packed dimer with a buried surface area of 7481 Å^2. These two interfaces play a major role in dimerization (Figure 3A–D). In the first interface which is formed with the N and PIWI domains, helix H1 from one *Pf*NurA protrudes into the pocket formed by strand s3 and helices h6 and h15 from another *Pf*NurA, and embraces another *Pf*NurA (Figures 2B, 2C and 3A–C). Here, helix H11 also interacts with helix h16 from another *Pf*NurA. In the second interface which is formed with the M domain, helices H7 and H8 from one *Pf*NurA interact with the equivalent helices from another *Pf*NurA (Figures 2C and 3D). In addition, strand S8 from a subunit interacts with strands s8 and s9 from another *Pf*NurA. The first interface is observed in many archaeal NurA proteins, whereas the second interface is likely to be unique to a few NurA proteins that only contain the M domain region (Supplementary Figure S2). Interestingly, although *Thermotoga maritima* NurA (*Tm*NurA) shares part of the conserved region in the first interface, it exhibits a different dimeric interface mode (see below), which suggests that even the first dimeric interface may not be structurally conserved in various NurA family members.

In the first interface, the H1-s3 interaction is largely formed by hydrogen (H)-bonds and hydrophobic interactions. Here, four H-bonds including Arg11–Ala83

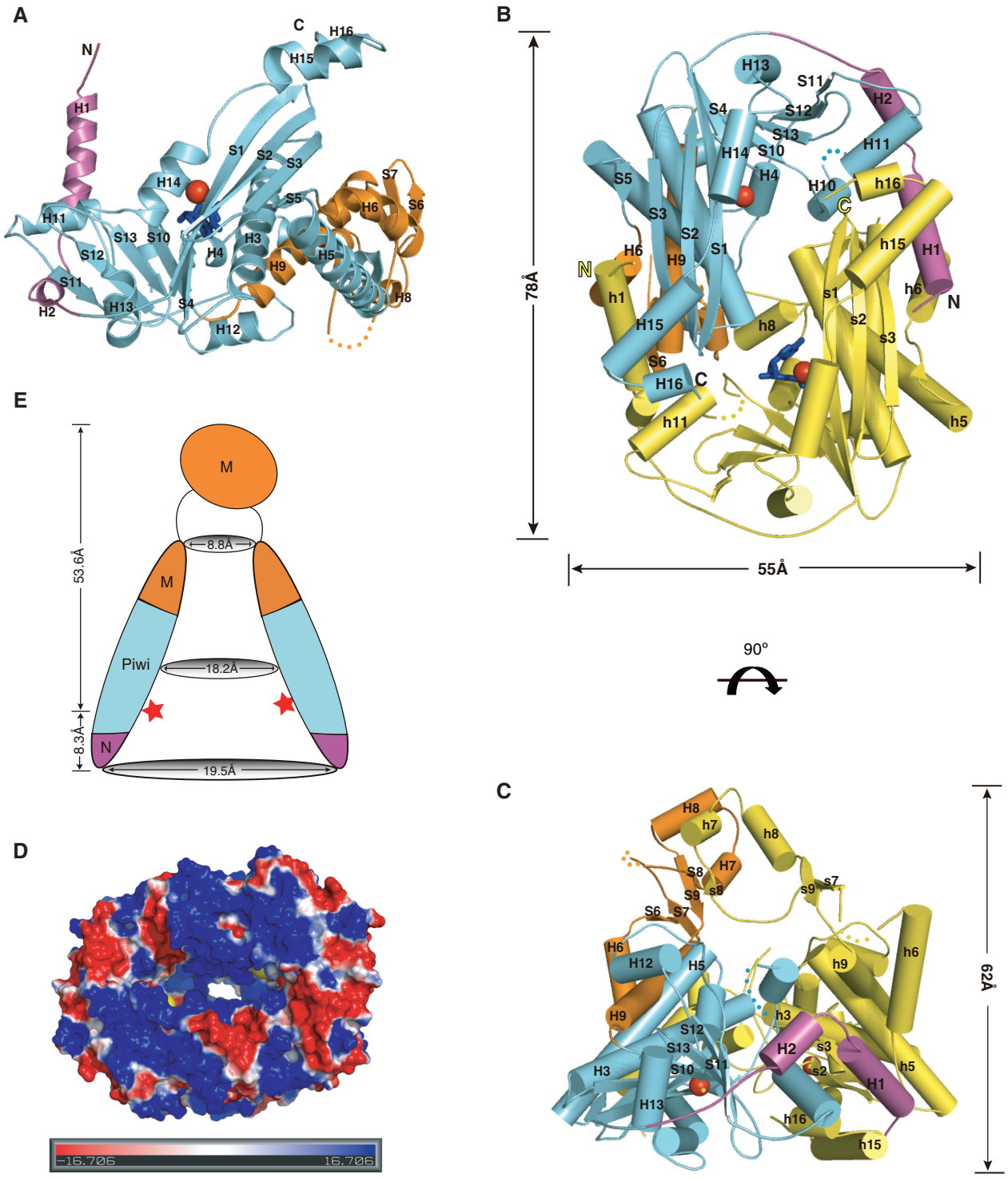

Figure 2. Overall structure of *Pf* NurA. (**A**) Structure of the *Pf* NurA monomer. The N-(magenta), PIWI (aqua), and M (orange) domains are shown with Mn^{2+} ions (red spheres). (**B**) The flat side of the *Pf* NurA dimer is viewed from the front and (**C**) from a 90° rotated view (along the horizontal axis of Figure 2B). Three domains of one monomer are shown in the same color scheme as in 2A (magenta, aqua and orange) and another monomer is shown in yellow for clarity. Secondary structures are labeled. Red spheres represent Mn^{2+} ions, and the two dAMP molecules are colored in blue. The disordered region is shown as dots. Dimension of the central hole is 29×19 Å^2. (**D**) Electrostatic potential of the *Pf* NurA dimer mapped onto the solvent-accessible protein surface (blue indicates positive regions; red indicates negative regions; the metal ion is in yellow). The figure is shown as a 90° rotated view of 2B along a 2-fold axis. (**E**) A schematic drawing of the *Pf* NurA central channel in the same view as that of 2C.The Mn^{2+} ions are represented as red stars.

