The EMBO Journal
August 2009 ; Volume 28 (16) : Pages 2479-2489
http://dx.doi.org/10.1038/emboj.2009.192

Structure and function of a novel endonuclease acting on branched DNA substrates

Bin Ren[1, 7, 8], Joelle Kühn[2, 7], Laurence Meslet-Cladiere[2], Julien Briffotaux[3, 4, 9], Cedric Norais[2], Regis Lavigne[5], Didier Flament[3, 4], Rudolf Ladenstein[1] and Hannu Myllykallio[2, 6, *]

[1] Center for Structural Biochemistry, Karolinska Institutet, NOVUM, Huddinge, Sweden
[2] Institute of Genetics and Microbiology, CNRS UMR8621, Universite Paris XI, Orsay, France
[3] Ifremer, UMR 6197, Laboratoire de Microbiologie des Environnements Extrêmes, Plouzané, France
[4] Laboratoire de Microbiologie des Environnements Extrêmes, Université de Bretagne Occidentale, UMR 6197, Plouzané, France
[5] INSERM U625, Université Rennes I, Campus de Beaulieu, Rennes, France
[6] Laboratoire d'Optique et Biosciences, INSERM U696, Centre National de la Recherche Scientifique Unité Mixte de Recherche 7645, Ecole Polytechnique, Palaiseau Cedex, France
[7] These authors contributed equally to this work
[8] Present address: Institute of Nano-biomedical Technology and Membrane Biology, State Key Laboratory of Biotherapy, Sichuan University, 610041 Chengdu, China
[9] Present address: Département de Microbiologie et d'Infectiologie, Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, Sherbrooke, Québec, Canada J1H 5N4

*: Corresponding author : Hannu Myllykallio, Telephone:+33169335010, Fax: +33169335084, email address : hannu.myllykallio@polytechnique.edu

Abstract:

We show that *Pyrococcus abyssi* PAB2263 (dubbed NucS (nuclease for ss DNA) is a novel archaeal endonuclease that interacts with the replication clamp PCNA. Structural determination of *P. abyssi* NucS revealed a two-domain dumbbell-like structure that in overall does not resemble any known protein structure. Biochemical and structural studies indicate that NucS orthologues use a non-catalytic ssDNA-binding domain to regulate the cleavage activity at another site, thus resulting into the specific cleavage at double-stranded DNA (dsDNA)/ssDNA junctions on branched DNA substrates. Both 3' and 5' extremities of the ssDNA can be cleaved at the nuclease channel that is too narrow to accommodate duplex DNA. Altogether, our data suggest that NucS proteins constitute a new family of structure-specific DNA endonucleases that are widely distributed in archaea and in bacteria, including *Mycobacterium tuberculosis*.

Keywords: branched DNA structures, DNA repair, novel endonuclease, RecB family, structure–function studies

Introduction

Branched DNA structures are created either by DNA damage or occur as intermediates during DNA replication, repair and recombination. To restore the genomic integrity, these DNA structures must be resolved and processed by specialized enzymes. In human cells, XPF-family endonucleases are key enzymes participating in processing of these anomalous DNA structures. These nucleases are often formed by an amino-terminal helicase domain that is fused to the ERCC4 nuclease domain. Strikingly, at least six different XPF-family members can be found in human cells (Ciccia et al., 2007). Among these, Mus81 and XPF are active endonucleases that preferentially act on branched DNA substrates, whereas the ERCC4 nuclease domains in EME1 and 2, FAAP24, ERCC1 and FANCM are predicted to be inactive. Eukaryotic XPF and Mus81 endonucleases form heterodimers with ERCC1 and Eme1/2, respectively (Newman et al., 2005;

Nishino et al., 2005b; Tsodikov et al., 2005). Although being evolutionarily related, Mus81 and XPF have different substrate specificities. For instance, Mus81 cleaves the 3' flap structures that could occur downstream of replication fork collapse and XPF cleaves the simple Y-forks containing unpaired 5' and 3' regions ("splayed arm") and stem-loop substrates. Recent studies have broadened the substrate specificity of the Mus81 complex to splayed arm (Y-forks) substrates as well and 5'-flaps (Ehmsen and Heyer, 2008). FANCM is a key component of the Fancomi anemia core complex, indicating a functional role in repair of DNA crosslinks. Recent studies have demonstrated that human FANCM interacts with FAAP24 (Ciccia et al., 2007) and binds various branched DNA molecules (Gari et al., 2008). Human FAAP24/FANCM complex does not seem to function as a nuclease but can promote branch migration of Holliday junctions and replication forks. These observations indicate that these numerous related proteins have evolved to counteract a variety of DNA damages.

The archaeal *Pyrococcus* and *Sulfolobus* species contain only one member of XPF-endonuclease family that *a priori* could correspond to the ancestor of eukaryotic XPF-family members. These proteins are very stable and provide an excellent model for understanding structure-function relationships of XPF family members. *P. furiosus* XPF-homolog, dubbed Hef (Nishino et al., 2005a; Nishino et al., 2005b), contains active helicase and nuclease domains that are related to human FANCM protein (Meetei et al., 2005). The nuclease domain of *P. furiosus* Hef is active but, differently from eukaryotic XPF/ERCC1 complex, it does not cleave splayed arm substrates (Komori et al., 2004). The substrate specificity of this archaeal protein is thus distinct from eukaryotic XPF-proteins. On the other hand, activity profile of *Sulfolobus* species XPF shares similarities both with the eukaryotic Mus81 and XPF (Roberts and White, 2005). However, the nuclease domain of *Sulfobus* XPF is not fused to the helicase domain but its activity is strictly regulated by the replication clamp (Roberts et al., 2003). An additional

72 difference with heterodimeric eukaryotic XPF-family members is that archaeal proteins function
73 as homodimers in their cellular context (Newman et al., 2005; Nishino et al., 2005b; Tsodikov et
74 al., 2005). Structural studies have indicated that a nuclease domain of XPF-family members is
75 followed by two consecutive DNA-binding helix-hairpin-helix (HhH) motifs. DNA binding often
76 requires conformational adjustment between different domains (Chapados et al., 2004; Newman
77 et al., 2005). Clearly, many functionally distinct but structurally related structure specific
78 nucleases have evolved for processing branched DNA substrates. To understand molecular basis
79 for the cleavage specificity of structure specific DNA endonucleases and their chromatin loading
80 require further experimental attention.

