Published in final edited form as:
Proteins. 2015 January ; 83(1): 188–194. doi:10.1002/prot.24704.

Structure and Specificity of FEN-1 from *Methanopyrus kandleri*

Santosh Shah[a,c], **Pete Dunten**[b], **Amanda Stiteler**[a], **Chad K. Park**[a], and **Nancy C. Horton**[a,*]

[a]Department of Chemistry and Biochemistry, University of Arizona, Tucson, Arizona, 85721, USA

[b]Stanford Synchrotron Radiation Lightsource, Stanford University, Menlo Park, CA 94025, USA

Abstract

DNA repair is fundamental to genome stability and is found in all three domains of life. However, many archaeal species, such as *Methanopyrus kandleri*, contain only a subset of the eukaryotic nucleotide excision repair (NER) homologues, and those present often contain significant differences compared to their eukaryotic homologues. To clarify the role of the NER XPG-like protein Mk0566 from *M. kandleri*, its biochemical activity and three dimensional structure were investigated. Both were found to be more similar to human FEN-1 than human XPG, suggesting a biological role in replication and long-patch base excision repair rather than in NER.

Keywords

protein-DNA complex; DNA nuclease; nucleotide excision repair

INTRODUCTION

DNA nucleases act in various biological processes including replication, recombination, and repair. DNA nucleases have high specificity and efficiency in performing activities such as endonucleolytic cleavage of double or single stranded DNA, and 5′ or 3′ exonucleolytic cleavage. Failure to properly perform these critical functions can lead to diseases such as premature aging, cancer, autoimmune disease, and repeat expansion disease[1].

The FEN-1 superfamily of DNA nucleases are structure specific nucleases that include the enzymes FEN-1 (Flap Endonuclease 1), GEN1 (Gap Endonuclease 1), Exo1 (Exonuclease 1), and XPG. These nucleases act in replication and long-patch base excision DNA repair (FEN-1), resolution of Holliday junction recombination intermediates (GEN1), recombination and repair including mismatch repair, double-strand break repair, and telomere maintenance (ExoI), and nucleotide excision repair (XPG)[2]. They share a common fold and cleave DNA specifically, according to specificities particular to their type. For example, FEN-1 enzymes cleave at the junction of single and double stranded DNA in 5’ flap or double flap (5’ and 3’ flap) substrates, as well as cleaving gapped substrates and/or performing exonucleolytic 5’->3’ cleavage on 5’ends in nicked and gapped DNA substrates. GEN1 enzymes recognize four-way junctions, Exo1 recognizes nicked, gapped, or blunt

*To whom correspondence should be addressed: Prof. Nancy C. Horton, Department of Chemistry, and Biochemistry University of Arizona, Tucson, AZ 85721, Telephone: (520) 626-3828,. nhorton@u.arizona.edu.
[c]Current address: Department of Medicine, University of Alabama, Birmingham, Alabama, 35294, USA

ends of DNA and performs 5'->3' exonuclease activity on the 5'end, and XPG has specificity for bubble substrates[2].

FEN-1 enzymes are conserved throughout the three domains of life, eukarya, bacteria, and archaea, and function in incising 5' flaps formed during Okazaki fragment maturation (during replication), as well as during long-patch base excision repair. The FEN-1 family of enzymes is critical for fundamental biological processes; mutation or deletion in FEN-1 results in repair deficiencies and genomic instability[3].

Archaeal enzymes are of interest to study due to their often closer homology to eukaryotic proteins than bacterial. In addition, the archaeal proteins are often simpler, more stable, and more easily isolated and purified for biochemical and biophysical study. The nucleotide excision repair pathway (NER) is an important DNA repair pathway involved in the excision of DNA strands containing damage from UV or bulky adducts. Mutations in proteins of this pathway cause the diseases xeroderma pigmentosum (XP), Cockayne Syndrome (CS), and Trichothiodystrophy (TTD) in humans[4]. The proteins performing this pathway in eukaryotes and bacteria are nonhomologous, and more than 16 polypeptides are involved in the eukaryotic pathway[4]. The study of this pathway in archaea could potentially simplify the study of NER. Some archaea contain homologues to the bacterial NER proteins, although many do not. Archaeal species without bacterial NER homologues contain some eukaryotic NER homologues, however are missing several key NER proteins, and the homologues that are present are often smaller than their eukaryotic counterparts.