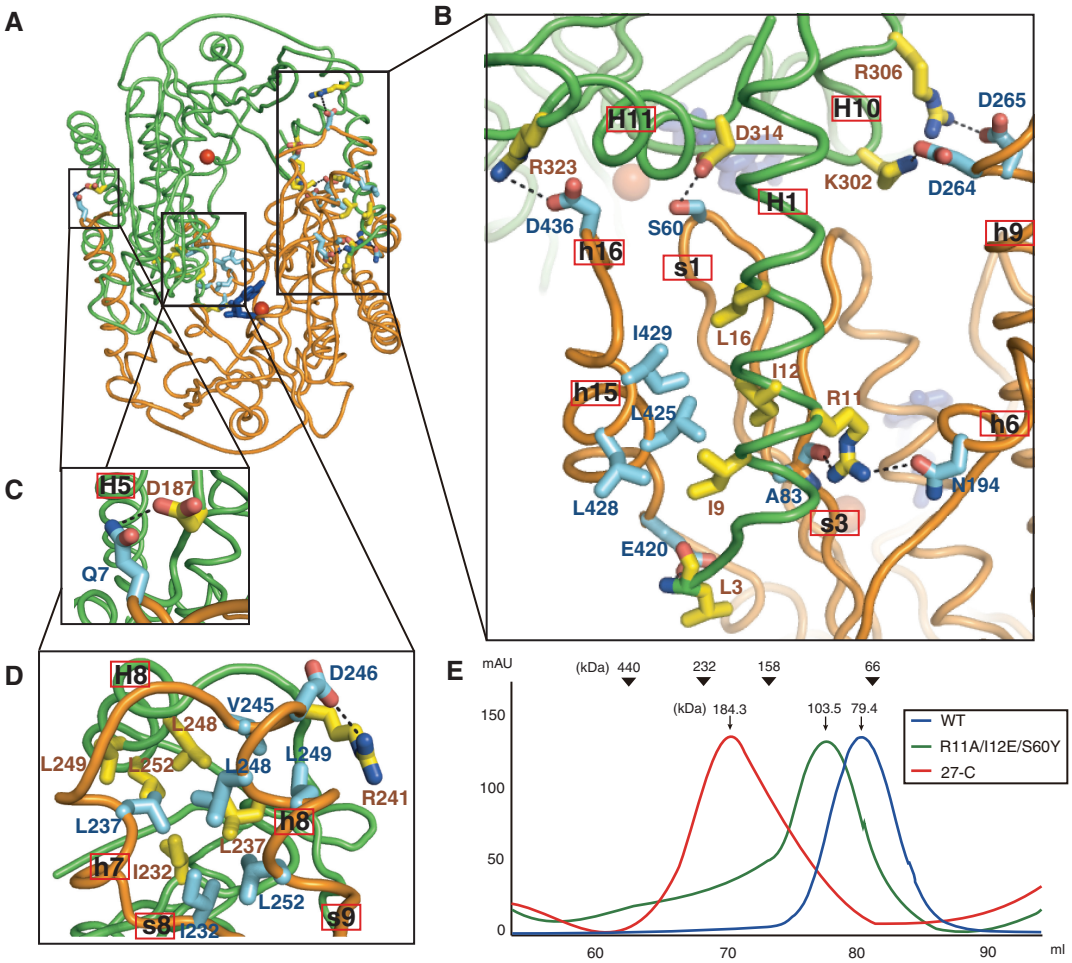

Figure 3. Structural analyses of the dimeric interface. (**A**) Overall view of the *Pf*NurA dimeric interface. Two different *Pf*NurA subunits are colored in green and orange.(**B**) A close-up view of the first interface of *Pf*NurA dimer. In the first interface, one subunit (green) and another *Pf*NurA (orange) interact primarily through the N- and PIWI domains as described in the text. (**C**) In addition to these interactions, an ion pair Gln7–Asp187 is present. (**D**) In the second interface, helices H7 and H8 of one *Pf*NurA interact with helices h7 and h8 of another *Pf*NurA. Strands s8 and s9 are also involved in dimer formation. Oxygen and nitrogen atoms are shown in red and blue, respectively. (**E**) Comparison of the oligomeric states of the wild-type and dimeric interface *Pf*NurA mutants using gel filtration chromatography. Standard molecular weights are shown at the top. Gel filtration analysis (Superdex 200) was performed using a buffer containing 25 mM Tris–HCl (pH 7.4), 200 mM NaCl, 5% glycerol and 5 mM DTT.

(main chain) are observed. In addition, three ion-pairs including Asp264–Lys302, Asp265–Arg306 and Arg323–Arg436 occur (Figure 3B). Hydrophobic interactions further stabilize this interface; Ile9, Ile12 and Leu16 from a subunit and are tightly packed against Leu425, Leu428 and Ile429 from another *Pf*NurA (Figures 2C and 3B).

To examine the importance of the N-terminal helix H1 in the dimer interface, we generated two *Pf*NurA mutant proteins by deleting the first 26 residues or by simultaneously replacing Arg11 with Ala, Ile12 with Glu, and Ser60 with Tyr. Both of these *Pf*NurA mutant proteins exhibited abrogated nuclease activities on the two different DNA substrates (Figure 1A). Gel filtration analysis revealed that a triple (R11A/I12E/S60Y) mutant eluted earlier than that of the wild-type *Pf*NurA at about a dimer size, and that the Δ*N26* mutant protein eluted at a tetramer size (Figure 3E), suggesting that the mutations on helix H1 affect the dimeric interface and alter the overall oligomer shape of *Pf*NurA.

In the second interface, hydrophobic interactions are highlighted by Ile232 (S8) and Leu237 (H7), which bind to the pocket formed by Ile232, Leu237, Val245, Leu248, Leu249 and Leu252 (Figure 3D). An ion-pair between Arg241 and Asp246 further stabilizes this interface.

Active site of *Pf*NurA

To identify the metal and substrate binding sites, we added 5 mM $MnCl_2$ and 10 mM dAMP to the *Pf*NurA crystal, which clearly revealed a Mn^{2+} ion site in each subunit, coordinated by Asp51 and Asp126, and lying close to Glu105 (Figure 4A). The distance between Mn^{2+} and the side chain of Asp51 or Asp126 is between 2.2 to 2.5 Å, whereas Glu105 is further from the Mn^{2+} ion. In the *Pf*NurA subunit, the active site is positioned in a cleft formed between the long (S1) and short (S4) strands. Figure 4B shows that the bottom of the cleft is formed by loop S4–H4, the right wall is formed by strands S1 and S4 and helix H3, the left wall is formed by helix H4 and strand S10. Helix H5 forms the back of the cleft.