81 Experimental observations indicating that *Pyrococcus* Hef does not cleave splayed arm
82 substrates containing unpaired 3' and 5' extremities raised a possibility that additional structure-
83 specific endonucleases specifically acting on ssDNA extremities exist in *Pyrococcus* species
84 (Komori et al., 2004). This possibility was further favored by the observation that the damage
85 recognition proteins XPA and XPC have not been identified in archaeal genomes (Kelman and
86 White, 2005), suggesting that detection of DNA damage could be mediated by detecting single-
87 stranded DNA carried by branched DNA replication/repair intermediates (Cubeddu and White,
88 2005). We show here that a previously uncharacterized member of RecB-family nuclease
89 *Pyrococcus abyssi* PAB2263 cleaves single-stranded regions in branched DNA structures. *In*
90 *silico* observations and biochemical studies proposed that NucS orthologs of hyperthermophilic
91 archaea may participate in processing of branched DNA structures. Crystallographic studies
92 revealed that *P. abyssi* NucS possesses unique structural features for the binding and processing
93 of the exposed single-stranded regions of the different types of branched DNA substrates.
94 Unexpectedly, sequence similarity searches revealed that, in addition to archaea, several bacteria,
95 including *Mycobacterium tuberculosis*, contain the NucS ortholog.

Results

Pyrococcus abyssi NucS (PAB2263) interacts with the replication clamp

Using a combination of motif searches and combinatorial peptide synthesis, we have recently shown that the **short** peptide corresponding to twelve carboxy-terminal amino acid residues of the *P. abyssi* NucS (indicated "PIP" in the Figure 1A) binds strongly to the homotrimeric replication clamp of this organism (Meslet-Cladiere et al., 2007). In *Pyrococcus* species NucS is encoded nearby the evolutionary conserved replication origin (Myllykallio et al., 2000), and genome sequence of *Thermococcus kodakarensis* KOD1_4 has indicated that the NucS ortholog is translationally coupled with RadA recombinase (Figure 1A).

Preliminary activity measurements indicated that NucS proteins posses a potent nuclease activity specific for ssDNA (data not shown, see also Figure 5). We thus refer to this protein as *P. abyssi* NucS (**nuc**lease specific for **s**sDNA). *P. abyssi* NucS (and other representatives of DUF91-family) contains the carboxy-terminal domain that carries the characteristic residues for the "RecB" family nucleases that are found in a variety of endonucleases and DNA repair enzymes (Aravind et al., 2000, Figure 1A). The related nuclease domain can also be found in a large number of additional bacterial, archaeal and viral proteins without the known function (see supplemental table 1). To study *in vitro* formation of the NucS-PCNA complex, we used surface plasmon resonance (SPR). First, NucS carrying a six-histidine tag at the N-terminus, was produced for SPR measurements and purified to greater than 99% homogeneity using Ni-NTA agarose and gel filtration chromatography (Figure 1B). SPR measurements with immobilized homotrimeric PCNA were performed to analyze the physical interactions of *P. abyssi* NucS with PCNA (Figure 1C). A K_D value of 15 nM, which reflects strong binding, was estimated for this interaction through global fitting analysis, accounting for both association and dissociation phases using a 1:1 interaction model. The high affinity for the observed interaction *in vitro* also

explains why the NucS-PCNA complex can be readily detected using the pull-down experiments and Western blotting analyses in *P. abyssi* cell free extracts (Figure 1C, inset).

P. abyssi NucS co-precipitates with PAB0190 (Hef) in cell-free extracts

We attempted identification of additional interaction partners for *P. abyssi* NucS (or a stable NucS-PCNA complex) in cell-free extracts. Pull-down experiments were initiated by incubating 20 μg of hexa-histidine tagged bait protein bound to Co^{2+} magnetic microbeads with the cell-free extracts prepared from the exponentially grown cultures. After extensive washes, proteins bound to the bait protein were eluted in SDS sample buffer, followed by 12% SDS-PAGE separation and Coomassie blue staining. Figure S1 shows the SDS-PAGE image of the proteins specifically interacting with *P. abyssi* NucS, revealing several candidate proteins that directly or indirectly interact with the bait protein. Notably, these proteins were absent from the negative control reaction (Figure S1) and pull-down experiments performed with additional DNA interacting proteins (data not shown). The seven major protein bands were excised from the gel and subjected to in-gel tryptic digestion. The obtained peptides were extracted from the gel and analyzed, leading to identification of nine proteins that co-precipitated with NucS and/or a highly stable (K_D=15 nM) NucS-PCNA complex [indicated *a* through *i* (note that some bands corresponded to two individual proteins)]. The following proteins were identified in this experiment: Hef (PAB0190), PCNA (confirming the above results shown in Figure 1C), small and large subunit of the replication factor C (the clamp loader) and two subunits of topoisomerase VI.

The two-domain subunit structure of NucS

To reveal the structural basis for its enzymatic activity, we determined the crystal structure of *P. abyssi* NucS. Each subunit of this protein clearly displays a dumbbell-like two-domain structure. The N- and C-terminal domains are separated from each other by a distance of about 28 Å and connected by a stretched polypeptide linker, consisting of residues 115 to 125 (Figure 2A). Neither the subunit as a whole nor any of the domains displays structural similarities to functionally characterized flap endonucleases.

The N-terminal domain (residues 1-114) basically has a unique half-closed β-barrel structure (Figure 2B). Eight β-strands form two anti-parallel β-sheets that are packed orthogonally onto each other, generating two layers of anti-parallel β-sheets: the six-stranded sheet β1β8β2β3β4β5 and the five-stranded sheet β6β7β8β2β3. In the middle of the β-barrel are three long β-strands (β8, β2, β3) that fold over and span into both sheets to constitute the backbones of the β-barrel. Their curved turns form a closed bottom of the barrel on one side. The other five β-strands and helix α1 are distributed equally in the two layers of the β-barrel surrounding the three long strands. Loops connecting the β-strands gather on one side of the β-barrel opposite the closed bottom. Although structural comparison did not reveal any apparent similar folds in the Protein Data Bank, the N-terminal β-barrel can be viewed remotely analogous to the Sm-fold of the eukaryotic RNA binding domain of small nuclear ribonucleoproteins (Sm proteins) and the OB-fold of single strand DNA binding proteins (Kambach et al., 1999; Theobald et al., 2003, see also Figure S2). Proteins in these two families also form a β-barrel or barrel-like structure but with a smaller scaffold consisting of one helix and five strands arranged in different topological placements.