We investigated the genome of *Methanopyrus kandleri* for eukaryotic NER homologues, and asked whether the potential homologue to the eukaryotic NER nuclease XPG, Mk0566, could possibly perform XPG specific DNA cleavage. However, multiple sequence alignment of the coding region indicated a closer relationship to eukaryotic FEN-1 than to XPG. Consistent with the alignment data, we found that the DNA substrate cleavage preference and three dimensional structure of Mk0566 also both show a closer relationship to those of FEN-1 than XPG. These results suggest that Mk0566 is unlikely to function in NER as it is understood in eukaryotes. In addition, these studies add to our growing understanding of FEN-1 enzyme structure and function.

MATERIALS AND METHODS

Protein and DNA preparation

Recombinant Mk0566 was prepared using an *E. coli* expression system and a synthetic codon optimized gene producing a fusion protein with *E. coli* maltose binding protein N-terminal to Mk0566. The protein was purified using amylose chromatography, protease cleavage to release Mk0566, then additional purification using Q-sepharose and butyl chromatography. The DNA oligonucleotides were prepared synthetically, purified using HPLC, and annealed using equimolar concentrations or other as indicated. The DNA constructs used in DNA cleavage assays were designed to anneal as in Figure S1. DNA constructs for crystallization were designed to form a truncated double flap structure, with a 1 nucleotide 3' flap, 2 nucleotide 5'flap, 8 bp downstream duplex and 9 bp upstream duplex. See Supplementary Information for more detail.

DNA Cleavage and Analytical Ultracentrifugation Analysis

See Supplementary Information for Methods.

Crystallization, Data Collection, Structure Solution, Refinement, and Analysis

The hanging drop vapor diffusion method5 was used to screen crystallization conditions. Drops were composed of 1.5 μl Mk0566 protein at varied concentrations mixed with a 1.5 molar excess of DNA and 1.5 μl of the precipitant solution. Crystals were obtained using 15% PEG 4K, 100 mM Tris-HCl (pH 8.5@RT), 150 mM NaCl, 10 mM $CaCl_2$ as the precipitating solution. Crystals reached full size in about 2 weeks at 17°C. The crystals were then exchanged into a cryoprotectant solution (25% PEG 4K, 100 mM Tris-HCl (pH 8.5@RT), 300 mM NaCl, 10 mM $CaCl_2$ and 30% glycerol) and flash-frozen in liquid nitrogen. X-ray diffraction was measured using synchrotron radiation at the Stanford Synchrotron Radiation Lightsource (SSRL) BL9-2. Data collection was performed while maintaining the crystal at 100K. Image processing and data reduction were performed with MOSFLM6 and SCALA7 respectively. The structure was solved by molecular replacement using PHASER8. The MR model was based on structurally aligning 1RXW and 1A76 (FEN-1 from *Archeoglobus fulgidus*9 and Methanococcus jannaschii10, which share 58% and 48% sequence identity with FEN-1 from Methanopyrus kandleri), and then removing those parts of 1RXW where the chain traces differed. The resulting MR model was 1RXW 3-37, 57-80, 129-185, 208-254, 272-324 (216 residues total). The PHASER Z-scores were RFZ 7.6, TFZ 16.8 for the first molecule placed, and RFZ 8.1, TFZ 18.4 for the second. The initial R-factors reported by REFMAC were R-work 49.1% and R-free 49.4%. Refinement was carried out using REFMAC and the model was built using Coot11 and PDB_REDO starting with the 1RXW model. All structure figures were prepared using PyMOL12. Structural comparisons and superpositions were performed with the DALI server13. Sequence alignment and secondary structure depiction of Figure 1B prepared with ESPript14.

RESULTS AND DISCUSSION

DNA Cleavage Specificity of Mk0566

DNA cleavage assays utilizing different flap, overhang, single stranded, double stranded and bubble substrates (Table S1) were performed with purified, recombinant Mk0566. In these assays, different enzyme concentrations (0–2 μM) were incubated with the radiolabeled DNA in the presence of buffer containing 10 mM Mg^{2+} at 57°C for 2 or more hours. The incubation temperature was chosen by testing different temperatures and observing that which gave maximal DNA cleavage. The high temperature optimum is consistent with the source organism, *Methanopyrus kandleri*, a hyperthermophile. This high temperature optimum also indicates that the observed activity does not derive from any contaminating, copurifying *E. coli* FEN-1. In the assays, only a single strand of the particular DNA substrate (which contain 2–4 strands) is radiolabeled to simplify interpretation, resulting in 14 different assays (Table S1). Of the eight different types of DNA substrate, cleavage by Mk0566 was observed in only a subset: substrates 5–6 (5'flap with upstream and downstream double stranded DNA, Fig. S1B), and substrates 9–10 (blocked flap or fork substrate, Fig. S1E)(Table S1). Follow-up assays showed that cleavage observed in