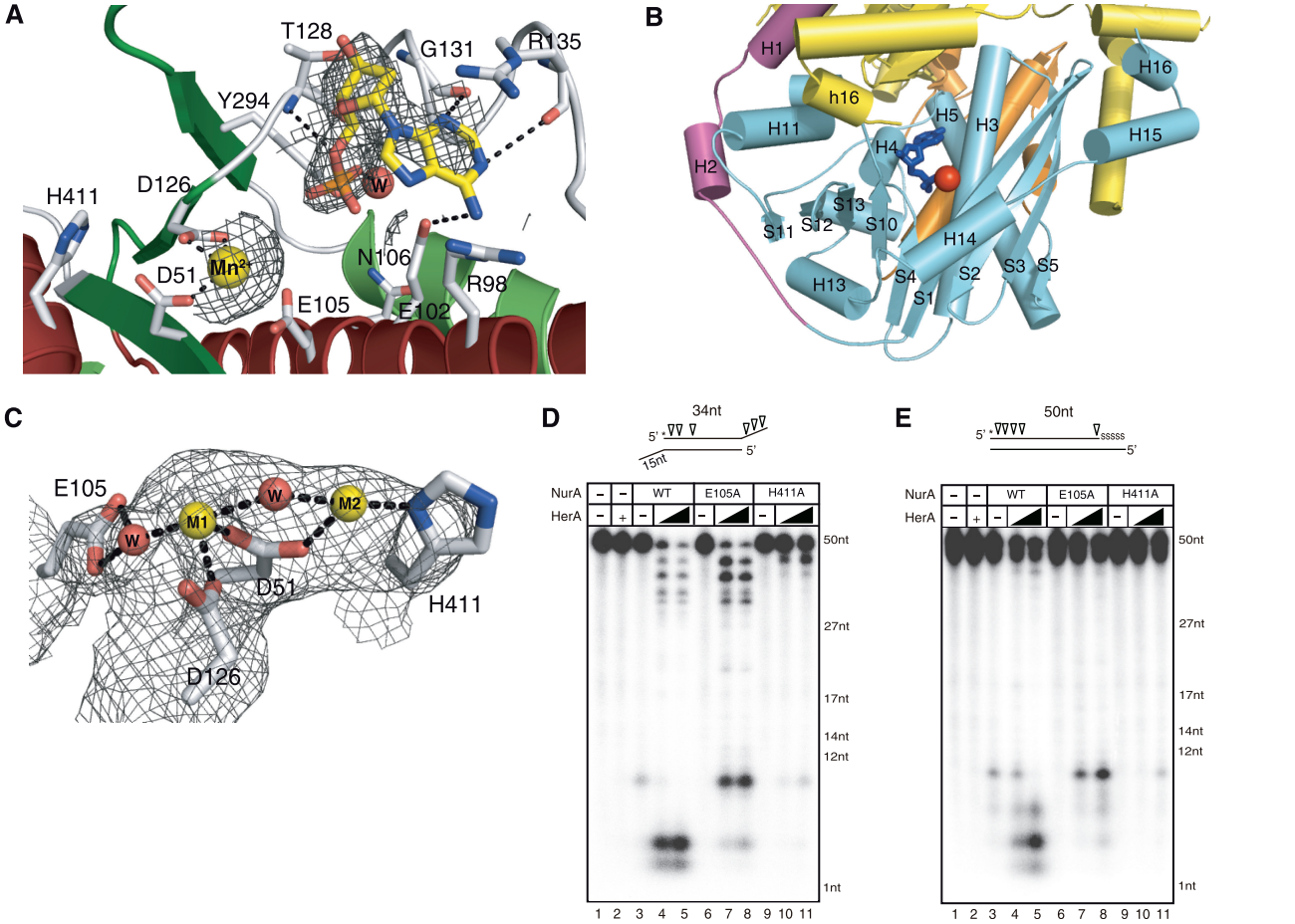

Figure 4. Active site of *Pf*NurA. (**A**) $2F_O$–F_C electron density map (1.2σ) showing dAMP, Mn^{2+} and the interacting residues. H-bonds and ion pairs are shown as dotted lines. (**B**) An active site cleft. An active site cleft with a Mn^{2+} ion (red sphere) and dAMP (blue stick). The C-terminal end of another *Pf*NurA monomer (yellow) is directed toward the active site cleft of the *Pf*NurA monomer. The three domains are colored as in Figure 2B. (**C**) $2F_O$–F_C electron density map (1.5σ) of two Mn^{2+} ions. Active site residues and water molecules are shown. H-bonds and ion pairs are shown as dotted lines. (**D**) Nuclease activity analysis of the two active site mutant proteins. TP 424/423 or (**E**) TP 580/124 was used as the substrate. Assay conditions were the same as those used in Figure 1B.

The C-terminus (yellow) of another subunit, formed with several highly conserved residues, is directed toward the active site cleft of the *Pf* NurA subunit (Figure 4B).

We observed another strong density near the metal-binding site where we added dAMP. The density was clear in one subunit (subunit B), while it is ambiguous in another subunit and we did not build dAMP. A phosphate group is directed to helix H3 and strand S4 and the base is exposed to the surface of the open channel, interacting with helix H4. The base interacts with the guanidinium group of Gly131(O) and Arg135(O) and is close to Arg98 and Glu102. The ribose ring is packed against the ring of Tyr90. A phosphate is located near Glu102, Glu105 and Asn106 on one side, the backbone of Asp126–Thr128, and the side chain of Tyr294. A Mn^{2+} ion is 6.5 Å separated from the phosphate of dAMP. The 3′ OH group points toward the metal and the entrance of the channel. Highly, but not absolutely, conserved Arg residues are lined along the open channel to the narrow end. It is likely that these Arg residues coordinate with phosphate backbones of a DNA substrate (see below).

To examine the possibility that additional Mn^{2+}-binding site may be present at the active site, we added an excess amount (100 mM) of $MnCl_2$. In addition to the Mn^{2+} I site (M1), we did observe additional density (at the opposite side of the dAMP-binding site from the M1 site), raising the possibility that a second Mn^{2+} ion may bind to the active site (Figure 4C). The added Mn^{2+} (M2) coordinated with His411 and Asp51, and a water molecule which also interacted with M1. The distance between the two Mn^{2+} ions is 4.2 Å. It is possible that the water molecule that interacts with both M1 and M2 ions functions as a nucleophile for the nuclease reaction.

Mutational analysis of the active site

The Asp51 mutation (or corresponding residues from *Sto*NurA) completely abolished nuclease activity, consistent with structural observations (15,30). We also observed that mutation of Glu105 to Ala abolished *Pf* NurA endonuclease activity on two different substrates (TP424/423 and TP580/124), suggesting that this residue is important for catalysis (Figure 4D and E, lane 6). Additional

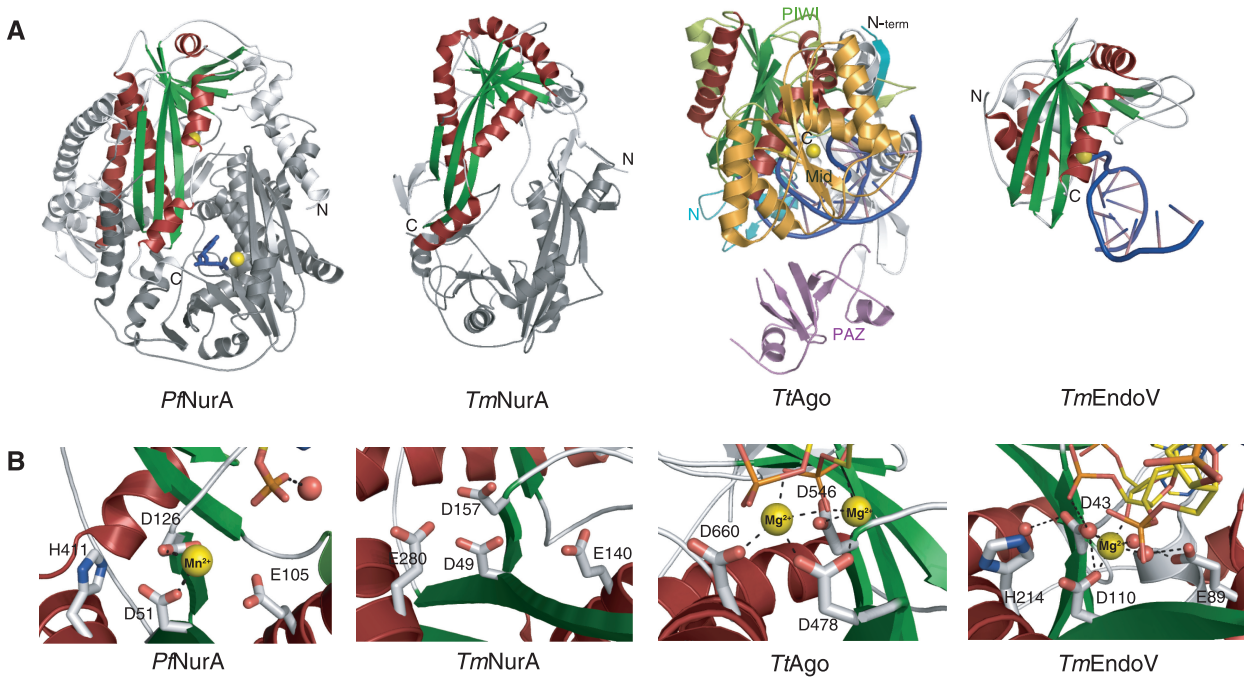

Figure 5. Structural comparison of RNaseH-like fold proteins. (**A**) Comparison of the *Pf*NurA, *Tm*NurA (1zup), *Tt*Ago (3hvr) and *Tm*Endo (2w35) structures. Conserved α-helices (red) and β-sheets (green) form a PIWI domain. *Pf*NurA and *Tm*NurA show a dimeric structure, whose interfaces significantly differ from each other. The *Tm*NurA structure does not contain any metal ions in its active site. The monomer structures of *Tt*Ago and *Tm*EndoV are shown and their metal ions are denoted as yellow spheres. DNA is shown in blue. (**B**) Comparison of active site in *Pf*NurA, *Tm*NurA, *Tt*Ago and *Tm*EndoV. His411 of *Pf*NurA is only conserved in *Tm*EndoV and replaced to Asp in other structures. Oxygen and nitrogen atoms are shown in red and blue, respectively. A water molecule is in red sphere.