The C-terminal domain of *P. abyssi* NucS (residues 126 to 233) possesses an α/β structure with a five-stranded central β-sheet and four flanking α-helices (Figures 2A and 2C), representing a minimal endonuclease fold (Pingoud et al., 2005). The C-terminal domain hosts an active site with a sequence motif conserved in the family of RecB-like nucleases (Aravind et al., 2000), although the size of the C-terminal domain (108 residues) is much smaller than that of the *E.coli* RecB nuclease domain (108 residues *versus* 274 residues) (Singleton et al., 2004). Structural alignment of their cores gives a root-mean-square deviation of 3.2 Å for 82 Cα atoms (Figure 2C).

Oligomerization of *P. abyssi* NucS in the crystal

The dumbbell-like subunit structure of NucS has a large hydrophobic patch exposed on the six-stranded N-terminal β-sheet. By domain swapping, the subunits assemble to form a dimer that constitutes the asymmetric unit. The dimer displays a crescent shape and the four domains of the two subunits are arranged in the order of $C_1N_2N_1C_2$ (Figure 2D). Upon dimer formation the hydrophobic patches are buried to form a large hydrophobic core in the middle between the two N-terminal β-barrels, resulting in the formation of an interface β-barrel due to the back-to-back packing of the two six-stranded sheets. From the structural point of view, dimer formation appears critical for the folding and stabilization of the NucS structure. In the crystal NucS dimers further aggregate to form tetramers because of crystallographic 2-fold symmetry. However, while the formation of dimer from monomer results in a buried accessible surface area of 2880 $Å^2$ for each subunit (20 % of the surface area of the subunit), tetramer formation only buries 424 $Å^2$ more for each subunit (which is about 3 % of the subunit surface area), indicative of a weak inter-dimer association that may result from crystal contacts.

DNA induced oligomerization in solution

To investigate the properties of *P. abyssi* NucS in solution, we performed dynamic light scattering and sucrose gradient centrifugation experiments. The predicted molecular weights of the *P. abyssi* NucS dimer and tetramer are 60 and 120 kDa, respectively. Figure 1D indicates that in 600 mM NaCl, pH 8.0, the mean diameter of NucS protein in solution is approximately 7.5 nm, which corresponds to an approximate molecule weight of 59 - 63 kDa. In agreement with this observation, *P. abyssi* NucS co-sediments with BSA (66 kDa) in 5-20 % sucrose gradients [300 mM NaCl, pH 8 (Figure 1E), also indicating dimeric solution structure. However, when a single stranded oligonucleotide [87-mer (30 kDa), molar protein monomer: oligonucleotide ratio 4:1] was combined with the protein sample, the particle size significantly increased to a value of 12.5 nm.

Identification of a putative ssDNA binding and the nuclease active site

On the concave surface of the dimer exists a groove that is enclosed by the loops gathering on one side of the N-terminal β-barrel (Figures 2B and 3A). Structural comparison (Figure S2) raised a possibility that this groove is equivalent to the single-stranded nucleic acid binding sites found in OB- / Sm- family proteins. This potential ssDNA binding site (site I) was verified by our mutagenesis studies (see below). On one side of the groove, a conserved tryptophane residue, W75, and residues 68 to 71 form a surface protrusion close to the dimer interface (Figure 3B). Two such neighboring protrusions separate the two site I clefts related by the non-crystallographic 2-fold axis. The bottom surface of site I is essentially hydrophobic and is surrounded by more than a dozen charged residues, notably among which are two clusters of basic residues that consist of K68, K69, R70 and R93, R94, R95, respectively. They are located

on different sides of the cleft (Figures 3A and B). R42 sticks out from the surface and is conserved in all NucS sequences (Figure 4D). It resides on the side neighboring R93-R95. The surface of the cleft is therefore rather basic (Figure S3). The site I cleft has a length of about 33 Å with a depth of 9 Å and a width of 13 Å, appropriately sized to accommodate ssDNA. It is plausible that upon the binding of ssDNA the phosphate backbone would form charged interactions or hydrogen bonds with the basic residues, whereas the DNA bases forge hydrophobic interactions and hydrogen bonds with the residues constituting the bottom and side surface of the cleft. The side chains of aromatic residues Y39 and W75 that participate in the formation of the walls of site I are ideally orientated and may form stacking interactions with DNA bases (Figures 3A and C). The aromatic rings of these two residues roughly face each other from the two sides of the cleft and both point outwards from the surface. Figure S4A shows the electron densities at this region. The residues Y39 and W75 of *P. abyssi* NucS are evolutionary conserved (Figure 4D), indicating their functional importance.

The nuclease active site (site II) sits in a dimer interface cleft that is capped by the flexible linker connecting the N- and C-terminal domains (Figure 3). Several active site residues that are conserved in the RecB motif (Figure 4D), including D160, E174, K176, Q187 and Y191, are located at the bottom of the cleft. A cluster of basic residues, including K176, R177, R178, K179, flank the cleft on one side. The linker region ranging from residue 121 to residue 125 shows weak electron densities and forms a bridge above the conserved active site residues, making site II a channel that penetrates through the wall of the protein (Figure 3B, C), which indicates that the substrate for the enzyme must have a free end, such as that of a broken DNA strand. Compared to site I, site II is more hydrophilic and has a length of about 20 Å with a depth of 13 Å. It is conceivable that upon the binding of ssDNA the phosphate backbone would probably make contacts with the bottom of the channel by forging hydrogen bonds or charged interactions

with the conserved active site residues of E127, D160, E174 and K176, to facilitate the digestion of phosphodiester bonds (Figure 3C).

The orientations of sites I (N-terminal ssDNA binding domain) and II (active site) are roughly perpendicular to each other with a cross angle of about 50 to 60 degrees. Upon superposition the active site region of NucS can be aligned well with that of the *E.coli* RecB nuclease domain (Figure 2C), suggesting that they share a similar nuclease mechanism (Pingoud et al., 2005). Consequently, the conserved acidic residues of E127, D160 and E174 in *P. abyssi* NucS likely function in catalysis to ligate a magnesium ion required for catalysis, allowing stabilization of a deprotonated water molecule (in fact a hydroxyl ion) that will hydrolyze the phosphodiester bond of DNA. Compared to the other enzymes, the N-terminal domain of NucS provides an extra basic residue, K44, to the active site, which together with the conserved K176 may help to stabilize negatively charged transition states in the reaction (Figure 3C). Electron densities at the active site region are shown in Figure S4B.

