6662–6670 *Nucleic Acids Research, 2005, Vol. 33, No. 20*
doi:10.1093/nar/gki974

DNA end-directed and processive nuclease activities of the archaeal XPF enzyme

Jennifer A. Roberts and Malcolm F. White*

Centre for Biomolecular Sciences, University of St Andrews, St Andrews, Fife KY16 9ST, UK

Received September 21, 2005; Revised and Accepted November 4, 2005

ABSTRACT

The XPF/Mus81 family of structure-specific nucleases cleaves branched or nicked DNA substrates and are implicated in a wide range of DNA repair and recombination processes. The structure of the crenarchaeal XPF bound to a DNA duplex has revealed a plausible mechanism for DNA binding, involving DNA distortion into upstream and downstream duplexes engaged by the two helix–hairpin–helix domains that form a dimeric structure at the C-terminus of the enzyme. A flexible linker joins these to the dimeric nuclease domain, and a C-terminal motif interacts with the sliding clamp, which is essential for the activity of the enzyme. Here, we demonstrate the importance of the downstream duplex in directing the endonuclease activity of crenarchaeal XPF, which is similar to that of Mus81-Eme1, and suggest a mechanistic basis for this control. Furthermore, our data reveal that the enzyme can digest a nicked DNA strand processively over at least 60 nt in a 3′–5′ direction and can remove varied types of DNA lesions and blocked DNA termini. This *in vitro* activity suggests a potential role for crenarchaeal XPF in a variety of repair processes for which there are no clear pathways in archaea.

INTRODUCTION

The two related structure-specific endonucleases XPF-ERCC1 and Mus81-Eme1 are found in most eukarya from yeast to humans. XPF-ERCC1 functions primarily in the Nucleotide Excision Repair (NER) pathway, where it cleaves on the 5′ side of the repair bubble formed around bulky DNA lesions such as photoproducts (1). Roles for the enzyme in recombinational repair (2), telomere maintenance (3) and processing of blocked 3′-termini (4) have also been suggested. *In vitro*, XPF-ERCC1 cleaves 3′ single strand arms from splayed duplex substrates (5). The Mus81-Eme1 endonuclease preferentially cleaves nicked Holliday junctions, D loops and 3′-flaps, and is thought to play a role in processing early recombination intermediates in the restart of stalled replication forks and during meiotic recombination (6,7). The two enzymes are related by the conservation of the core nuclease domain, which is itself a member of the nuclease superfamily that includes the type II restriction enzymes, Holliday junction resolving enzymes and many other nucleases (8).

Most archaea have a single homologue of the XPF/Mus81 endonuclease. In euryarchaea such as *Pyrococcus furiosus*, the nuclease domain is fused to an N-terminal helicase domain and a C-terminal helix–hairpin–helix (HhH_2) domain (8–11), an organization that matches that of the eukaryal XPF polypeptide. The crenarchaeal XPF homologue has a different domain arrangement, with an N-terminal nuclease domain and a C-terminal HhH_2 domain that in turn interacts specifically with the sliding clamp PCNA (12). Both forms of the archaeal enzyme are homodimeric, and the heterodimeric structure of the eukaryal enzymes probably represents an evolutionary adaptation to accommodate further protein:protein interactions (13), as only a single active nuclease domain is required (14). The substrate specificity of the archaeal XPF is more closely related to Mus81-Eme1 than to XPF-ERCC1. Similar to Mus81-Eme1, substrates with a 'downstream duplex' (Figure 1) are cleaved more efficiently than splayed duplex-type substrates (9,15,16). Recently, the crystal structure of XPF from *Aeropyrum pernix* in complex with duplex DNA has been solved. The structure suggests a model for DNA binding by the archaeal XPF whereby the dimeric HhH_2 domain binds to an upstream and downstream DNA duplex, bending the DNA substrate through 90°C and allowing a 3′ DNA flap to be cleaved at the active site of the nuclease domain (17).