mutation studies showed that His411 is also important for nuclease activity, as the His411 mutation or equivalent residue abrogated *Pf*NurA and *Sto*NurA nuclease activity (Figure 4D, lanes 10 and 11; 30). The H411A *Pf*NurA mutant exhibited weak nuclease activity in the presence of *Pf*HerA (Figure 4D and E, lane 11). Since this residue is exposed, its effect on nuclease activity is unlikely due to the structural perturbation. The *Tm*NurA structure does not contain any metal ion in its active site. Three residues (Asp49, Asp157 and Glu140) are conserved, His411 is not present in *Tm*NurA, and this residue is replaced with an acidic residue (Glu280) in *Tm*NurA (Figure 5A, B and Supplementary Figure S2).

Interestingly, the E105A *Pf*NurA mutant showed efficient 3′ exo- and 5′-endonuclase activities (10-nt product) in the presence of *Pf*HerA (Figure 4D, lanes 7 and 8). However, this mutant showed very weak 5′ exonuclease activity. Efficient 5′-endonuclease activity of the E105A *Pf*NurA mutant in the presence of *Pf*HerA was further confirmed with the TP580/124 substrate (Figure 4E, lanes 7 and 8). Taken together, our mutational analysis demonstrates the significance of Glu105 and His411 of *Pf*NurA for nuclease activity, and that *Pf*NurA nuclease activity can be changed in the presence of *Pf*HerA.

Structural comparison with *Tm*NurA

The crystal structure of the *Tm*NurA homolog (PDB ID; 1zup) has been deposited in the PDB data base (unpublished). *Tm*NurA consists of 330 residues which is significantly shorter than that of *Pf*NurA (451 residues), and shares 12% sequence identity. *Tm*NurA also forms a dimer and the r.m.s.d. between *Pf*NurA and *Tm*NurA is 3.3 Å for 183 Cα atoms. Although the overall core structure containing the eight stranded-sheet (S1–S3, S4, S10–S13, green) and flanking helices (helices H3, H13–H15, red) in the PIWI domain is similar between *Pf*NurA and *Tm*NurA, significant differences are observed in the dimeric architecture and dimeric interface.

First, in contrast to *Pf*NurA, the two N-terminal strands of *Tm*NurA protrude into the central channel and divide the channel (Figure 5A). Second, the dimeric interface of *Tm*NurA is not as tight as that of *Pf*NurA: the dimeric interface of *Tm*NurA is formed through the interaction between N-terminal S1 and s2, between loops S1–S2 and s1–s2, and between loops S1–S2 and s2–s3. Third, although both N- and C-terminal regions are involved in dimerization in *Pf*NurA and *Tm*NurA, the contents and modes of interactions between these domains are totally different: the N-terminal helix H1 of *Pf*NurA is not present in *Tm*NurA, suggesting that the dimeric interaction of *Tm*NurA might be less tight than that of *Pf*NurA. In addition, the M domain containing several helices (H4 and H6–H8) and strands (S5–S9) of *Pf*NurA (residues 130–278) is not present in *Tm*NurA. Consequently, the central channel in the *Tm*NurA dimer is wide-open at both the front and back sides, whereas the channel becomes narrower toward one side, and is partly blocked by helices H7 and H8 in *Pf*NurA (Figure 5A).

Structural comparison with other nucleases

When compared with its functional homolog, human Exo1, strands S10–S13 of *Pf*NurA can be superimposed with the corresponding part of the structure, whereas other regions are completely different (34). The PDB data base search using the DALI server revealed that *Pf*NurA is most similar to the structures of argonaute (3hvr, z score 8.1, rmsd 3.5 Å for 182 Cα; 35) and endoV nuclease (2w35, 7.7, 3.4 Å for 152 Cα; 36), in addition to the *Tm*NurA homolog (10.1, 3.8 Å for 194 Cα). The core of these structures shares high similarity, with the RNase HII fold, in which the five strands (3, 2, 1, 4, 10) are packed against helix H14 on one face and helices H3, H4 and H9 on another face (35, 36, Supplementary Figure S5). However, large differences are observed in the overall monomeric structure and dimeric arrangement.

Figure 5A shows the structural comparison between *Pf*NurA and the proteins with structural homology. The red colored helices and green colored strands of *Pf*NurA denote high similarity with the PIWI domain of *Thermus thermophilus* Ago (*Tt*Ago; 35). However, the lengths of several secondary elements (S1, S2, S3, H3 and loops S1–S2 and S3–H3) are much longer than the equivalent parts of *Tt*Ago. Unlike the PIWI domain, an equivalent region to the Mid domain (orange, residues 328–546) of *Tt*Ago is not present in the *Pf*NurA dimer. The PAZ domain (magenta) of *Tt*Ago can overlay portions including strands S1, S2, S3, helix H4 and loop H13–H14 from another *Pf*NurA subunit. Furthermore, the N-domain of *Tt*Ago (blue) can be superimposed on a part of the PIWI domain and parts of the M domain (residues 144–154, 215–224 and 252–261) of *Pf*NurA.

EndoV is present as a monomer; thus, DNA binds to the more open interface of Endo V compared to *Tt*Ago or *Pf*NurA (37). Here, a central β-sheet comprising S1, S2, S3, S4, S10, S12 and S13 (green), which is flanked by helices H13 and H14, and H3, H9 and H12 (red), which shares structural conservation with *Pf*NurA.

The two Asp (Asp51 and Asp126) are conserved in RNase H, Ago, endoV and *Tm*NurA (Figure 5B). The Glu residue (Glu105) is also conserved in all RNase H and NurA family members. In *Tt*Ago, Asp660 (His411 of *Pf*NurA) coordinated the Mg^{2+} ion; thus, three catalytic acidic residues (Asp478, Asp546 and Asp660) coordinated the two Mg^{2+} ions. It has been suggested that an OH group bound to the Mg^{2+} ion in *Tt*Ago, which is bound to Asp660, plays a role as a nucleophile (38). *Tm*EndoV exhibits high similarity to the active site of *Pf*NurA–dAMP–Mn^{2+} as it contains one Mg^{2+} ion and a conserved His, two Asp and a Glu residue in its active site (Figure 5B).

