An Archaeal Endonuclease Displays Key Properties of Both Eukaryal XPF-ERCC1 and Mus81*

Jennifer A. Roberts and Malcolm F. White‡
From the Centre for Biomolecular Sciences, University of St. Andrews, KY16 9ST, Scotland, United Kingdom

Structure-specific nucleases of the XPF/Mus81 family function in several DNA recombination and repair pathways in eukaryotes, cleaving a variety of flap and branched DNA substrates. Mus81 and XPF are clearly related evolutionarily but differ markedly in their substrate specificity and protein partners. We demonstrate that the XPF endonuclease from Sulfolobus solfataricus, which is dependent on the sliding clamp proliferating cell nuclear antigen for activity, represents an ancestral form of the XPF/Mus81 family, with key properties in common with both enzymes. The archaeal XPF has a domain organization and sequence preference very similar to eukaryal XPF-ERCC1. However, the archaeal enzyme has a pronounced preference for Mus81-type substrates such as D loops, nicked four-way junctions, and 3’ flaps. These all have in common a 5’-DNA end next to the cleavage site. The availability of the sliding clamp proliferating cell nuclear antigen may dictate the activity of Sulfolobus XPF in vivo.

Branched DNA structures are created in many circumstances, arising both as a direct consequence of DNA damage and as intermediates in cellular DNA recombination and repair pathways. These structures must be processed correctly by structure-specific nucleases prior to DNA replication. In eukaryotes, the related nucleases XPF-ERCC1 and Mus81-Eme1/Mms4 (abbreviated here to Mus81*) process a variety of DNA substrates with 3’ branches and bubbles (see Table II for a definition of the DNA substrates). The cellular functions of these important enzymes are still being delineated, as is the molecular basis for their interaction with DNA substrates.

Most euryarchaeae encode a clear homolog of eukaryal XPF, including the N-terminal helicase and C-terminal nuclease domains. The homolog from Pyrococcus furiosus, Hef, cleaves a variety of flap and structure that are substrates for eukaryal Mus81* (1). The crystal structure of the nuclease domain of Pyrococcus Hef unveiled the relationship of this domain with a superfamily of nucleases including the restriction enzymes, Holliday junction resolving enzymes Hjc and T7, endonuclease I, and other DNA repair enzymes (2). Crenarchaeal XPF homologs lack the N-terminal helicase domain altogether and consist only of the nuclease and HhH2 domains (see Fig. 3). In this respect, they may represent the simplest and perhaps ancestral version of the XPF/Mus81 nuclease. We showed previously that the XPF homolog from Sulfolobus solfataricus (SsoXPF) cuts flap or splayed duplex structures (3) and that its activity is crucially dependent on an association with the sliding clamp PCNA, via a C-terminal PCNA interaction motif (3). Here we report the detailed kinetic analysis of the substrate specificity of Sulfolobus XPF. We demonstrate key similarities with both the eukaryal XPF and Mus81* nucleases.

MATERIALS AND METHODS

Recombinant Proteins— S. solfataricus XPF and PCNA heterotrimer were expressed and purified as described previously (3, 4).

DNA Substrates—The oligonucleotides used for substrates (Table I) were 5’-32P-end labeled and assembled into various structures (see Table II) by slow cooling from 85 °C to room temperature for 3 h or overnight. Substrates were purified on a native 6% acrylamide gel followed by electro-elution and ethanol precipitation. Size markers (G and A) were prepared from labeled substrates using standard protocols.

Endonuclease Assays—Nuclease reactions (10 μl) were assembled in 30 mM HEPES, pH 7.6, 5% glycerol, 40 mM KCl, 0.1 mg/ml bovine serum albumin, and 0.1 mg/ml calf thymus DNA with 80 mM DNA substrate and 1 μM XPF-PCNA and equilibrated at 55 °C. Cleavage was initiated by adding MgCl2 to a final concentration of 10 mM, mixed briefly, and incubated at 55 °C. 5-μl aliquots were taken at selected time points and added to chilled stop solution (10 mM Tris-Cl, pH 8, 10 mM EDTA, 0.1 mg/ml calf thymus DNA) to terminate the reaction. DNA was ethanol-precipitated and analyzed on a denaturing polyacrylamide/urea/Tria-borate-EETDA gel. Uuncut substrate and products were quantified (Imaging System, Fuji), and the reaction rate was obtained by linear regression. In the experiment, to determine the influence of a 5’-phospho group on activity, the PCNA-XPF concentration was reduced to 0.1 μM to allow more accurate quantitation.