The function of the archaeal enzyme *in vivo* is not clear. A few archaea have clear homologues of the UvrABC proteins

*To whom correspondence should be addressed. Tel: +44 1334 463432; Fax +44 1334 462595; Email: mfw2@st-and.ac.uk
Present address:
Jennifer A. Roberts, Cancer Research UK Beatson Laboratories, Garscube Estate, Switchback Road, Glasgow G61 1BD, UK

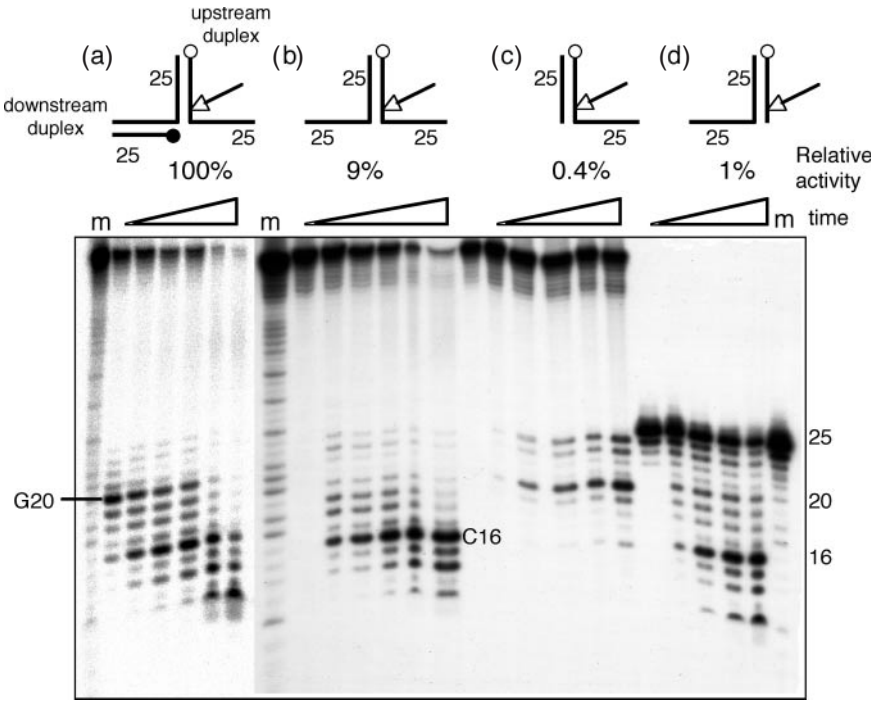

Figure 1. Cleavage of minimal DNA flap substrates by SsoXPF at 55°C. White circles show the ^{32}P-labelled 5′ DNA end of the substrate strand, and a black circle indicates the 5′ end of the downstream duplex. The arrows show the position of cleavage of the substrate. The rate of cleavage of each substrate under single turnover conditions, expressed as a percentage of the rate observed with the 3′-flap substrate (16) is shown. Time points were as follows: (**a**) 3′-flap 5, 10, 20, 40, 90, 240 s; (**b**) splayed duplex 0, 0.5, 1, 2, 5 and 10 min; (**c**) 3′-overhang 0, 1, 5, 10, 20 and 40 min; (**d**) 5′-overhang 0, 1, 5, 10, 20 min; m = A and G markers.

for bacterial NER arising from a lateral gene transfer event (18), but most lack these genes and encode homologues of several of the eukaryal-type NER enzymes, such as the helicases XPB and XPD, and the flap endonuclease Fen1, as well as XPF (19). However, an NER-type patch repair activity has not been demonstrated for archaea lacking UvrABC. A role for archaeal XPF similar to that of Mus81 in the restart of stalled replication forks has also been suggested (9). In this paper, we report that, in addition to a DNA end-directed endonuclease activity like Mus81, XPF from *Sulfolobus solfataricus* (SsoXPF) can function as a processive nuclease, digesting a DNA strand 5′ of a nick in a DNA duplex over scores of nucleotides. This activity is unaffected by the presence of a variety of DNA lesions in the substrate strand and suggests a potential novel mechanism for DNA repair in the archaea.

MATERIALS AND METHODS

Recombinant proteins

Recombinant *S.solfataricus* XPF and PCNA proteins were expressed and purified as described previously (20). Recombinant *S.solfataricus* SSB was expressed and purified according to (21).