*Pf*NurA–DNA model

Based on the *Pf*NurA–dAMP–Mn^{2+}, *Tt*Ago–DNA and EndoV–DNA crystal structures, we modeled both dsDNA and ssDNA on an active site of *Pf*NurA (Figure 6A and Supplementary Figure S5). DNA primarily binds to PIWI and PAZ in *Tt*Ago, and the 3′-end is recognized by the N domain and the 5′-end is recognized by the Mid domain (38). When the PIWI domain of *Tt*Ago is superimposed on the equivalent domain of *Pf*NurA, dsDNA bound to a sheet comprised of strands (S3, S2, S1, S4, S10, H13, S12 and S11), and regions, including helices H11 and H12 and strands S11 and S12 from another *Pf*NurA, encircled the DNA completely. However, helix H3 from one monomer and strands S1–S3 from another *Pf*NurA collided with the modeled dsDNA, implying that a conformational change in this region or dsDNA is necessary to accommodate the dsDNA (Supplementary Figure S5).

In contrast, ssDNA, metal and large parts of the *Tt*Ago structure including PIWI could be readily superimposed with the metal, dAMP and *Pf*NurA PIWI domain (Figure 6A). When PIWI from *Tt*Ago–DNA or EndoV–DNA is superimposed on *Pf*NurA, ssDNA fits nicely into a cleft formed by strand S1, helix H4 and loops S4–H4, and S10–H10 from one *Pf*NurA, and strand S1 and loop H14–H15 from another *Pf*NurA. The modeled ssDNA threads through the channel from the 3′–5′ direction, where the 3′ OH of a ribose ring is directed to the open channel and the 5′ phosphate group is directed to the narrow part of the channel (Figure 6A, box). From the Mn^{2+} ion(s) to the narrow end of the channel (H7 and H8) is ∼50 Å in which ∼10–11 nt can be located (Figure 2E). Several conserved Arg and Lys residues (Arg54, Arg58, Arg98, Arg135, Lys297, Arg323, Lys415, Lys419 and Arg422) are positioned along the phosphate group of the modeled DNA and are likely to stabilize the substrate or reaction intermediate.

DNA binds to the surface of the open channel

Based on the structural comparison of the PIWI domains between *Pf*NurA and *Tt*Ago, and the information of the metal and dAMP-binding site of *Pf*NurA, it is likely that DNA binds to the groove formed by strand S1, helix H4 and loop S4–H4, S10–H10 and H14–H15 of *Pf*NurA. These regions are rich in positively charged residues and ring-containing residues, which could interact with a DNA substrate, including Lys297, Arg323, Tyr380, Tyr403, Lys415, Lys419 and Arg435 (Figure 6C).

To examine if these residues affected DNA binding and cleavage by *Pf*NurA, charge-inversion mutant proteins were constructed and their nuclease activities were assessed. We made three double- or triple-mutant proteins. These mutant proteins included Lys297Glu–Tyr380Phe–Tyr403Phe, Arg323Glu–Arg435Glu and Lys415Glu–Lys419Glu. In the absence of *Pf*HerA, the Lys297Glu–Tyr380Phe–Tyr403Phe and Lys415Glu–Lys419Glu mutants failed to show any nuclease activities on two different DNA substrates, whereas Arg323Glu–Arg435Glu showed similar (or higher) nuclease activity compared to that in wild-type *Pf*NurA, despite this mutant exhibited slight change in circular dichroism (CD) spectra near 210 nm (Figure 6C–E, lanes 3, 6, 9, 12 and Supplementary Figure S6). However, the Lys415Glu–Lys419Glu mutant showed strong nuclease activity in the presence of *Pf*HerA, which was comparable to wild-type *Pf*NurA, and only the Lys297Glu–Tyr380Phe–Tyr403Phe mutant failed to show any nuclease activity (Figure 6C–E). The CD analysis revealed that the