RESULTS

First, we determined the structure specificity and requirements of Sulfolobus XPF. To pin down the function of structure-specific nucleases, it is necessary to define their specificity for the different branched DNA structures that arise in vivo due to DNA damage and repair pathways. Quantitative data for eukaryal XPF-ERCC1 and Mus81* are not yet available. We have carried out a systematic quantitative analysis of the substrate preference of SsoXPF using a panel of branched DNA substrates (Table II). In each case, the sequence of the strand cleaved by XPF is the same, although its length varies in some cases.

In general terms, the order of substrate preference of SsoXPF matches that of the Mus81* family rather than the eukaryal XPF family. The best substrate is a 3’ flap followed by a nicked duplex, nicked three-way junctions, and nicked four-way junctions. The efficient cleavage of a nicked duplex substrate demonstrates that no single-stranded flap is required, rather a discontinuity in the DNA duplex is enough to promote the nuclease activity. These DNA species are all the preferred

* This work was supported in part by the Biotechnology and Biological Sciences Research Council. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ A Royal Society University Research Fellow. Tel.: 44-1334-463432; Fax: 44-1334-462595; E-mail: mfw2@st-and.ac.uk.

1 The abbreviations used are: SsoXPF, the XPF homolog from Sulfolobus solfataricus; PCNA, proliferating cell nuclear antigen; NER, nucleotide excision repair.
substrates of Mus81* and share in common a 5’-DNA end near the branch point or point of discontinuity that has been shown to be essential in directing the activity of Mus81* (5). The 3’ flap and nicked three-way substrates are analogs of a stalled and reversed replication fork, respectively (6). In contrast, a substrate analog of an active replication fork with a 3’-DNA end near the branch point is cut much more slowly, again in common with Mus81* (6). A splayed duplex species, the preferred substrate of eukaryotic XPF-ERCC1, is cleaved 10-fold more slowly than a 3’ flap substrate. Lastly, four-way junctions are very poor substrates for SsoXPF, cut more than 2 orders of magnitude more slowly than the preferred substrates.

The clear preference of SsoXPF for a 5’-DNA end near the cleavage site mirrors that reported previously for Mus81* from Saccharomyces cerevisiae and Schizosaccharomyces pombe (5, 6). In vivo, most substrates of this type would have a phosphate group at this 5’ position, although this was not present in the synthetic substrates used for Mus81*. To test whether phosphorylation of this terminus was important for the activity of SsoXPF, we compared the rates of cleavage with a nicked duplex substrate with and without a 5’-phosphate at the nick (Fig. 1). We found a measurable although modest increase in the cleavage rates of a phosphorylated substrate as compared with a non-phosphorylated one, suggesting that the archaeal XPF (and probably also Mus81*) does not require the presence of a 5’-phosphate to recognize this DNA end.

Recent studies of Mus81* suggest that it can play an important role in the generation of crossover recombination events by cleavage of D loop (displacement loop) substrates formed as a consequence of strand invasion after a chromosome sustains a double strand break (7, 8). A D loop bubble is a poor substrate for SsoXPF in the absence of an invading 3’-ssDNA end (Fig. 2B, left). This structure is equivalent to an NER bubble, the favored substrate of eukaryal XPF-ERCC1 (9). Weak cleavage of the b75 strand is seen at the splayed duplex junction (Fig. 2B, right).
**Fig. 2.** *Sulfolobus* XPF cleaves a D loop substrate. *A*, schematic diagram showing the D loop strand nomenclature used under "Results." Open circles indicate 5' DNA ends. *B*, denaturing polyacrylamide gel showing the cleavage of the D loop DNA substrate. The schematics above the gel represent the substrate, with the labeled strand indicated by a black circle at its 5' end. The major cleavage sites are indicated with black and gray arrows. Minor cleavage of the splayed duplex junction on strand b75 is indicated by the white arrow. The time points were 1, 5, and 10 min; *m*, A and G size markers of the labeled strand; *control c1*, DNA alone; *c2*, DNA + SsoXP3; *c3*, DNA + PCNA. *C*, schematic representation of the central 60 bp of the D loop substrate showing the invading strand in bold. Black and gray arrows indicate the major sites of cleavage by SsoXP3. All strand scissions are mapped 3' to pyrimidine residues. *D*, a schematic of D loop formation and cleavage by *Sulfolobus* XPF. Following a double strand break of the gray DNA duplex and generation of 3' tails, strand invasion generates substrates for SsoXP3, which cuts most efficiently at site a. Second end capture and DNA synthesis generate a second SsoXP3 substrate (a 3' flap or nicked Holliday junction) at c. Cleavage at sites a and c, coupled with strand exchange, gap filling, flap removal, and ligation, would generate crossover recombinant products. Cleavage at site b may be disfavored in *vivo* due to the requirement for PCNA, which would be loaded during DNA replication near sites a and c.
Fig. 3. Structure and evolution of the XPF/Mus81 nuclease family. The conserved nuclease (nuc) domain is found in many other nucleases involved in DNA restriction and repair pathways. In the XPF/Mus81 family, the nuclease domain is linked to a helix hairpin helix (HhH2) DNA binding domain. The archaeal XPF enzymes are homodimers, whereas the eukaryotic enzyme has duplicated to form heterodimeric proteins, allowing subunit specialization, and to form nucleases with differing substrate specificity. The archaeal and Mus81 enzymes show a distinct preference for substrates with a 5’ DNA terminus (red circle), whereas eukaryal XPF-ERCC1 has specialized for cleavage of splayed duplex or bubble substrates.