substrates 6 and 10 was occurring on the unlabeled strand (NER2), and altering the mobility of the incompletely denatured DNA in the gels. The fact that several bands are seen for the annealed DNA (Fig. S2) suggests greater than one cleavage event, likely corresponding to 5'->3' exonuclease activity on NER2 following endonucleolytic cleavage, as seen with other FEN-1 enzymes[6]. In addition, follow-up assays with substrate 9 (Table S1, blocked flap or fork, Fig. S1E) in the presence of 50 fold excess NER4 show diminished cleavage activity relative to the 5'flap substrate (compare lane 8 to lane 6, Fig. S3), suggesting either that the duplex nature of the 5' flap reduced cleavage activity by Mk0566, or that activity was due to a small amount of substrate missing the NER4 strand (and hence equivalent to the 5'flap substrate of Figure S1B). Follow-up assays also suggested that the presence of a nucleotide 3' of the junction on the upstream duplex (on NER3, i.e. NER3+1nt, Materials and Methods, creating the double flap substrate, Fig. S1D) enhanced cleavage activity by Mk0566 on 5'flap and blocked flap constructs (compare lanes 10 and 12 to lanes 6 and 8, respectively, Fig. S3), as seen for other FEN-1 nucleases[7]. Finally, no evidence of cleavage of single stranded DNA, fully double stranded DNA (Fig. S1F), double stranded DNA containing a bubble (Fig. S1G), pseudo-Y DNA (Fig. S1A), or 3'flap DNA substrates (Fig. S1C) by Mk0566 was found.

The previously reported preferred substrate of FEN-1 enzymes is a double flap substrate containing duplex regions both upstream and downstream, and both 5' and 3' flaps (1 nucleotide in the case of the 3'flap, 1–6 nucleotides in the case of the 5' flap although longer 5' flaps can be cleaved)[8] (Fig. S1D). The preferred substrate of XPG is a bubble which contains no flaps, but instead a region of single stranded DNA on both strands flanked by duplex DNA on both sides of the bubble (Fig. S1G). In addition, the pseudo-Y substrate representing a "half-bubble" is also cleaved efficiently by XPG[2] (Fig. S1A). Mk0566 did not cleave the bubble or pseudo-Y substrates; instead, cleaved the 5' flap (Fig. S1B)(substrates 5–6, Table S1), double flap construct (Fig. S1D)(Lane 10, Fig. S3), and the blocked flap (Fig. S1E), consistent with reported FEN-1 cleavage specificity. These results indicate that Mk0566 has an activity profile more consistent with FEN-1, rather than the XPG-type enzymes.

Sedimentation Velocity & Equilibrium

In order to characterize the oligomeric state of Mk0566, sedimentation velocity and equilibrium experiments were performed using an analytical ultracentrifuge. Sedimentation velocity estimated the molecular weight as 42.0 kDa (Fig. S4), and global fitting of data obtained from sedimentation equilibrium to a single species model estimated the molecular weight of the enzyme to be 39.5 kDa (Fig. S5). Both estimates are very close to the calculated monomeric molecular weight of Mk0566 (41.0 kDa), therefore Mk0566 is monomeric under these solution conditions.

Crystal Structure of Mk0566

Crystals of purified Mk0566 were prepared in the presence of double-flap DNA, and a 2.20 Å diffraction data set was collected at SSRL beamline 9–2 (Table 1). The space group of the diffraction data was determined to be $P2_12_12_1$ with cell dimensions of 55.33 Å, 86.19 Å, and 147.62 Å. The final merged, scaled, and reduced data show an I/σ of 10.6 and R_{merge} of