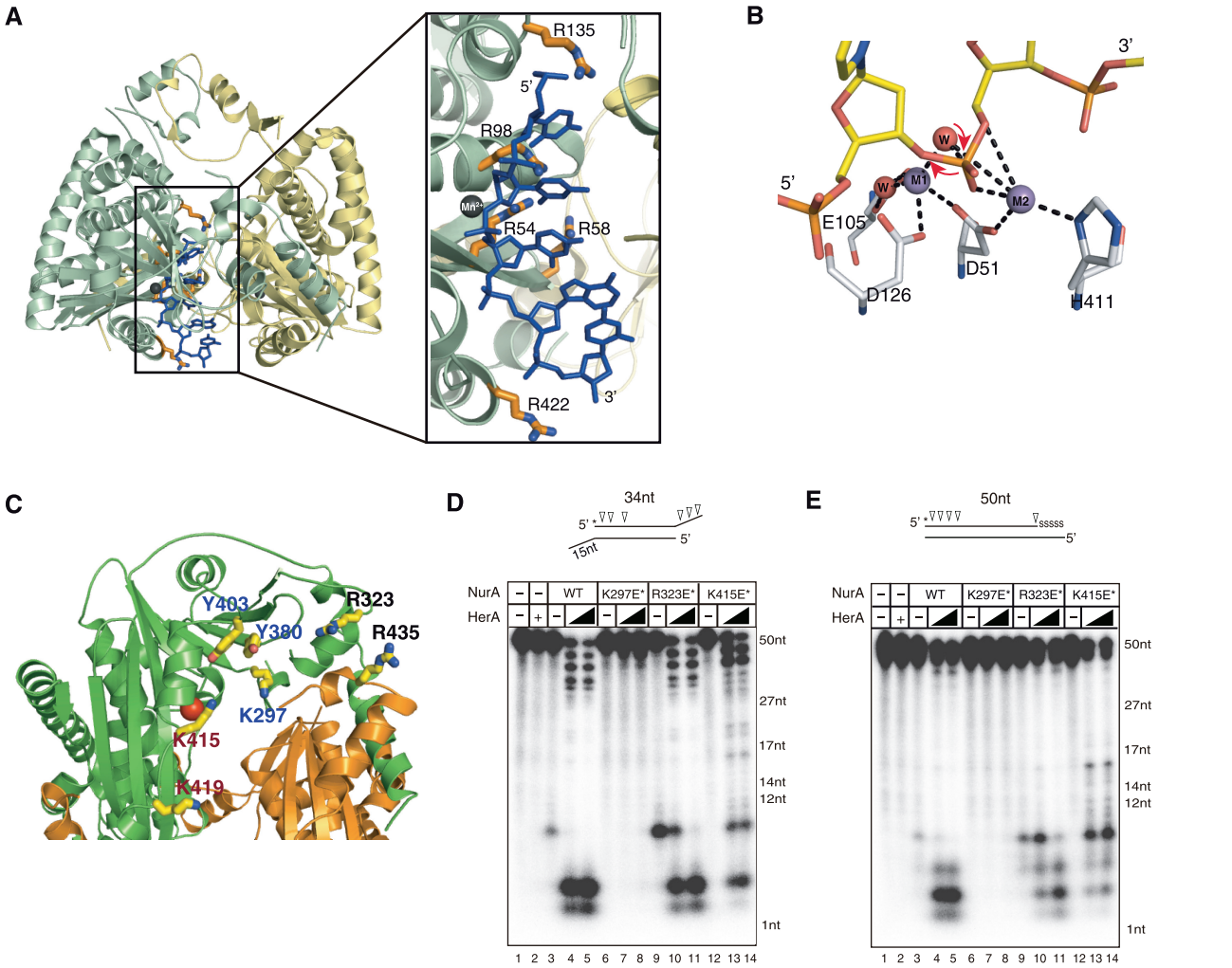

Figure 6. *Pf*NurA–ssDNA-binding model. (**A**) A ssDNA (blue) is modeled on *Pf*NurA based on the *Tt*Ago–DNA and *Pf*NurA–dAMP–Mn^{2+} structures. Box: a close up view of the model showing the direction of the ssDNA, where the 3′ OH of a ribose ring points toward the open channel (entrance) and the 5′-phosphate group is directed to the narrow part of the channel. (**B**) Schematic drawing of the *Pf*NurA nuclease mechanism. Mn^{2+} (M1) is coordinated by Asp51, Asp126 and two water molecules. Mn^{2+} (M2) interacts with Asp51, His411 and one water molecule. In this model, a water nucleophile (red) bound to both metal ions attacks the phosphate group and M1 stabilizes the leaving 3′ oxyanion, which allows the 5′–3′ digestion.(**C**) Structure of the *Pf*NurA showing the position of some residues on the surface of the elliptical channel. Residues containing K297E/Y380F/Y403F (blue label), R323E/R435E (black label) and K415E/K419E (red label) are shown. (**D**) Nuclease activity assay of the three DNA-binding mutant using TP 424/423 or (**E**) TP 580/124. Each DNA substrate (20 nM) was incubated for 120 min at 65°C with 350 nM of *Pf*NurA, increasing amounts of *Pf*HerA (35 and 70 nM), 5 mM $MnCl_2$ and 1 mM ATP. In lane 2, 70 nM of *Pf*HerA was used as a control. Abbreviations: K297E*, K297E/Y380F/Y403F; R323E*, R323E/R435E; K415E*, K415E/K419E.

Lys297Glu–Tyr380Phe–Tyr403Phe triple mutation did not alter the overall *Pf*NurA conformation (Supplementary Figure S6). The mutational analysis suggested that Lys297, Tyr380 and/or Tyr403 on the flat face of the *Pf*NurA dimer are involved in DNA binding. In addition, Lys415 and Lys419 may also participate in DNA substrate binding.

DISCUSSION

In this study, we determined the crystal structure of *Pf*NurA and identified substrate binding sites to reveal its nuclease mechanism. NurA is a conserved nuclease in archaea and has been proposed as a functional homolog of eukaryotic Exo1 nuclease (15). Although it is only present in archaea, it has drawn substantial interest because (i) it has no reported structural *Pf*NurA homologs and exhibits weak sequence similarities between paralogs and within orthologs, (ii) it is likely to play an important role in archaeal HR, and (iii) it specifically processes DNA in the 5′ to 3′ direction. We have provided important insights into these questions in the present study. We have provided important insights on these questions in the present study. First, we have shown that *Pf*NurA contains a PIWI domain with an RNase HII fold and forms a pyramid-shape dimer with a central channel that can be used for both endo- and exonucleolytic substrate cleavage. Second, we have shown from the *Pf*NurA–dAMP–Mn^{2+} structural studies that *Pf*NurA resembles argonaute and provides the basis of the 5′–3′ nuclease activity of NurA.

HerA alters the *Pf*NurA product

We have shown that *Pf*NurA together with *Pf*HerA generates 2- to 4-nt products from both ends of a substrate, displaying both 5′ and 3′ exonuclease activities, whereas *Pf*NurA alone produced 6–10 nt from the 5′-end of substrates in an endonucleolytic manner (Figure 1A and 1B). Although previous studies clearly showed that *Pf*HerA facilitates *Pf*NurA nuclease activities in the presence of Mn^{2+} or Mg^{2+} ions, no detailed biochemical characterizations of the *Pf*NurA–*Pf*HerA complex have been described (15). Previous studies showed that the *Pf*NurA–*Pf*HerA complex exhibits both exo- and endonuclease activities on a 2.5-kb DNA substrate (15, Paull, personal communication). Although the same enzymes were used, we note some differences between current experiments and previous analyses by Hopkins and Paull (15). First, we used a 17.5:1 ratio of *Pf*NurA (350 nM) and the DNA substrate (20 nM), whereas a 1600:1 ratio of enzyme (96.2 nM) to substrate (0.06 nM) was used in previous assays (15). Second, Hopkins and Paull used a 2.5-kb substrate, whereas we used short-length DNA (50 bp) substrates. It is likely that *Pf*HerA, a helicase that unwinds DNA in both directions, melts the DNA substrates to make them more susceptible to the *Pf*NurA dimer, which possesses two active sites and alters the generation of product.

*Pf*NurA dimer is formed by two interfaces

The dimeric form of *Pf*NurA appears to be essential for its nuclease activity (Figures 1A and 3A–E). *Pf*NurA forms a dimer through two interfaces: one is formed by the N- and PIWI domains and another by the M domain (Figure 2B and C). Our mutational analysis revealed that the N-terminal H1 helix is critical for dimer formation. Interestingly, when we removed helix H1 or mutated three residues (Arg11, Ile12 and Ser60), the *Pf*NurA dimer did not dissociate into a monomer. Instead, it assembled into a tetramer or formed a less compact dimer compared to that in wild-type *Pf*NurA (Figure 3E). These data suggest that the structural change in the N-terminal H1 helix alters the dimeric arrangement or overall *Pf*NurA structure.

Sequence alignment of archaeal NurA members divides NurA into those with or without the M domain (residues 190–286). NurA proteins from *Pyrococcus* and *Methanococcus* species contain the M domain, whereas those from *Sulfolobous* and *Thermotoga* do not possess the corresponding region (Supplementary Figure S2). The M domain plays an important role in *Pf*NurA dimerization (Figure 2B and C). Furthermore, it is located on one face of the channel (rear side), which may imply an additional role for this region such as regulation of the release of a product(s) and/or to provide additional binding surface to other partners of *Pf*NurA.

Interestingly, a dimeric interface formed by helices H1 and H2 from a *Pf*NurA subunit and strand s3, helices h6 and h15 from another *Pf*NurA is significantly different from that of *Tm*NurA (Figure 5A and B). In the *Tm*NurA dimer interface, the N-terminal S1 and S2 strands, and loops S1–S2 and S2–S3 of a subunit directly interact with its equivalent region from another *Tm*NurA subunit. The structural divergence of the dimeric interface in NurA members may be due to the fact that NurA has very weak sequence similarity among homologous proteins. Further structural studies from another archaeal NurA are essential to understand conservation of the dimeric interface in NurA members at the structural level.

Interaction between HerA and NurA

Although NurA can function independently as a nuclease, it coordinates with HerA to generate both 5′ and 3′ ssDNA molecules; at least *Pf*HerA and *Pf*NurA interact stably enough to maintain their interactions during gel filtration (15).

The structures of *Pf*NurA and MlaA/HerA determined by electron microscopy (EM) provide insights into understanding how these proteins function together (18). The flat face of the *Pf*NurA dimer contains an open channel (29×19 Å) that becomes narrower toward the opposite side (Figure 2E). EM studies of MlaA/HerA revealed that it forms a hexamer with a large hole of ~30 Å diameter, which matches the size of the channel at the flat face of *Pf*NurA (18). Furthermore, the metal-binding active site is located near this flat face of *Pf*NurA. Thus, we propose that the flat face of *Pf*NurA is a binding region for *Pf*HerA.

The *Pf*HerA hexamer interacts with one to three *Pf*NurA molecules (15). Based on our structural analysis, we propose that the stoichimetry of *Pf*HerA and *Pf*NurA is 3:1, which is consistent with the data from the aforementioned study (Figure 7). Furthermore, the surface (helices H14–H16) in which *Pf*HerA is proposed to bind contains several conserved residues in NurA family members. These include Pro406, Leu407, Ala410, Ile416, Ser417 and Ala427–Asn430. Taken together, we suggest that a HerA hexamer and a NurA dimer bind on the surface of the elliptical channel (Figure 7).