2B, white arrows, rate 0.02 min⁻¹. The z75 strand is also cut weakly at the equivalent site on the other side of the bubble (rate 0.054 min⁻¹). Once the invading strand is present, however, SsoXPF is strongly activated to cut the b75 strand at a new position (Fig. 2B, a, and C, black arrows). The invading strand creates a good substrate for SsoXPF, as observed for Mus81*, probably stimulated by the 5’-DNA end paired with the invading strand (Fig. 2A, 5’ end of oligo r26–50) (7). The rate of cleavage at this new position is at least 50-fold faster than those for the substrate without the invading strand. Cleavage at site b on the z75 strand (Fig. 2B, b, and C, gray arrows) is also stimulated by the presence of the invading duplex but is cut significantly more slowly than site a. Second end capture and DNA synthesis result in the formation of a 3’ flap substrate (site c) that is the most preferred substrate for SsoXPF (Table II). This is the second site proposed by Osman et al. (7) for Mus81*. Cleavage at sites a and c, coupled with capture of the second strand, DNA synthesis, and strand exchange, would result in crossover recombination products (Fig. 2D). Any 3’ or 5’ flaps generated as a result of this process could be removed by XPF or Fen1. Cleavage at site a alone would dissociate a D loop and could be functionally relevant for resetting of collapsed replication forks (10). Every strand scission was mapped 3’ to pyrimidine residues, confirming that although SsoXPF cuts preferentially at Mus81*-type substrates, it retains key characteristics of XPF-ERCC1.

**DISCUSSION**

In addition to the clear similarities in their domain organization (Fig. 3), the archaeal XPF has the same sequence preference for cleavage at pyrimidines as eukaryal XPF (3). In contrast, Mus81* has little or no sequence dependence (5). SsoXPF has a significant activity against splayed duplex substrates that are found at the 5’ side of lesions repaired by the eukaryal XPF protein during NER. These substrates are cut very poorly by Mus81*. Taken together, these data suggest a close relationship between the eukaryal and archaeal XPF enzymes.

On the other hand, our kinetic analysis of the substrate preference of SsoXPF demonstrates clearly that the best substrates resemble stalled replication forks, D loops, and nicked Holliday junctions, with a 5’-DNA end near the branch point or strand discontinuity (Table II and Fig. 3), a feature in common with Mus81*. Cleavage of 3’ flap substrates by Mus81-Mms4 is strongly stimulated by the presence of a 5’ terminus within 4 nucleotides of the branch point (5). The highly efficient cleavage of a nicked duplex substrate demonstrates that the archaeal XPF protein seeks out strand discontinuities that allow DNA deformation, flanked on either side by duplex DNA, rather than branched DNA substrates per se. This is also the case for Mus81*, which cleaves nicked duplexes efficiently to generate 5-nucleotide gapped duplexes (5). These substrates are cut very poorly by eukaryal XPF-ERCC1. D loop substrates are also cut by SsoXPF and Mus81* in a similar way, with strong stimulation observed when strand invasion occurs.

Therefore, the archaeal XPF clearly has the potential to act as both a Mus81-like flap endonuclease and an XPF-like repair endonuclease. Previous comparisons of the activity of Mus81-Mms4 with Rad1-Rad10 have come to the conclusion that the specificities of the two enzymes do not overlap. The presence of a 5’-DNA terminus within 4 nucleotides of the branch point results in strong stimulation of Mus81-Mms4 and reduces the activity of Rad1-Rad10 such that the two enzymes have at least a 100-fold preference for their specific substrates (5). SsoXPF bridges these two diverged proteins and likely represents the ancestral form of both enzymes.