Site I is sufficient and necessary for high affinity binding of ssDNA

We tested the hypothesis that site I of NucS proteins could contain a high affinity binding site for ssDNA through biochemical studies. This study was performed either using SPR analysis on the oligonucleotide immobilized through the biotin moiety carried in the 3' terminus of a ssDNA oligonucleotide (87-mer) or by electrophoretic mobility shift assays (EMSA) (oligonucleotides used are indicated in supplemental Table 2). SPR measurements indicated that *P. abyssi* NucS physically interacted with ssDNA that was immobilized trough the 3' extremity at nanomolar concentrations [K_D=4 nM at 25 ^oC (Figure 4A)]. This result was confirmed by EMSA assays where DNA-protein complexes were allowed to form at 40 ^oC in the presence of varying concentrations of the single stranded oligonucleotide. Under these experimental conditions, we

detected a formation of a high affinity ssDNA–NucS complex with an apparent K_D value of 150 nM (Figure 4B).

Guided by structural data, we constructed and purified several mutant proteins (Figures 4C and D) that were expected to show reduced binding for ssDNA. Indeed, we demonstrate that, differently from the wild-type protein, the mutant proteins W75S, R42A, R70A and a double mutant R42A W75S failed to bind ssDNA under stoichiometric binding conditions used (Figure 4C). These experiments provide experimental proof that the ssDNA binding site suggested by structural studies is necessary and sufficient for the high affinity ssDNA binding. Note that the residues shown to be crucial for high affinity ssDNA binding in this study (filled circles in Figure 4D) are conserved in the archaeal or bacterial NucS homologs.

NucS slides onto 3' or 5' flaps before a specific cleavage at the ssDNA/dsDNA junction

Keeping in mind that the active site of NucS proteins is not suited to accommodate double stranded DNA, we used single-stranded 5' and/or 3' flap substrates (Figure 5) in nuclease assays where protein quantities were varied over a wide range. Figures 5A and B show that *P. abyssi* NucS can cleave either 5' or 3' flaps. Another case of bipolar cleavage activity has been demonstrated for a *Saccharomyces cerevisiae* Mus81-Mms4/Eme1 endonuclease recently (Ehmsen and Heyer, 2008). The appearance of the specific reaction products of approximately 25 (Figure 5A) and 22 nucleotides (Figure 5B) indicates specific cleavages of the 5' and 3' flaps around the ssDNA/dsDNA junction. This activity does not result from the contaminating exo- or endonuclease activity, as substitutions in the conserved "RecB" nuclease motif residues are catalytically inactive (Figure S5 and data not shown). We also observed that addition of excess protein resulted into the formation of a "smear" corresponding to reaction products ranging from 8 to 12 nucleotides. This finding is analogous to what has been observed earlier for the *P.*

furiosus Hef protein that is able to introduce unpairing of the base pairs close the junction (Nishino et al., 2005a; Nishino et al., 2005b). To test whether *P. abyssi* NucS directly binds to the ssDNA/dsDNA junction, or whether the 3'or 5' flaps must slide into the nuclease active site before the cleavage, we repeated the above activity measurements with biotinylated 3' and 5' flap substrates. Figures 5A and B show that streptavidin conjugation of the biotinylated flap substrates abolished the detectable nuclease activity. In agreement with our structural data, these observations suggest the blockage of the tracking reaction along the flap.

ssDNA binding and formation of a PCNA complex modulate the cleavage specificity of *P. abyssi* NucS

We performed additional enzymatic tests using a single stranded oligonucleotide and when it was hybridized to form the splayed arm structure that carries both the 3' and 5' protruding flaps. Figure 5C shows that *P. abyssi* NucS can specifically cleave ssDNA substrates, suggesting that there is some specific end recognition. This notion is also supported by the observation that the long single stranded oligonucleotide is cleaved into the several, apparently regularly spaced, DNA fragments (Figure S5). Although *P. abyssi* NucS was able to cleave also the splayed arm structure, the reaction products observed for ssDNA and splayed arm substrate were different. First, the sizes of the specific reaction products are not identical for single stranded and splayed arm substrates (Figure 5C; compare the left and the middle panel). Moreover, a "smear" similar to what we putatively assumed above for melting the base pairs nearby the junction was only observed for the splayed arm substrate. These measurements thus indicate that the unpaired DNA strand opposite to the strand that is cleaved and/or duplex regions in the substrate modulates the cleavage specificity.

To investigate the functional role of the ssDNA binding mediated by the site I, we performed enzymatic tests with the splayed arm template, using the double mutant R42A W75S that failed completely to bind single stranded oligonucleotide in our EMSA assays (Figure 4D). This mutant was still catalytically active, as the formation of "smear" of 8 – 12 nucleotides was still detected. However, we did not detect the formation of the specific cleavage at the junction (Figure 5C, the right panel), thus revealing that the high-affinity ssDNA binding at the site I regulates the catalytic activity of this class of enzymes. We found that the addition of PCNA, which encircles dsDNA in the cellular context, increases the cleavage specificity of NucS proteins (Figures 5D and S6). This weak activation results from inhibition of non-specific activity in our nuclease assays using splayed arm substrate. We have not observed a modulation of the NucS activity by the replication clamp on the 5' single stranded extremity of the splayed arm template used in Figure 5D.

Discussion

Our structural and functional studies indicate that NucS proteins constitute a novel family of DNA endonucleases that act on branched DNA structures. *P. abyssi* NucS does not show any overall structural similarities to other proteins, including functionally characterized flap endonucleases. The two identified DNA-binding sites are constructed to accommodate single strand nucleic acids only and serve as regulatory and ssDNA digestion sites, respectively. The aromatic residues of W75 and Y39 at site I in one subunit can form strong stacking interactions with exposed DNA bases so as to hold the intact strand, which would in turn help to feed the damaged strand into the nuclease site of the other subunit. We have observed that the ssDNA binding at the site I regulates the cleavage of splayed arm substrates (Figure 5C), indicating that the non-catalytic DNA binding site regulates the cleavage of ssDNA at the active site channel. It is feasible that the distinct ssDNA arms (i.e. 3' and 5' flaps) could be bound in the two

neighboring antiparallel clefts of site I related by the non-crystallographic symmetry. It can be inferred from the *P. abyssi* NucS structure that a continuous strand can only make contacts with the site I while a broken strand with a free end can further slide into the nuclease channel (site II) and then be digested. Taken together, our biochemical and structural analyses indicate that the coordination between the two DNA binding sites and subunits is crucial for regulating NucS activity (Figure 5C).