8.8% (Table 1). Molecular replacement was used to solve the structure, and the asymmetric unit was found to contain 2 copies of Mk0566 (Fig. 1A). No electron density for the DNA was observed. Iterative refinement, density modification, and model building was able to produce a model with all residues with the exception of residues 88–123, 340–348 at the C-terminus, and the additional 7 amino acids on the amino terminus produced from cloning artifacts. The disordered region at 88–123 contains the helical arch region of FEN-1. Representative electron density is shown in Figure S6. A total of four solvent peaks (two bound to each copy of Mk0566) were identified as Cl^- ions based on their scattering (9 σ in the omit difference map), surroundings (nearest neighbors include the amide backbone nitrogen of Ser 317, side chain of Asn 67, tyrosyl hydroxyl of Tyr 63, and the positively charged side chains of Arg 320 and Arg 327), and refined temperature factors. When modeled as Cl^-, their refined temperature factors are similar to those of surrounding water molecules (average temperature factor for refined Cl^- ions: 44 $Å^2$, average temperature factor of 10 nearest water molecules: 40 $Å^2$). When modeled as water molecules, the temperature factors refine to values much lower than other water molecules (average of 22 $Å^2$). A sequence alignment of Mk0566 with selected FEN-1 sequences, showing conserved residues and Mk0566 secondary structure elements, is shown in Figure 1B.

The structure of Mk0566 is most similar to other FEN-1 enzymes from archaea, with RMSD between 1.5–2.6 Å over 266–293 residues (Table S2). The RMSD between Mk0566 and human FEN-1 is 1.9 Å, and to human ExoI is 2.7 Å. Unfortunately, no structures of XPG or GEN1 type enzymes have been reported so a comparison could not be made. Comparison of the structure of Mk0566 to the closest archaeal FEN-1 structures shows some interesting differences; the topology and secondary structure around residue 265 is different in Mk0566 and *P. furiosus* FEN-1 (Fig. S7A), but very similar in the unique β pin structure (residues 189–202, Mk0566 numbering) of both enzymes (Fig. S7B). The topology around residue 265 in *A. fugidus* FEN-1 is similar to that of Mk0566, however different in its secondary structure at residues 266–273 (Fig. S7C).

Although no DNA was found in the electron density of the Mk0566 structure, superposition of Mk0566 onto the structures of human FEN-1 bound to DNA (PDB ID 3Q8M, blue, Fig. S8A, with Mk0566 in green) and T4 RNase H (PDB ID 2IHN, blue, Fig. S8B, with Mk0566 in green) suggest how the flap DNA likely binds Mk0566. Most of the important secondary structure elements of these enzymes are conserved in Mk0566 (Fig. S8A-B). The electron density of Mk0566 showed two electron density features (four total in the asymmetric unit) that have been interpreted as chloride ions (red spheres, Fig. S8A–B). These occur very close to the locations of two phosphate groups, one from each strand, from the upstream duplex near the 3'flap binding site (Fig. S8A and Fig. 2B).

The structure of FEN-1 enzymes has been likened to a "left handed boxing glove"[9] (Fig. 2A), with the helical arch forming the thumb, the upstream DNA binding region as the fingers, and the downstream DNA binding region as the wrist part of the glove. A superposition of the structure of human FEN-1 bound to DNA missing the 5' flap but containing a 1 nt 3' flap as well as upstream and downstream duplex DNA[9] (magenta and wheat color, Fig. 2A) onto Mk0566 (green, Fig. 2A) shows how some of the unique features of Mk0566 might be related to substrate DNA binding. The helical arch is an important

structure as it forms over the active site where DNA cleavage occurs, and contains a gap through which the single stranded 5' end of the 5'flap DNA has been proposed to thread[8, 9]. The structure of human FEN-1 bound to DNA provides the best snapshot to date of how this may occur, although the 5'flap is too short to actually thread through the arch. The issue of threading (or not) has been debated in the literature[9], and is not possible in the case of the related XPG enzyme since no free end occurs in the bubble substrate cleaved by XPG. It may be that part of the XPG sequence that forms the arch in FEN-1 enzymes takes on a different structure in XPG which could isolate the single stranded DNA into the active site using a "gateway" common to FEN-1 and XPG enzyme sequences[9, 10]. In fact, sequence alignments indicate that XPG contains a repeat of these sequences on either side of a large insert (Fig. S9). The helical arch is disordered in the structure of Mk0566, which does not contain bound DNA, and is consistent with the idea that a disorder-to-order transition occurs upon DNA binding that may facilitate threading[11]. The disordering of the helical arch may allow 5' flaps with bulky groups or annealed DNA, such as the blocked flap (Fig. S1E) which was observed to be cleaved by Mk0566 (Fig. S3), to be bound by Mk0566 as suggested for other FEN-1 enzymes[8].