Hyperthermophiles such as *S. solfataricus* suffer a higher frequency of DNA damage such as hydrolytic deamination, oxidation, and single strand breaks as a consequence of their higher growth temperatures (11). Many of these lesions are known to stall or collapse replication forks; for example, a single strand break is converted to a double strand break if replicated. Stalled and collapsed forks are a frequent event in bacteria and eukarya, which consequently have efficient mechanisms to recover from these potentially catastrophic events. One mechanism for fork rescue is via the formation of a Holliday junction, which can be resolved to allow replication restart (reviewed in Ref. 12). The substrate specificity we have described for SsoXPF is consistent with a role in the cleavage of stalled replication forks to generate substrates for recombination and/or resolution of early recombination intermediates (D loops), similar to Mus81*. In this regard, the requirement of SsoXPF for PCNA as an essential cofactor may be relevant as the target sites of the enzyme may be determined in part by PCNA loading. This is well illustrated by consideration of the D loop substrate (Fig. 2D). The recombination scheme requires extension by DNA replication of both the invading 3’ strand (strand x50) and of the second DNA end after capture. This replication will result in loading of PCNA on both these duplexes. This may direct SsoXPF to cut at sites a and c, thus generating recombinogenic DNA ends. In contrast, cleavage at...
site b could be disfavored in vivo as PCNA will not be loaded on this duplex.

The data presented here do not rule out the possibility that SsoXPF also functions in an archaeal NER pathway. The function of eukaryal XPF-ERCC1 is largely determined by the protein:protein interactions it makes. In NER, this includes specific interactions between ERCC1 and XPA (13) and between XPF and the C terminus of the helicase XPB, a component of the multiprotein complex TFIH (14). Thus XPF-ERCC1 is recruited into the assembling repair complex and positioned to introduce a single strand cleavage 5′ of a DNA lesion. XPF-ERCC1 also has roles in the resolution of DNA interstrand cross-links (15, 16) and in targeted gene replacement (17). These functions link XPF-ERCC1 to recombinational pathways and necessitate recruitment of the enzyme into different DNA processing complexes from those formed during NER. Similarly, interactions of SsoXPF mediated by PCNA may help to define its role in vivo.

REFERENCES
1. Komori, K., Fujikane, R., Shinagawa, H., and Ishino, Y. (2002) Genes Genet. Syst. 77, 227–241
2. Nishino, T., Komori, K., Ishino, Y., and Morikawa, K. (2003) Structure (Camb.) 11, 445–457
3. Roberts, J. A., Bell, S. D., and White, M. F. (2003) Mol. Microbiol. 48, 361–371
4. Dionne, I., Nookala, R. K., Jackson, S. P., Doherty, A. J., and Bell, S. D. (2003) Mol. Cell 11, 275–282
5. Bastin-Shanower, S. A., Fricke, W. M., Mullen, J. R., and Brill, S. J. (2003) Mol. Cell Biol. 23, 3487–3496
6. Whitby, M. C., Osman, F., and Dixon, J. (2003) J. Biol. Chem. 278, 6928–6935
7. Osman, F., Dixon, J., Doe, C. L., and Whitby, M. C. (2003) Mol. Cell 12, 761–774
8. Hollingsworth, N. M., and Brill, S. J. (2004) Genes Dev. 18, 117–125
9. de Laat, W. L., Appeldoorn, E., Jaspers, N. G., and Hoeijmakers, J. H. (1998) J. Biol. Chem. 273, 7835–7842
10. Doe, C. L., Osman, F., Dixon, J., and Whitby, M. C. (2004) Nucleic Acids Res. 32, 5570–5581
11. Grogan, D. W. (2000) Trends Microbiol. 8, 180–185
12. McGlynn, P., and Lloyd, R. G. (2002) Nat. Rev. Mol. Cell. Biol. 3, 859–870
13. Li, L., Lu, X., Peterson, C. A., and Legerski, R. J. (1995) Mol. Cell. Biol. 15, 5396–5402
14. Evans, E., Maggs, J. G., Hwang, J. R., Egly, J. M., and Wood, R. D. (1997) EMBO J. 16, 6559–6573
15. Kuraoka, I., Kobertz, W. R., Ariza, R. R., Biggerstaff, M., Essigmann, J. M., and Wood, R. D. (2000) J. Biol. Chem. 275, 26632–26636
16. Niedernhofer, L. J., Odijk, H., Budzowska, M., van Drunen, E., Maas, A., Theil, A. F., de Wit, J., Jaspers, N. G., Beverloo, H. B., Hoeijmakers, J. H., and Kanaar, R. (2004) Mol. Cell. Biol. 24, 5776–5787
17. Niedernhofer, L. J., Essers, J., Weeda, G., Beverloo, B., de Wit, J., Muijtjens, M., Odijk, H., Hoeijmakers, J. H., and Kanaar, R. (2001) EMBO J. 20, 6540–6549
An Archaeal Endonuclease Displays Key Properties of Both Eukaryal XPF-ERCC1 and Mus81

Jennifer A. Roberts and Malcolm F. White

*J. Biol. Chem.* 2005, 280:5924-5928.
doi: 10.1074/jbc.M412766200 originally published online December 8, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M412766200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 17 references, 9 of which can be accessed free at http://www.jbc.org/content/280/7/5924.full.html#ref-list-1