Figure 6A shows two models of a possible interaction of the flapped (Model I) and splayed (Model II) DNA substrates with NucS dimers based on our experimental data. These models emphasize the contacts of NucS with single-stranded regions of the DNA substrates. Although the NucS dimer carrying the two binding sites is sufficient for processing both the 3' and 5' extremities of ssDNA, our dynamic light scattering experiments indicate that two NucS dimers may bind to ssDNA substrate. This observation suggests that the binding of two individual NucS dimers to the single-stranded 5' and 3' extremities could occur simultaneously, or, alternatively, *P. abyssi* NucS dimers could form a ring-shaped nuclease assembly similar to that observed in the crystals. Strikingly, the observed "tetrameric" structure [i.e. dimer of dimers (Figure 3)] somewhat resembles the ring-shaped structures of the sliding clamps of DNA polymerases and the prokaryotic DNA repair mediator protein RecR (Hingorani and O'Donnell, 1998; Lee et al., 2004)**.** Thus, the topological restraints brought about by stabilization of a ring structure may increase the processivity of nuclease activity by ensuring that the protein remains bound to its substrate (Hingorani and O'Donnell, 1998).

Although the molecular function of NucS proteins waits for further exploration, the genomic context data of *Pyrococcus* and *Thermococcus* species *nucS* (Figure 1A) have raised a possibility that hyperthermophilic NucS orthologs could function in recombinatorial DNA repair. Moreover, *P. abyssi* NucS interacts strongly with the replication clamp both *in vitro* and *in vivo*

(Figure 1C) and our pull-down experiments (Figure S1) suggested that the NucS proteins may interact directly or indirectly with Hef, an archaeal FANCM ortholog. *P. furiosus* Hef nuclease domain is thought to participate in endonuclease processing of double stranded regions of stalled replication forks (Komori et al., 2004; Nishino et al., 2005a; Nishino et al., 2005b) and vertebrate orthologs promote branch migration of replication forks (Gari et al., 2008). NucS proteins may thus play a specialized role in processing of ssDNA extremities created by archaeal Hef proteins (Komori et al., 2004). Finally, gene expression omnibus records GDS2677 and GDS326 (www.ncbi.nih.gov) indicate that in *M. tuberculosis* NucS (Rv1321) is expressed in stressed cells.

We propose a functional model of how the interactions of *P. abyssi* NucS with PCNA could assist in appropriate processing of the splayed arm substrate (Figure 6B). First, the sliding replication clamp PCNA, in association with other proteins, is bound at the duplex region of the splayed arm substrate, whereas NucS dimers are not chromatin-bound. Then, the high affinity of *P. abyssi* NucS for PCNA and ssDNA could permit the loading of NucS at the single stranded unpaired extremities of the model substrate. After that, PCNA and the high affinity ssDNA binding site I will direct the cleavage at the 3' protruding strand to the junction of ss/dsDNA. The effect of PCNA in increasing cleavage specificity likely results from inhibition of the melting of base pairs at the vicinity of the junction (Figures 5D and S6). As the concave surface of the *P. abyssi* NucS dimer is predominantly basic, and duplex regions modulate NucS activity (Figure 5), NucS may, together with PCNA, interact with the duplex DNA. Note also that the very high affinity of NucS for single-stranded DNA could allow direct chromatin loading of NucS without PCNA, as well as dissociation of RPA (or other proteins) from ssDNA.

In conclusion, we have identified a novel class of DNA endonucleases that in some species co-operates with the replication clamp and possibly with additional proteins [such as the Hef, RF-C and topoisomerase VI (not shown in Figure 6B)] in the processing of branched DNA

structures. Strikingly, NucS homologs can be found not only in archaea, but also in bacteria, suggesting unexpected functional parallels between archaeal and bacterial species (particularly in high GC% Gram positive bacteria).

Materials and methods

Further experimental details are given in supplemental material.

Protein purification, site-directed mutagenesis and molecular genetic techniques

P. abyssi PCNA and NucS (PAB2263) were purified as described previously (Meslet-Cladiere et al., 2007; Ren et al., 2007). Site-directed mutagenesis used the single primer mutagenesis kit (Stratagene) following manufacturer's instructions. Sequences of all oligonucleotides used are indicated in Supplemental table 2.

Surface Plasmon resonance experiments

SPR analyses were performed in a BIAcore X apparatus (Biacore, Sweden). The interactions were monitored at 25 ^{o}C at a flow rate of 30 µl/min in HBS-P buffer [10 mM Hepes, pH 7.4; 150 mM NaCl and 0.005% (v/v) P20]. Experiments with the PCNA chip were performed as already described (Meslet-Cladiere et al., 2007). When DNA was fixed on a chip, 20 nM of 3' biotinylated single strand oligonucleotide was injected over a Streptavidin chip (Biacore) during 2 minutes at the flow rate 5 µl/min. The affinity constant (K_D) values were estimated from the simultaneous fit of association and dissociation phases using the Biaevaluation Software version 3.2, assuming 1:1 binding model. Pull-down experiments were essentially performed as previously (Meslet-Cladiere et al., 2007).