Other important landmarks on the FEN-1 structure include the hydrophobic wedge (Fig. 2A), which helps to position the DNA with the 5' flap towards the active site (Fig. 2A) and the 3' flap towards the 3'flap binding pocket (Fig. 2A). All three of these regions are well conserved structurally in Mk0566 compared to human FEN-1 and other reported FEN-1 structures (Table S2). The H3TH (or in human FEN-1 H2TH[9]) is also well conserved, however some interesting unique features occur at 265 region and the β pin (Fig. 2A, Fig. S10A–B, Fig. 2C–D). The region around residue 265 forms a small alpha helical domain that is near the binding site of the downstream duplex DNA (Fig. 2A). The path followed by the protein in the region in Mk0566 (green, Fig. S10A–B) deviates from that of *P. furiosus* FEN-1 (red, Fig. S10B), but is more similar to that followed by *A. fulgidus* FEN-1 (red, Fig. S7C) and human FEN-1 (magenta, Fig. S10A), although the human FEN-1 has a long loop that more closely approaches the DNA (Fig. S10A), and neither *A. fulgidus* nor human FEN-1 form an alpha helix between residues 266–273 (Mk0566 numbering) as in Mk0566 (green, Fig. S7C, Fig. S10A). The differences in structure may be related to substrate binding, may change upon DNA binding, or they may have other implications for substrate and/or partner protein binding.

The β pin is a two β stranded protuberance that is located (after superposition with the DNA bound human FEN-1 structure) in the major groove of the upstream duplex of the bound DNA, as well as proximal to the backbone of the "template strand" (the strand that is continuous between the two duplex regions) opposite from the helical arch (Fig. 2A). Both *P. furiosus* FEN-1 and Mk0566 share this feature (Fig. S7B, Fig. 2D) which is extended relative to the related structures in human FEN-1 (magenta, Fig. 2C) and *M. jannaschii* (PDB ID 1A77) (see also sequence alignments at residues 184–202, Mk0556 numbering, Fig. 1B). This feature, being much longer in the Mk0566 and *P. furiosus* FEN-1 structures, reaches further into the major groove and may result in tighter substrate binding, or perhaps even greater specificity for double stranded in the upstream DNA (i.e. greater specificity for 5'flap vs. pseudo-Y structures). Although some FEN-1 enzymes cleave both pseudo-Y and

5'flap substrates, Mk0566 did not cleave a pseudo-Y substrate in our assays (Table S1). The binding of single stranded DNA in the pseudo-Y substrate bound to T4 RNase H (blue, Fig. S8B) in this region shows a different path that is inconsistent with the observed position of the β pin in Mk0566, although a shift in either the β pin or DNA could avoid steric collision. Still the β pin (and perhaps the putative phosphate binding sites marked by Cl^- ions (red spheres, Fig. 2B)), may be important in substrate discrimination given that Mk0566 (and *P. furiosus* FEN-1) both show greatly increased specificity for duplex regions in the upstream DNA relative to other FEN-1 enzymes, including those from human and from *M. jannaschii* which do not possess the elongated β pin found in Mk0566 and *P. furiosus* FEN-1[6, 12–14].

In conclusion, we find that Mk0566 displays a substrate preference and three dimensional structure similar to other archaeal FEN-1 enzymes, as well as to human FEN-1. The absence of cleavage of bubble substrates makes it an unlikely candidate for the XPG activity required by NER. It is possible that Mk0566 could act in repairing DNA lesions through a long-patch repair system, where 5'flap DNA occurs as an intermediate and can act as a substrate for Mk0566. Alternatively, the enzyme Bax1 may instead perform the role of XPG in some archaeal species[15], although this enzyme does not appear to be present in *M. kandleri*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). In addition, research reported in this publication was supported by the Office of the Director, National Institutes of Health of the National Institutes of Health under Award Number S10OD013237. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

Abbreviations

bp	base pair(s)
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
HOAc	acetic acid
nt	nucleotide(s)
OAc	acetate
NER	nucleotide excision repair
PMSF	phenylmethylsulfonyl fluoride
RT	room temperature