DNA substrates

The oligonucleotides used to make the DNA structures (Table 1) were purchased from Operon, with the exception of the cyclobutane pyrimidine dimer (CPD)-containing oligo which was purchased from Phoenix BioTechnologies. The

Table 1. Oligonucleotides used for DNA substrates

Oligo	Sequence (5′ to 3′)
b25	CCTCGAGGGATCCGTCCTAGCAAGC
b50	CCTCGAGGGATCCGTCCTAGCAAGCCGCTGCT ACCGGAAGCTTCTGGACC
b50-bio	CC[bio∼dT]CGAGGGATCCGTCCTAGCAAGCCGCTGC TACCGGAAGCTTC-[bio∼dT]GGACC
b75	GGAGCGGTGGTTGAATTCCTCGACGCCTCGAGGG ATCCGTCCTAGCAAGC CGCTGCTACCGGAAGC TTCTGGACC
x26-50	GCTTGCTAGGACGGATCCCTCGAGG
x50	GCTCGAGTCTAGACTGCAGTTGAGAGCTTGCTA GGACGGATCCCTCGAGG
x60	GGCAATCCCTGCTCGAGTCTAGACTGCAGTTGAGAG CTTGCTAGGACGGA TCCCTCGAGG
x100	GCTCGAGTCTAGACTGCAGTTGAGAGGTCCAGAAG CTTCCGGTAGCAGCG GCTTGCTAGGACGGATCC CTCGAGGCGTCGAGGAATTCAACCACCGCTCC
r26-50	TCTCAACTGCAGTCTAGACTCGAGC
r36	TTCTCAACTGCAGTCTAGACTCGAGCAGGGATTGCC
r35	TCTCAACTGCAGTCTAGACTCGAGCAGGGATTGCC
r34	CTCAACTGCAGTCTAGACTCGAGCAGGGATTGCC
r33	TCAACTGCAGTCTAGACTCGAGCAGGGATTGCC
r32	CAACTGCAGTCTAGACTCGAGCAGGGATTGCC
r31	AACTGCAGTCTAGACTCGAGCAGGGATTGCC
h25	GGTCCAGAAGCTTCCGGTAGCAGCG
y35	CGGGATCGAGCACCAGAATTCACGAGTACCTGCGG
y35 CPD	CGGGATCGAGCACCAGAAT^TCACGAGTACCTGCGG
y35 F	CGGGATCGAGCACCAGAAT[Fluorocein∼dT] CACGAGTACCTGCGG
y35 3′phos	CGGGATCGAGCACCAGAATTCACGAGTACCT GCGG-phosphate
z60	GCTTGCTAGGACGGATCCCTCGAGGCCGCAGGT ACTCGTGAATTCTGGTGCTCGATCCCG

Table 2. Construction of nuclease substrates

Substrate	^{32}P-labelled oligo	Other oligos
3′-Flap	b50	x50, r26–50
Splayed duplex	b50	x50
3′-Overhang	b50	x26–50
5′-Overhang	b25	x50
Biotinylated substrate	b50-bio	x26–50
100mer nicked duplex	b75	x100, r26–50
100mer 5′-overhang	b75	x100
60mer nicked duplex	y35	z60, b25
60mer fluorescein	y35 F	z60, b25
60mer CPD	y35 CPD	z60, b25
60mer 3′ phosphate	y35 3′phos	z60, b25
3′-Flaps with gaps	b50	x60 and r36, r35, r34, r33, r32 or r31

oligos forming the substrate strand were 5′-[^{32}P]end labelled and assembled into various structures by slow cooling from 85°C to room temperature overnight. Substrates were assembled by including the oligos indicated in Table 2 and were purified on a native 6% acrylamide gel followed by elution and ethanol precipitation as described previously (22). Size markers (A and G) were prepared from labelled substrates under standard protocols. To make the biotin–streptavidin conjugated substrates [biotin∼dT] was substituted at positions 3 and 45 in the b50 oligonucleotide (b50-bio), which was annealed with oligo x26–50 and incubated a 2.5-fold molar excess of streptavidin over biotin prior to use in assays.

Nuclease assays

Reactions were assembled in 30 mM HEPES, pH 7.6, 5% glycerol, 40 mM KCl, 0.1 mg/ml BSA and 0.1 mg/ml calf thymus DNA with 80 nM DNA substrate and 1 μM *Sso*XPF plus PCNA and equilibrated at 55 or 35°C as indicated. Cleavage was initiated by adding $MgCl_2$ to a final concentration of 10 mM, mixed briefly and incubated at 55 or 35°C as indicated. Aliquots (5 or 10 μl) were taken at selected time points and added to chilled stop solution (10 mM Tris–HCl, pH 8, 10 mM EDTA, 0.1 mg/ml calf thymus DNA) to terminate the reaction. DNA was ethanol precipitated and analysed by denaturing polyacrylamide/urea/TBE gel.