Dynamic light scattering and sucrose gradient centrifugation

For dynamic light scattering experiments, NucS (1mg/mL) in 50 mM Hepes pH 8, 600 mM NaCl was used. When DNA was present in the samples, 0.2 nmoles of a 87 bases single strand oligonucleotide was added to obtain a NucS (monomer): DNA ratio of 4 : 1. The samples were filtered through a 0.1-µm filter (Ultrafree-MC, millipore). A 25 µl aliquot was transferred to the sample container in the Zetasizer Nano ZS (Malvern Instruments Ltd). Dynamic light

scattering was measured at a temperature of 20 °C (2 individual rounds of 11 runs in 10 seconds). The results shown are the values for Z-averages of the hydrodynamic size corresponding to intensity weighted mean diameter. Number-distributions were calculated using the integrated software (Malvern). The size of free oligonucleotide was measured within the experimental pH and salt range as controls. The molecular weight of NucS assemblies were estimated using a calibration curve plotting the Z-averages against the molecular weight of cytochrome C (12,4 kDa), carbonic anhydrase (29 kDa), stanniocalcin (54 kDa), BSA (66 kDa) and a glycoside hydrolase (100 kDa). 5%-20% sucrose gradients were run in 50 mM HEPES, pH 8, 300 mM NaCl using a SW41 rotor (35 000 RPM, 43 hours, +4 °C). 250 µl fractions were collected from the top of the gradients. The proteins were localized and quantified using A_{280} readings and SDS-PAGE. Carbonic anhydrase (29 kDa), BSA (66 kDa) and PBCV-1 ThyX (92 kDa) were used as standards.

Crystallization and data collection

Due to the low content of methionines (3 in 257 residues) and local structural disorders, an initial attempt of structural solution by the multi-wavelength anomalous diffraction method (MAD) was not successful using seleno-methionine substituted wild type protein. To improve the anomalous diffraction signal, a triple mutant (I89M/L188M/L246M) was prepared using site-directed mutagenesis. The SeMet-substituted mutant was crystallized by the sitting-drop vapour diffusion method at 293 K by mixing 2 µl of protein at 4 mg/ml (30 mM HEPES pH 8.0 and 0.57 M NaCl) with 2 µl reservoir solution (0.1 M Tris-HCl pH 8.5, 0.2 M ammonium sulfate and 25 % PEG 3350). A three-wavelength Se MAD data set was collected at 100 K on beamline BM14 at ESRF (Grenoble, France) with a MAR225 CCD detector (Table 1). The reservoir solution, added with 20 % (v/v) glycerol, was used as the cryo-protectant. The data were processed using the

HKL2000 package. The crystals belonged to the space group $C222_1$ with cell dimensions of a = 78.9 Å, b = 100.6 Å and c = 157.6 Å.

Structure determination and refinement

The structure of *P. abyssi* NucS triple mutant was solved by MAD using the program suite CNS(Brunger et al., 1998). Heavy atom search with anomalous Patterson function calculation identified eight sites using data up to 3.2 Å, corresponding to four sites per subunit of the dimer that constitutes the crystallographic asymmetric unit. There were still two sites that could not be located in each subunit due to structural disorders. While phase calculation at 3.2 Å gave an overall figure of merit of 0.68, density modification was employed to further improve and extend the phases to 2.8 Å resolution. The calculated electron density was of good quality, enabling the tracing of polypeptide chain.

Model building was carried out with the program O (Jones et al., 1991). The models were refined with REFMAC5 in CCP4 (Murshudov et al., 1997), using a second data set collected at the remote wavelength. 5% of the diffraction data were set apart to monitor the free R-factor (R_{free}) throughout the refinement process. Electron density averaging was employed using local 2-fold symmetry to improve the maps. Strict non-crystallographic symmetry restrains were applied during refinement. The regions that were apparently disordered and not visible in the electron density maps include the N-terminal His-tag and C-terminal residues of 234-251 for both subunits and residues of 1-2 for subunit B. Weak and discontinuous electron densities were observed for the residues of 121 to 125, which constitute a part of an extended flexible linker. These five residues were set to alanine or glycine according to the protein sequence and their weights were put to zero during refinement. The final model of the dimer consists of 464 residues and 124 waters with a calculated R-factor of 22.6 % and R_{free} of 23.8 % for all the diffraction data

up to 2.6 Å resolution. Structure determination and refinement statistics for the triple mutant are shown in Table 1.

Nuclease and EMSA assays

Nuclease assays at 46 or 56 ^{o}C were performed using substrates previously described (Komori et al., 2004), with reduced incubation times (10 minutes). Reaction mixtures (20 μl in all experiments) contained *P. abyssi* NucS and labeled oligonucleotides in 20 mM HEPES, pH 8; 50 mM NaCl and 5 mM $MgCl_2$. $MnCl_2$ was able to replace $MgCl_2$ required for catalytic activity. Reactions were analyzed on 20% denaturing polyacrylamide gels. The effect of the PCNA on the NucS nuclease activity was quantified through formation of specific cleavage product, reflecting the cleavage at the ss/ds DNA junction.

Reaction mixtures (20 μl in all experiments) for EMSA assays contained 10 mM Tris, pH 7.5, 20% (v/v) glycerol, 1 mM EDTA, (Komori et al., 2004) 100 μg/ml of bovine serum albumin and 1 mg/ml of Orange G. DNA-protein complexes were formed during 30 minute incubation at 40 ^{o}C. Magnesium was omitted from reaction mixtures to prevent degradation of oligonucleotide substrates that were marked either with ^{32}P or Cy5.5 fluorophor. Native 10% polyacrylamide gels were run in the cold room. Gel images were acquired, analyzed and quantified using either Storm Phosphorimager (Amersham) or Licor Odyssey imaging systems.

Accession Numbers

The atomic coordinates of *P. abyssi* NucS (PAB2263) have been deposited with the Protein Data Bank (accession code 2VLD).

Acknowledgements

We thank M. Czjzek (Station Biologique de Roscoff) for advice with the experiments and helpful discussions and Taku Yamashita and Andreea Sodolescu for help with art work. This work was supported by European Union REPBIOTECH program to R. L and H. M (contract number QLK3-CT-2002–02071) and the ANR program ARCREP to D.F and H.M. D.F. thanks the French National Institute of Marine Genomics for financial support (contract number: FNS-500-370). H. M. also acknowledges INSERM AVENIR program and Fondation Bettencourt Schueller for financial support.