References

1. Thoms KM, Kuschal C, Emmert S. Lessons learned from DNA repair defective syndromes. Experimental dermatology. 2007; 16:532–544. [PubMed: 17518994]
2. Shen B, Singh P, Liu R, Qiu J, Zheng L, Finger LD, Alas S. Multiple but dissectible functions of FEN-1 nucleases in nucleic acid processing, genome stability and diseases. Bioessays. 2005; 27:717–729. [PubMed: 15954100]
3. Tsutakawa SE, Lafrance-Vanasse J, Tainer JA. The cutting edges in DNA repair, licensing, and fidelity: DNA and RNA repair nucleases sculpt DNA to measure twice, cut once. DNA repair. 2014; 19:95–107. [PubMed: 24754999]
4. Gillet LC, Scharer OD. Molecular mechanisms of mammalian global genome nucleotide excision repair. Chem Rev. 2006; 106:253–276. [PubMed: 16464005]
5. Chapados BR, Hosfield DJ, Han S, Qiu J, Yelent B, Shen B, Tainer JA. Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. Cell. 2004; 116:39–50. [PubMed: 14718165]
6. Harrington JJ, Lieber MR. The characterization of a mammalian DNA structure-specific endonuclease. EMBO J. 1994; 13:1235–1246. [PubMed: 8131753]
7. Finger LD, Blanchard MS, Theimer CA, Sengerova B, Singh P, Chavez V, Liu F, Grasby JA, Shen B. The 3'-flap pocket of human flap endonuclease 1 is critical for substrate binding and catalysis. J Biol Chem. 2009; 284:22184–22194. [PubMed: 19525235]
8. Sobhy MA, Joudeh LI, Huang X, Takahashi M, Hamdan SM. Sequential and multistep substrate interrogation provides the scaffold for specificity in human flap endonuclease 1. Cell Rep. 2013; 3:1785–1794. [PubMed: 23746444]
9. Tsutakawa SE, Classen S, Chapados BR, Arvai AS, Finger LD, Guenther G, Tomlinson CG, Thompson P, Sarker AH, Shen B, Cooper PK, Grasby JA, Tainer JA. Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. Cell. 2011; 145:198–211. [PubMed: 21496641]
10. Hohl M, Dunand-Sauthier I, Staresincic L, Jaquier-Gubler P, Thorel F, Modesti M, Clarkson SG, Scharer OD. Domain swapping between FEN-1 and XPG defines regions in XPG that mediate nucleotide excision repair activity and substrate specificity. Nucleic Acids Res. 2007; 35:3053–3063. [PubMed: 17452369]
11. Patel N, Atack JM, Finger LD, Exell JC, Thompson P, Tsutakawa S, Tainer JA, Williams DM, Grasby JA. Flap endonucleases pass 5'-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5'-ends. Nucleic Acids Res. 2012; 40:4507–4519. [PubMed: 22319208]
12. Hosfield DJ, Mol CD, Shen B, Tainer JA. Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. Cell. 1998; 95:135–146. [PubMed: 9778254]
13. Kaiser MW, Lyamicheva N, Ma W, Miller C, Neri B, Fors L, Lyamichev VI. A comparison of eubacterial and archaeal structure-specific 5'-exonucleases. J Biol Chem. 1999; 274:21387–21394. [PubMed: 10409700]
14. Hosfield DJ, Frank G, Weng Y, Tainer JA, Shen B. Newly discovered archaebacterial flap endonucleases show a structure-specific mechanism for DNA substrate binding and catalysis resembling human flap endonuclease-1. J Biol Chem. 1998; 273:27154–27161. [PubMed: 9765234]
15. Rouillon C, White MF. The evolution and mechanisms of nucleotide excision repair proteins. Res Microbiol. 162:19–26. [PubMed: 20863882]