RESULTS

Cleavage of minimal substrates by SsoXPF

We have previously characterized the substrate specificity of SsoXPF against a range of branched DNA substrates and shown that 3′-flap structures and nicked duplexes are cleaved most quickly (16). We analysed a subset of these substrates to determine whether any support processive cleavage by SsoXPF *in vitro* (Figure 1). Each has a labelled substrate strand (white circle at 5′ end) which is 50 nt long for a–c, and 25 nt for substrate d. The first 25 nt form a DNA duplex with an unlabelled complementary strand. The 3′-flap substrate is cut most quickly by SsoXPF, due to the presence of the downstream duplex (strand with black circle at 5′ end, Figure 1a). For this substrate we observed cleavage initially at nucleotide G20 followed by the disappearance of the G20 product and the accumulation of processively smaller products such as C16 and G14 (Figure 1a). A splayed duplex substrate lacking the downstream duplex is cleaved ∼10-fold more slowly than a 3′-flap (16). For this substrate, strong initial cleavage was no longer seen predominantly at G20. Rather, a background of weaker bands were seen from position 24 to 17 with the strongest cleavage observed at nucleotide C16 (Figure 1b). This suggests that the presence of the downstream duplex influences the choice of cleavage position by SsoXPF, as has been observed for Mus81 (15). A similar situation holds for a 5′-overhang substrate, which has a 5′-flap but no 3′-flap (Figure 1d). This substrate is cleaved slowly but with a similar pattern to the splayed duplex. This reinforces the observation that a 3′-flap is not an essential feature of substrates for SsoXPF, presumably because the enzyme generates a small single strand flap on binding DNA substrates (see later). For all three of these substrates (a, b and d), there was evidence of processive cleavage, with initial cleavage products close to the branch point being processed further to produce shorter substrate strands. This observation agrees with initial observations with 3′-flap substrates (12). Finally, a substrate with a 3′-overhang represents the poorest cleavable substrate we have tested for SsoXPF (16). Cleavage was seen at a number of positions including G24, A23 and predominantly at G20 (Figure 1c). For this substrate each product increased linearly over time with no hint of conversion to smaller products. The lack of processivity correlates with the absence of a downstream duplex or single-stranded DNA (ssDNA) flap (see later).

SsoXPF can function as a processive nuclease on long DNA substrates

The data presented in Figure 1 suggest that over extended assay times *Sulfolobus* XPF cuts a 3′-flap substrate in a processive manner, with an initial cleavage site 4–5 nt 5′ of the single-stranded flap followed by further processing of the cleavage products yielding gapped duplex products. We have extended these studies by using much larger substrates to determine the ability of SsoXPF to sustain this exonuclease-type activity over longer distances. Large nicked duplex substrates with an 'upstream duplex' 75 bp in length were cleaved by SsoXPF to generate products with larger and larger gapped duplexes (Figure 2). The activity profile was biphasic, with a very rapid initial cleavage removing ∼5 nt of the substrate strand (giving a product of ∼70 nt near the top of the gel), followed by a slower activity which resulted in digestion to ∼12 nt from the 5′ end of the upstream duplex. The initial rapid cleavage reflects the stimulatory effect of the downstream duplex, which is present in all the preferred SsoXPF substrates (16). When the strand that allows formation of the downstream duplex was absent the initial rapid cleavage was abolished, but processive nuclease activity was still observed (Figure 2). The control lanes in Figure 6 show that there is no contaminating nuclease activity in our preparations of recombinant PCNA and XPF. Thus, it appears that SsoXPF has the potential to act as a processive nuclease *in vitro*. The end point at around 12 bp was seen with all substrates regardless of length and is probably dictated by the instability of DNA duplex of this length at the assay temperature of 55°C, as melting of the remaining substrate would prevent further digestion. Therefore, the enzyme might act like an exonuclease over much longer stretches *in vivo*. Our data