FIGURE LEGENDS:

Figure 1: *Biochemical characterization of P. abyssi NucS (PAB2263)*: *(A)* The domain structure of the *P. abyssi* NucS protein. The conserved residues of the RecB family nucleases are shown. Note also that protein contains the putative PCNA interacting peptide (PIP) at its extreme carboxy-terminus. In addition, the genomic organization of the *P. abyssi* (*Pab*) NucS in the vicinity of the replication origin (*oriC*) and a likely translational coupling of *Thermococcus kodakarensis* KOD1_4 (Tkod) NucS with the RadA recombinase are shown below the dotted line. *(B)* The proteins used in this study: 1 μg of *P. abyssi* NucS wild type and mutant proteins were separated on 11 % SDS-PAGE. *(C)* SPR experiment on the PCNA chip. Indicated amounts of *P. abyssi* NucS were injected at time 0 over the PCNA chip. All data are indicated in resonance units (RU) as a function of time. The background resulting from the buffer injection was subtracted from the observed signals. The inset shows the pull-down experiment using *P. abyssi* NucS (PAB2263) attached into magnetic beads. Western blot analysis of PCNA using antibody raised against *P. abyssi* PCNA is shown. *(D)* Dynamic light scattering *experiment* of *P. abyssi* NucS (1 mg/ml) with and without a single stranded oligonucleotide. Mean particle diameters 12.5 nm and 7.5 nm were observed with and without single stranded oligonucleotide, respectively. *(E)* Sucrose gradient centrifugation analysis of the *P. abyssi* NucS. 1 mg of protein was loaded on the top of 5%-20% sucrose gradient in 300 mM NaCl (pH 8). Gradients indicate that the *P. abyssi* NucS co-sediments with BSA (66 kDa).

Figure 2: *The structure of P. abyssi NucS.* *(A)* The dumbbell-like two domain structure of PAB2263 subunit. A water molecule at the C-terminal nuclease active site is shown as a ball. *(B)* The β-barrel structure of the N-terminal domain, with a half-closed bottom of one side and gathered loops on the opposite side that form a shaped ssDNA binding site. *(C)* Superposition of

the NucS C-terminal domain and *E.coli* RecB nuclease domain, which are colored in light brown and blue, respectively. The bound magnesium ion in RecB and a water molecule in NucS at the active sites are shown as balls. *(D)* The assembly of dimer and tetramer in the crystal due to crystallographic contacts. Each subunit chain is colored differently to illustrate domain swapping. The side chains of the conserved W75 and nearby basic residues are shown. R42 sticks out from dimer surface into the elliptical hole of the tetramer. All structural images in this publication were created by the program Pymol (DeLano, W.L. 2002. The PyMOL Molecular Graphics System, http://www.pymol.org).

Figure 3: *The ssDNA binding sites. (A)* Two different ssDNA binding sites (site I and II) are formed in the dimer. The basic residues surrounding the binding sites and a water molecule at the nuclease active sites are displayed. *(B)* The concave surface of the dimer, viewed with a 90° rotation relative to (A). The site I groove is outlined with dotted lines and the positions of the charged residues surrounding the ssDNA-binding sites are highlighted. Residues from a different subunit are marked by asterisks. *(C)* A close view of the nuclease channel together with a part of site I. The nuclease active site residues in the channel are shown, as well as W75 and Y39 of site I that may forge stacking interactions with DNA.

Figure 4: *ssDNA binding studies of P. abyssi NucS (A)* SPR study using a ssDNA chip with an 87-mer oligonucleotide that was immobilized onto a Streptavidin chip through a biotin moiety located at the 3' extremity. Various amounts of NucS were injected over the chip at the time 0. Binding data are reported in RU as a function of time. *(B)* EMSA assay (20 µl) using an oligonucleotide marked with a Cy5.5 fluorophor. Dotted gray lines indicate the 95% confidence interval for the non-linear fit assuming 1:1 binding. *(C)* Quantification of ssDNA binding for indicated mutant proteins are shown. Experiments were performed with two different protein

concentrations using 1 pmol of an oligonucleotide indicated in panel C. Results shown are averages of two independent measurements *(D)* Structure guided amino acid sequence alignment of the representative NucS orthologs from archaea and bacteria. The secondary structural regions of *P. abyssi* NucS (PAB2263) are indicated. The homologous regions of the sequences are boxed and the conserved residues are shaded. The active sites residues in the conserved sequence motif of RecB-like nuclease family are indicated by filled triangles. The positions that have been tested by mutagenesis are indicated by filled circles. Note that the PCNA binding peptide in the C-terminus of *P. abyssi* NucS is not conserved in all orthologs. The picture was prepared using the program ALSCRIPT (Barton, 1993).

Figure 5: *Nuclease activity of P. abyssi NucS with model substrates.* The sizes of the predicted reaction products are also indicated at the left side of each panel. *(A)* NucS activity on the 5' flap (27 nucleotides) using 1 pmol in 20 μl of non-modified substrate depicted at the bottom of the figure (the left panel). Reactions were performed with indicated protein amounts. The right panel shows an experiment where the 3'-biotin moiety had reacted with the streptavidin. The asterisk refers to a ^{32}P label used to mark the substrate. The oligo names include their lengths in nucleotides. *(B)* NucS activity on the 3' flap (22 nucleotides). The experiment was performed as indicated in Panel A. *(C)* NucS activity on a single stranded oligonucleotide (the left panel) and the splayed arm substrate (the middle and the right panel). Where indicated, experiments were performed using either wild-type or the mutant R42A W75S that does not bind ssDNA. *(D)* PCNA increases the cleavage specificity of *P. abyssi* NucS at the ssDNA/dsDNA junction. 2.5 pmol of NucS, 10 pmol of PCNA and 10 pmol of BSA were included in reaction mixtures. The bars indicate the standard deviations of three independent measurements.

Figure 6: *(A)* Hypothetical models showing the binding of a flapped (Model I) and splayed (Model II) DNA substrate by the *P. abyssi* NucS dimer. Note that in these models only the single-stranded region of the DNA substrates makes close contacts with NucS. The surface areas corresponding to W75 in the DNA binding site I are colored in blue. *(B)* A proposed mechanism suggesting the coordination between PCNA and NucS during the processing of the splayed arm substrate. For detailed description of distinct steps, see text.

Table 1. Structure determination and refinement statistics.