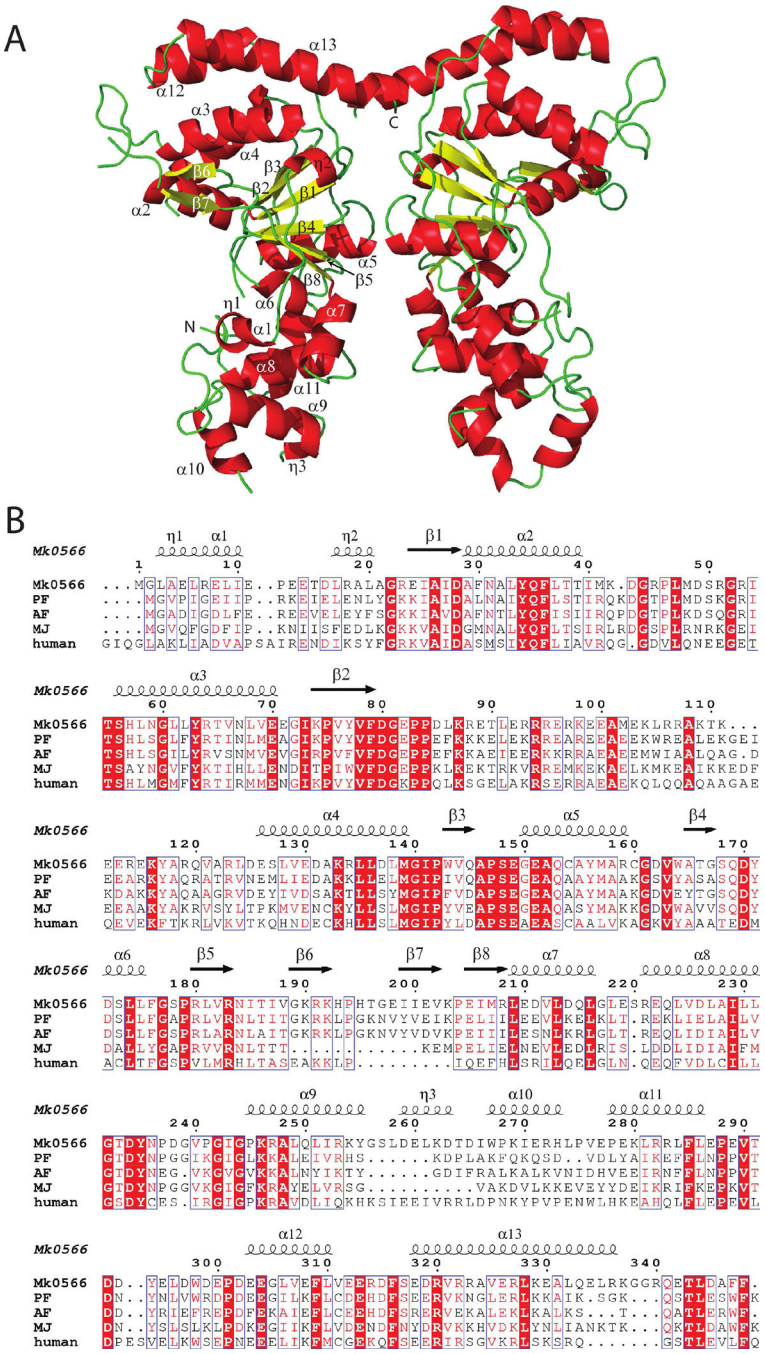

Figure 1.

A. Ribbon diagram of the two copies of Mk0566 in the asymmetric unit, showing alpha helices (red), beta sheets (yellow) and loops (green). Secondary structural elements labeled on one of the two copies in the asymmetric unit (α=alpha helix, β=beta strand, η=3_{10} helix, N=amino terminus, C=carboxy terminus) **B.** Alignment of Mk0566 (top row, Mk0566) sequence with selected FEN-1 enzymes: PF, *P. furiosus* FEN-1, AF, *A. fulgidus* FEN-1, MJ, *M. jannaschii* FEN-1, human, human FEN-1, with corresponding secondary structure of Mk0566 (above aligned sequences). Conserved residues are boxed and shown in red font,

strictly conserved residues are shown in white font and highlighted in red. (Note: the C-terminal R residue of PF is not shown). Secondary structural elements for Mk0566 shown above the sequences, with definitions as in A.