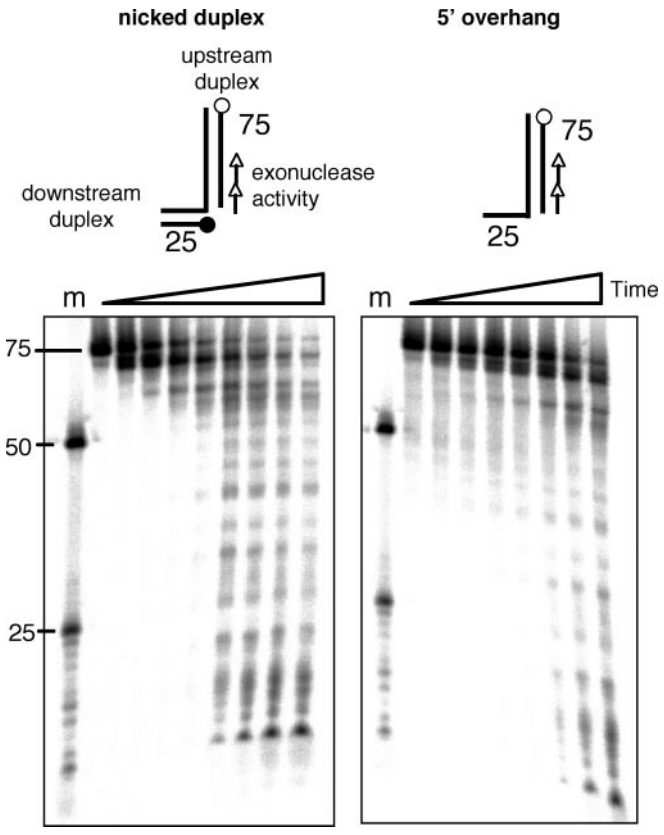

Figure 2. Processive nuclease activity of SsoXPF. White circles show the ^{32}P-labelled 5′ DNA end and a black circle indicates the 5′ end of the downstream duplex. Time points were 0, 10, 30 s, 1, 2, 5, 10, 20 and 30 min with the exception of no 10 min time point for the 5′-overhang. m = 50 and 25 nt DNA markers (some degradation of the 25 nt DNA). The cleavage pattern observed probably reflects the sequence specificity of the enzyme. Processive cleavage was not due to a contaminating exonuclease in the PCNA or *Sso*XPF preparations (see controls in Figures 5 and 6).

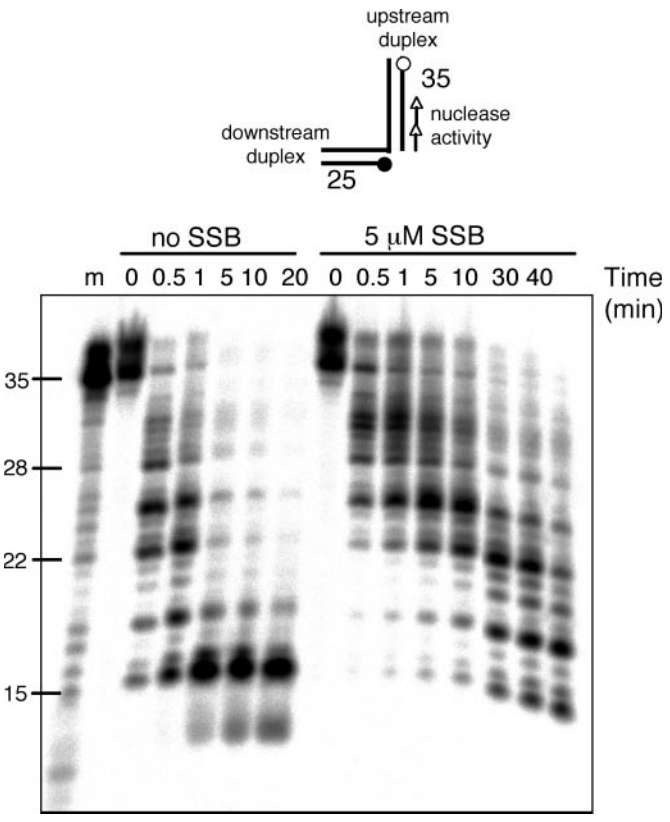

Figure 3. SSB inhibited processive cleavage by SsoXPF at 55°C. *Sulfolobus* SSB protein was added to a final concentration of 5 μM, representing a large excess over that required to saturate the DNA present in the assay. White circles show the ^{32}P-labelled 5′ DNA end and a black circle indicates the 5′ end of the downstream duplex. m = A and G size marker.

do not rule out the possibility that the XPF protein dissociates after each cleavage reaction then re-associates to catalyse another reaction, but functionally a processive action of the enzyme is observed. It is relevant in this regard to note the requirement of SsoXPF for the sliding clamp PCNA, which is a processivity factor for many enzymes. Clearly, PCNA may remain bound to the DNA even if XPF dissociates, facilitating renewed binding of the latter.