The nuclease activity analysis revealed that the presence of *Pf*HerA altered the *Pf*NurA product, from 10 to 2 nt (Figure 1B). In the presence of *Pf*HerA, which unwinds dsDNA in both directions, the ssDNA molecule is expected to enter the front channel of *Pf*NurA. As we have shown in our model, dsDNA did not fit into the *Pf*NurA dimer and required a conformational change of a protein, whereas ssDNA which is generated by *Pf*HerA can thread into the central channel of *Pf*NurA. However, we do not completely exclude the possibility that *Pf*HerA induces the conformational change of *Pf*NurA for efficient binding of ssDNA with the 2-nt product generation.

DNA-binding model and nuclease mechanism of NurA

The mutational analysis revealed that replacing residues on the open elliptical channel inhibited *Pf*NurA nuclease activity, and that the size of this channel matched with the hole in MlaA/HerA (Figure 7; 18). Thus, we speculate that the DNA molecule passes the wide-open face of the channel and is directed toward the relatively narrow end of the *Pf*NurA dimer through the metal-binding site.

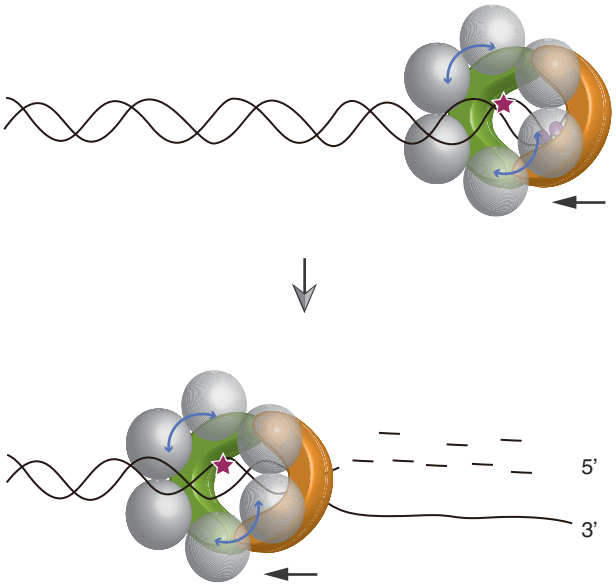

Figure 7. Schematic model of NurA-HerA for 5′–3′ processing. A hexameric *Pf*HerA (grey) and the 'flat' face of the *Pf*NurA dimer (orange and green) interact and function together to generate 5′ resected DNA. In the presence of *Pf*HerA which unwinds dsDNA, the ssDNA molecule is expected to enter the front channel of *Pf*NurA to produce 3′ ssDNA for the RadA binding. An active site of *Pf*NurA is shown as a star.

It is unclear if only one or two divalent metal ions are involved in the cleavage process by *Pf*NurA. Although we observed only one Mn^{2+} ion (M1) in the *Pf*NurA–dAMP–Mn^{2+} structure, adding an excess amount of $MnCl_2$ clearly revealed additional possible Mn^{2+} binding (M2) at an active site, and a catalytically important residue, His411, was coordinated with this Mn^{2+} (M2) ion. The importance of His411 has been demonstrated by *Sto*NurA (30) and the present *Pf*NurA results (Figure 4D–E). When superimposed with *Tt*Ago, the two metal sites of *Pf*NurA positioned well to those of the Mg^{2+} ion site in *Tt*Ago. The relatively high B value of this Mn^{2+} (M2) ion (B: 131.6 Å^2, average protein B: 115.7 Å^2) suggests that this metal weakly binds to the active site (through Asp51 and His411). A water molecule interacts with both metal ions, although this water and the M2 ion are more distantly located (3.0 Å). We presume that the M2 metal positions and activates this water molecule for nucleophilic attack on the phosphate group and M1 metal stabilizes the leaving 3′ oxyanion (Figure 6B). While the 5′ phosphate group is directed to the narrow end of the channel, the 3′ OH group points toward the open elliptical channel. This catalytic reaction may allow 5′–3′ processing of the nuclease reaction.

The distance between the putative catalytic water to end of the channel (H7 and H8) is about 50 Å, which is the length of 10–11 nt (Figure 2E). This length is consistent with the size of the dsDNA cleavage product by *Pf*NurA (Figure 1A and B). It is possible that helices H7 and H8 interact with dsDNA and stabilize substrate binding for efficient cleavage.

It is unresolved at this stage why two active sites are present in the *Pf*NurA dimer. For example, whether or not both active sites function equally is unknown. Interestingly, in the presence of *Pf*HerA, we observed 3′-end cleavage products as well as 5′-end cleavage products (Figure 1B). Thus, it is possible that both active sites recognize each dsDNA strand (one for the 5′–3′ direction and another for the 3′–5′ direction). Since HerA has no preference in directionality, it could generate both 3′ and 5′ ssDNA ends. When the two unwound ssDNAs enter the channel of the *Pf*NurA dimer, one strand with the 3′-end may bind the active site, whereas another strand with the 5′-end could interact with another active site, explaining the cleavage reaction in both directions (Figure 7). Another remaining question is how *Pf*NurA displayed 3′–5′ exonuclease activity in the presence of *Pf*HerA. Since our structure only provides the basis for 5′–3′ nuclease activity, further structural analyses using various substrates or ligands are required to resolve this question.

In summary, we initiated this work to understand if archaeal NurA is structurally conserved relative to its eukaryotic counterpart Exo1. Interestingly, *Pf*NurA does not share structural similarity with Exo1. Instead, it resembles bacterial Ago, a microRNA processing enzyme, and EndoV, both in structural (including active site) and functional aspects. The *Pf*NurA–dAMP–Mn^{2+} complex structure in conjunction with the *Tt*Ago–DNA model and biochemical data clearly provide insight into 5′–3′ nuclease process. Further structural and functional analyses on the NurA–DNA complex should help to understand the role of this enzyme both *in vivo* and *in vitro*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1, Supplementary Figures 1–6 and Supplementary Methods.

ACKNOWLEDGEMENTS

The authors thank to Young Bong Park and Gwang Hyun Gwon for technical help and discussion. Coordinates and structure factors have been deposited with RCSB data base with 3TAI (apo *Pf*NurA), 3TAL (*Pf*NurA-Mn^{2+}) and 3TAZ (*Pf*NurA-dAMP-Mn^{2+}). Diffraction data from apo- and Mn^{2+}-bound crystals were collected on beamline 4A at the Pohang Advanced Light Source (PAL) and Argonne National Laboratory (APS).

FUNDING

Funding for open access charge: National R&D Program for Cancer Control, Ministry for Health and Welfare (1020280); National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2010-0019706 and No. 2010-0029766); a rising star program (POSTECH) and the BK21 program (Ministry of Education).

Conflict of interest statement. None declared.