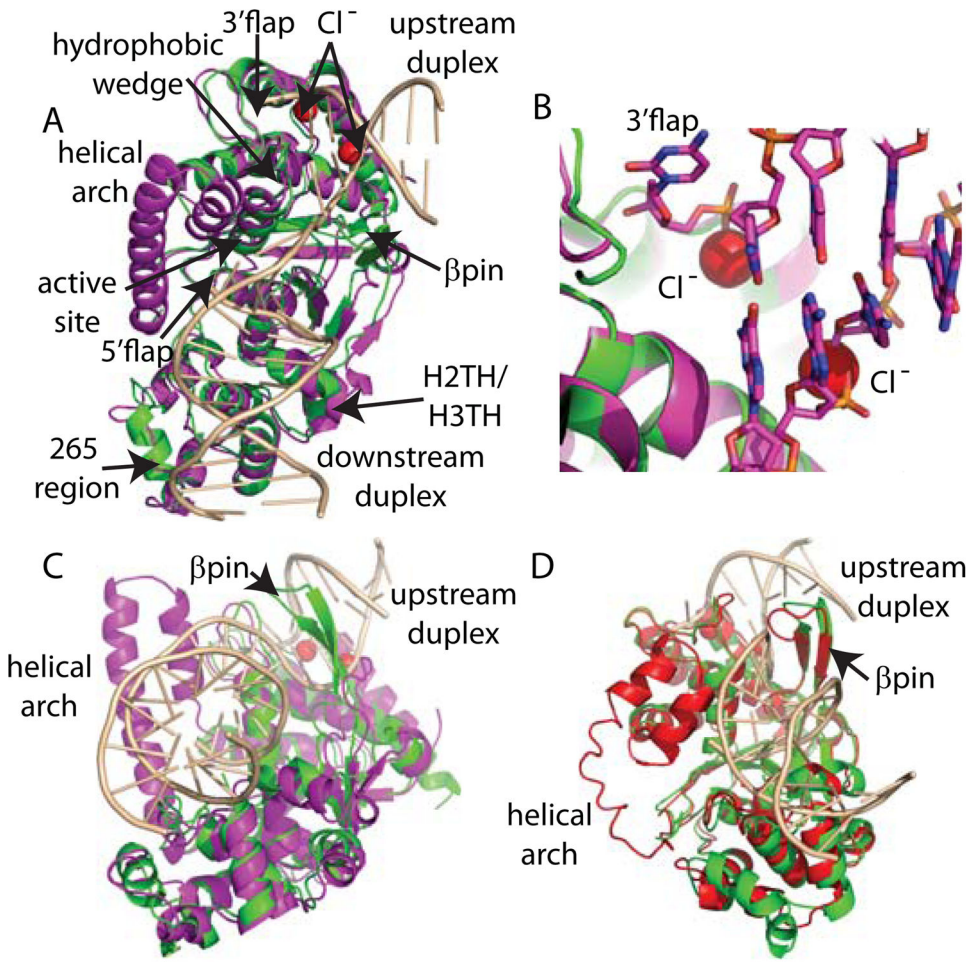

Figure 2.

Comparison of the Mk0566 structure with human FEN-1 bound to DNA (PDB ID 3Q8M) and *P. furiosus* FEN-1 (PDB ID 1B43). **A.** Superposition of human FEN-1 (magenta) bound to DNA (wheat) with Mk0566 (green). Red spheres mark the position of the Cl^- ions bound to Mk0566. **B.** As in A, zoom in near the Cl^- binding sites. **C.** As in A, perspective to emphasize position of the β pin region of Mk0566. **D.** Superposition of *P. furiosus* FEN-1 (red) and Mk0566 (green) with perspective to emphasize β pin orientation. DNA (wheat) is derived from the human FEN-1 structure.

Table 1

Diffraction Data and Structure Refinement Statistics for Mk0566

Space Group	$P2_12_12_1$
Cell	55.33 Å, 86.19 Å, 147.62 Å,
Resolution	2.20 Å (2.26-2.20 Å)[1]
Total Observations	208,925
Unique Observations	35,363
Completeness	96.4 % (88.5%)[1]
I/sigma	10.6 (2.2)[1]
Multiplicity	5.9 (5.9)[1]
Rmerge[2]	8.8 % (74.1%)[1]
Rcryst[3]	20.0 %
Rfree[4]	23.1 %
Overall B factor (Wilson plot)	39.2 $Å^2$
RMSD-bonds	0.008 Å
RMSD-angles	1.19°
Asymmetric unit	2 monomers
Number of protein atoms	4752
Numbers of waters	86
Number of Cl^-	4
PDB entry	4WA8

[1] Values in parentheses are those for the highest resolution shell.

[2] $R_{merge}=\Sigma_{hkl}(|<I_{hkl}>-I_{hkl}|)/(\Sigma_{hkl} I_{hkl})$ where $<I_{hkl}>$ is the average intensity over symmetry related and equivalent reflections and I_{hkl} is the observed intensity for reflection hkl.

[3] $R_{cryst}=\Sigma_{hkl}(||F_{obs}|-|F_{calc}||)/(\Sigma_{hkl}|F_{obs}|)$ where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitude for reflection hkl. The sum is carried out over the 95% of the observed reflections which are used in refinement.

[4] R_{free} refers to the R factor for the test reflection set (5% of the total observed) which was excluded from refinement.