	Refinement dataset	inflection	peak	remote
Data collection				
Wavelength (Å)	0.9077	0.9789	0.9787	0.9077
Resolution range (Å)	25 – 2.6	25 – 2.8	25 – 2.8	25 – 2.8
Outmost shell (Å)	(2.7 – 2.6)	(2.9 – 2.8)	(2.9 – 2.8)	(2.9 – 2.8)
Measured reflections	94709	116432	116358	116564
Unique reflections	19639	16090	15977	16030
Average redundancy	4.8 (4.7)	7.2 (7.1)	7.3 (7.3)	7.3 (7.2)
Average $I / \sigma (I)$	37.3 (2.9)	39.9 (2.6)	45.2 (4.5)	44.1 (3.6)
Completeness (%)	99.7 (99.4)	100 (100)	99.9 (100)	99.9 (99.9)
R_{merge} (%)	5.3 (41.0)	5.7 (60.2)	6.7 (38.0)	5.6 (43.8)
Phasing				
Resolution cut-off (Å)		3.2	3.2	3.2
Phasing power		1.72	1.49	1.38
Figure of merit	0.68			
Refinement				
Resolution range (Å)	25 – 2.6			
R-factor/R_{free} (%)	22.6 / 23.8			
Average B-factor (Å^2)	50.3			
No. of protein atoms	3727			
No. of water molecules	124			
R.m.s.d. bond length (Å)	0.016			
R.m.s.d bond angle (°)	1.7			
Ramachandran plot with residues in:				
most favored regions (%)	84.2			
additionally allowed regions (%)	15.8			

REFERENCES:

Aravind, L., Makarova, K.S. and Koonin, E.V. (2000) SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res*, 28, 3417-3432.

Barton, G.J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng*, 6, 37-40.

Brunger, 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., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, 54, 905-921.

Chapados, B.R., Hosfield, D.J., Han, S., Qiu, J., Yelent, B., Shen, B. and Tainer, J.A. (2004) Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell*, 116, 39-50.

Ciccia, A., Ling, C., Coulthard, R., Yan, Z., Xue, Y., Meetei, A.R., Laghmani el, H., Joenje, H., McDonald, N., de Winter, J.P., Wang, W. and West, S.C. (2007) Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. *Mol Cell*, 25, 331-343.

Cubeddu, L. and White, M.F. (2005) DNA damage detection by an archaeal single-stranded DNA-binding protein. *J Mol Biol*, 353, 507-516.

Ehmsen, K.T. and Heyer, W.D. (2008) Saccharomyces cerevisiae Mus81-Mms4 is a catalytic, DNA structure-selective endonuclease. *Nucleic Acids Res*, 36, 2182-2195.

Gari, K., Decaillet, C., Stasiak, A.Z., Stasiak, A. and Constantinou, A. (2008) The fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks. *Mol Cell*, 29, 141-148.

Hingorani, M.M. and O'Donnell, M. (1998) Toroidal proteins: running rings around DNA. *Curr Biol*, 8, R83-86.

Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*, 47 (Pt 2), 110-119.

Kambach, C., Walke, S., Young, R., Avis, J.M., de la Fortelle, E., Raker, V.A., Luhrmann, R., Li, J. and Nagai, K. (1999) Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell*, 96, 375-387.

Kelman, Z. and White, M.F. (2005) Archaeal DNA replication and repair. *Curr Opin Microbiol*, 8, 669-676.

Komori, K., Hidaka, M., Horiuchi, T., Fujikane, R., Shinagawa, H. and Ishino, Y. (2004) Cooperation of the N-terminal Helicase and C-terminal endonuclease activities of Archaeal Hef protein in processing stalled replication forks. *J Biol Chem*, 279, 53175-53185.

Lee, B.I., Kim, K.H., Park, S.J., Eom, S.H., Song, H.K. and Suh, S.W. (2004) Ring-shaped architecture of RecR: implications for its role in homologous recombinational DNA repair. *Embo J*, 23, 2029-2038.

Meetei, A.R., Medhurst, A.L., Ling, C., Xue, Y., Singh, T.R., Bier, P., Steltenpool, J., Stone, S., Dokal, I., Mathew, C.G., Hoatlin, M., Joenje, H., de Winter, J.P. and Wang, W. (2005) A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet*, 37, 958-963.

Meslet-Cladiere, L., Norais, C., Kuhn, J., Briffotaux, J., Sloostra, J.W., Ferrari, E., Hubscher, U., Flament, D. and Myllykallio, H. (2007) A novel proteomic approach identifies new interaction partners for proliferating cell nuclear antigen. *J Mol Biol*, 372, 1137-1148.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, 53, 240-255.

Myllykallio, H., Lopez, P., Lopez-Garcia, P., Heilig, R., Saurin, W., Zivanovic, Y., Philippe, H. and Forterre, P. (2000) Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science*, 288, 2212-2215.

Newman, M., Murray-Rust, J., Lally, J., Rudolf, J., Fadden, A., Knowles, P.P., White, M.F. and McDonald, N.Q. (2005) Structure of an XPF endonuclease with and without DNA suggests a model for substrate recognition. *Embo J*, 24, 895-905.

Nishino, T., Komori, K., Ishino, Y. and Morikawa, K. (2005a) Structural and functional analyses of an archaeal XPF/Rad1/Mus81 nuclease: asymmetric DNA binding and cleavage mechanisms. *Structure*, 13, 1183-1192.

Nishino, T., Komori, K., Tsuchiya, D., Ishino, Y. and Morikawa, K. (2005b) Crystal structure and functional implications of Pyrococcus furiosus hef helicase domain involved in branched DNA processing. *Structure*, 13, 143-153.

Pingoud, A., Fuxreiter, M., Pingoud, V. and Wende, W. (2005) Type II restriction endonucleases: structure and mechanism. *Cell Mol Life Sci*, 62, 685-707.

Read, R.J. (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr D Biol Crystallogr*, 57, 1373-1382.

Ren, B., Kuhn, J., Meslet-Cladiere, L., Myllykallio, H. and Ladenstein, R. (2007) Crystallization and preliminary X-ray analysis of a RecB-family nuclease from the archaeon Pyrococcus abyssi. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, 63, 406-408.

Roberts, J.A., Bell, S.D. and White, M.F. (2003) An archaeal XPF repair endonuclease dependent on a heterotrimeric PCNA. *Mol Microbiol*, 48, 361-371.

Roberts, J.A. and White, M.F. (2005) An archaeal endonuclease displays key properties of both eukaryal XPF-ERCC1 and Mus81. *J Biol Chem*, 280, 5924-5928.

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.

Theobald, D.L., Mitton-Fry, R.M. and Wuttke, D.S. (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct*, 32, 115-133.

Tsodikov, O.V., Enzlin, J.H., Scharer, O.D. and Ellenberger, T. (2005) Crystal structure and DNA binding functions of ERCC1, a subunit of the DNA structure-specific endonuclease XPF-ERCC1. *Proc Natl Acad Sci U S A*, 102, 11236-11241.