REFERENCES

1. Khanna,K.K. and Jackson,S.P. (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nat. Genet.*, **27**, 247–254.
2. Kowalczykowski,S.C., Dixon,D.A., Eggleston,A.K., Lauder,S.D. and Rehrauer,W.M. (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.*, **58**, 401–465.
3. Dillingham,M.S. and Kowalczykowski,S.C. (2008) RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol. Mol. Biol. Rev.*, **72**, 642–671.
4. Stracker,T.H. and Petrini,J.H. (2011) The MRE11 complex: starting from the ends. *Nat Rev Mol. Cell. Biol.*, **12**, 90–103.
5. Lengsfeld,B.M., Rattray,A.J., Bhaskara,V., Ghirlando,R. and Paull,T.T. (2007) Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex. *Mol. Cell*, **28**, 638–651.
6. Nicolette,M.L., Lee,K., Guo,Z., Rani,M., Chow,J.M., Lee,S.E. and Paull,T.T. (2010) Mre11-Rad50-Xrs2 and Sae2 promote 5' strand resection of DNA double-strand breaks. *Nat. Struct. Mol. Biol.*, **17**, 1478–1485.
7. Mimitou,E.P. and Symington,L.S. (2008) Sae2/CtIP, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature*, **455**, 770–774.
8. Cejka,P., Cannavo,E., Polaczek,P., Masuda-Sasa,T., Pokharel,S., Campbell,J.L. and Kowalczykowski,S.C. (2010) DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. *Nature*, **467**, 112–116.
9. Niu,H., Chung,W.H., Zhu,Z., Kwon,Y., Zhao,W., Chi,P., Prakash,R., Seong,C., Liu,D., Lu,L. *et al.* (2010) Mechanism of the ATP-dependent DNA end resection machinery from Saccharomyces cerevisiae. *Nature*, **467**, 108–111.
10. Tran,P.T., Erdeniz,N., Symington,L.S. and Liskay,R.M. (2004) EXO1-a multi-tasking eukaryotic nuclease. *DNA Repair*, **3**, 1549–1559.
11. Mimitou,E.P. and Symington,L.S. (2011) DNA end resection-unraveling the tail. *DNA Repair*, **10**, 344–348.
12. Connelly,J.C., Kirkham,L.A. and Leach,D.R.F. (1998) The SbcCD nuclease of *Escherichia coli* is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. *Proc. Natl Acad. Sci. USA*, **95**, 7969–7974.
13. Williams,R.S., Williams,J.S. and Tainer,J.A. (2007) Mre11-Rad50-Nbs1 is a keystone complex connecting DNA repair machinery, double-strand break signaling, and the chromatin template. *Biochem. Cell. Biol.*, **85**, 509–520.
14. Lim,H.S., Kim,J.S., Park,Y.B., Gwon,G.H. and Cho,Y. (2011) Crystal structure of the Mre11-Rad50-ATPγS complex: Understanding the interplay between Mre11 and Rad50. *Genes Dev.*, **25**, 1091–1104.
15. Hopkins,B.B. and Paull,T.T. (2008) The P. furiosus Mre11/Rad50 complex promotes 5′ strand resection at a DNA double-strand break. *Cell*, **135**, 250–260.
16. Constantinesco,F., Forterre,P. and Elie,C. (2002) NurA, a novel 5′-3′ nuclease gene linked to rad50 and mre11 homologs of thermophilic Archaea. *EMBO Rep.*, **3**, 537–542.
17. Constantinesco,F., Forterre,P., Koonin,E.V., Aravind,L. and Elie,C. (2004) A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nurA genes in thermophilic archaea. *Nucleic Acids Res.*, **32**, 1439–1447.
18. Manzan,A., Pfeiffer,G., Hefferin,M.L., Lang,C.E., Carney,J.P. and Hopfner,K.P. (2004) MlaA, a hexameric ATPase linked to the Mre11 complex in archaeal genomes. *EMBO Rep.*, **5**, 54–59.
19. Frols,S., Gordon,P.M.K., Panlilio,M.A., Duggin,I.G., Bell,S.D., Sensen,C.W. and Schleper,C. (2007) Response of the hyperthermophilic archaeon Sulfolobus solfactaricus to UV damage. *J. Bacterol.*, **189**, 8709–8718.
20. Rolfsmeier,M.L., Laughery,M.F. and Haseltine,C.A. (2010) Repair of DNA double-strand breaks following UV damage in three Sulfolobus solfataricus strains. *J. Bacterol.*, **192**, 4954–4962.
21. Quaiser,A., Constantinesco,F., White,M.F., Forterre,P. and Elie,C. (2008) The Mre11 protein interacts with both Rad50 and the HerA bipolar helicase and is recruited to DNA following gamma irradiation in the archaeon *Sulfolobus acidocaldarius*. *BMC Mol. Biol.*, **9**, 25.
22. Zhang,S., Wei,T., Hou,G., Zhang,C., Liang,P., Ni,J., Sheng,D. and Shen,Y. (2008) Archaeal DNA helicase HerA interacts with Mre11 homologue and unwinds blunt-ended double-stranded DNA and recombination intermediates. *DNA Repair*, **7**, 380–391.
23. Wei,T., Zhang,S., Zhu,S., Sheng,D., Ni,J. and Shen,Y. (2008) Physical and functional interaction between archaeal single-stranded DNA-binding protein and the 5′–3′ nuclease NurA. *Biochem. Biophys. Res. Commun.*, **369**, 523–529.
24. Fujikane,R., Ishino,S., Ishino,Y. and Forterre,P. (2010) Genetic analysis of DNA repair in the hyperthermophilic archaeon, *Thermococcus kodakaraensis*. *Genes Genet. Sys.*, **85**, 243–257.
25. Lyer,L.M., Makarova,K.S., Koonin,E.V. and Aravind,L. (2004) Comparative genomics of the FtsK–HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. *Nucleic Acids Res.*, **32**, 5260–5279.
26. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
27. Adams,P.D., Afonine,P.V., Bunkóczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D*, **66**, 213–221.
28. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D*, **60**, 2126–2132.
29. Brünger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D*, **54**, 905–921.
30. Wei,T., Zhang,S., Hou,L., Ni,J., Sheng,D. and Shen,Y. (2011) The carboxyl terminal of the archaeal nuclease NurA is involved in the interaction with single-stranded DN A-binding protein and dimer formation. *Extremophiles*, **15**, 227–234.
31. Ma,J.B., Yuan,Y.R., Meister,G., Pei,Y., Tuschl,T. and Patel,D.J. (2005) Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature*, **434**, 666–670.
32. Parker,J.S., Roe,S.M. and Barford,D. (2004) Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO J.*, **23**, 4727–4737.
33. Song,J.J., Smith,S.K., Hannon,G.J. and Joshua-Tor,L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science*, **305**, 1434–1437.
34. Orans,J., McSweeney,E.A., Iyer,R.R., Hast,M.A., Hellinga,H.W., Modrich,P. and Beese,L.S. (2011) Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell*, **145**, 212–223.
35. Lai,L., Yokota,H., Hung,L.W., Kim,R. and Kim,S.H. (2000) Crystal structure of archaeal RNase HII: a homologue of human major RNase H. *Struct. Fold. Des.*, **8**, 897–904.
36. Yang,W. and Steitz,T.A. (1995) Recombining the structures of HIV integrase, RuvC and RNase H. *Structure*, **3**, 131–134.
37. Dalhus,B., Arvai,A.S., Rosnes,I., Olsen,O.E., Backe,P.H., Alseth,I., Gao,H., Cao,W., Tainer,J.A. and Bjoras,M. (2009) Structures of endonuclease V with DNA reveal initiation of deaminated adenine repair. *Nat. Struct. Mol. Biol.*, **16**, 138–143.
38. Wang,Y., Juranek,S., Li,H., Sheng,G., Wardle,G.S., Tuschl,T. and Patel,D.J. (2009) Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature*, **461**, 754–761.