The *Sulfolobus* single-stranded DNA-binding protein inhibits the activity of SsoXPF

We tested the ability of the *Sulfolobus* single-stranded DNA-binding protein (SsoSSB) to modulate the activity of SsoXPF. For these assays and those in the Figure 4, we used a nicked duplex DNA substrate with a 35 bp upstream duplex and 25 bp downstream duplex (Figure 3). In both Figures 3 and 4, there is a variable amount of a smeared species present above the full-length 35 nt oligonucleotide substrate strand. As this was present in the marker lane as well as the reaction lanes it is clearly not due to the presence of the proteins in the assay, and as the marker lane yielded clear Maxam–Gilbert sequencing products the oligonucleotide had a defined 5′ end, suggesting that the band or smear above the uncut 35 nt substrate strand is an artefact of this particular substrate construct. For the time points investigated in these experiments, the substrates are mostly processed to smaller products that again range down to 12–13 nt from the 5′ end of the substrate strand. When the *Sulfolobus* single-stranded DNA-binding protein SsoSSB was included in the assay, it had a clear inhibitory effect on the exonuclease activity of XPF, with gapped duplex products appearing significantly more slowly when SSB was present in the assay (Figure 3). This confirms previous studies of the effect of SSB on XPF activity (12). As *Sulfolobus* SSB is an abundant protein *in vivo* (21), it is reasonable to expect that the production of gapped duplex DNA by SsoXPF (arising from processive cleavage of a nicked duplex substrate) would allow binding by SSB that in turn could inhibit the further activity of the nuclease. This could provide a means for the control or limitation of DNA degradation by XPF *in vivo*. A similar situation has been observed for the control of the human nuclease EXO1 by the ssDNA-binding protein RPA (23). *Sulfolobus* SSB has been proposed to act as a DNA damage sensor, recruiting repair proteins to sites of DNA damage via interactions with a C-terminal tail domain (24).

SsoXPF exonuclease activity can remove a variety of DNA lesions *in vitro*

Starting with the same substrate used in Figure 3 as a template, we investigated the ability of SsoXPF to degrade a substrate

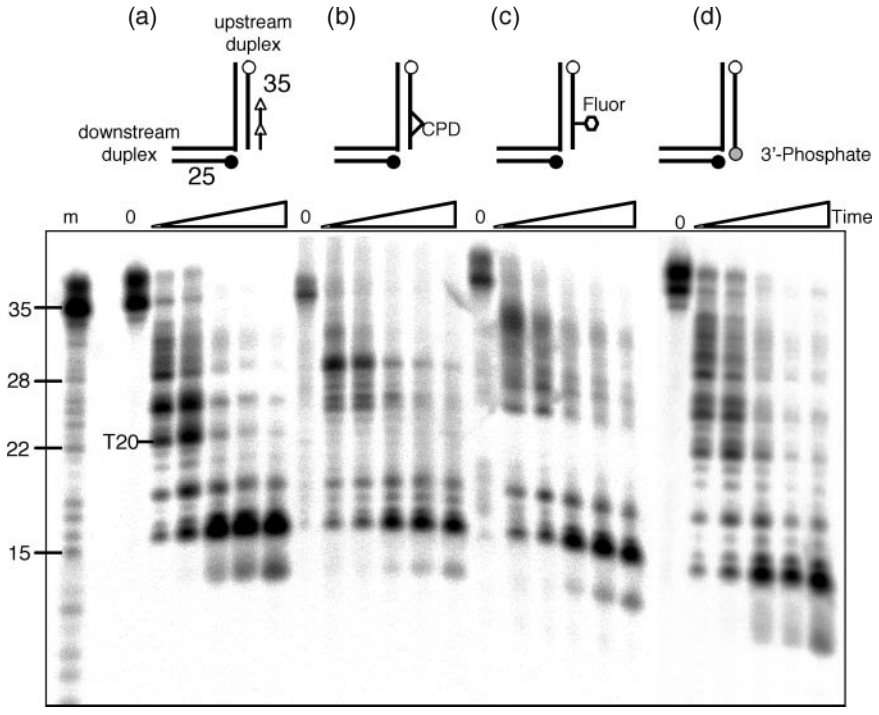

Figure 4. SsoXPF cleaved past lesions in DNA. A 35mer oligonucleotide modified with either a thymine dimer (CPD) at positions 19–20, a fluorescein at position 20, or a phosphate group at the 3′ end, was used to make a 60 bp nicked duplex structure where the nick was 3′ of the lesion. White circles show the ^{32}P-labelled 5′ DNA end, black circles the 5′ end of the downstream duplex, and a grey circle indicates the 3′ phosphate. Time points were 30 s, 1, 5, 10 and 20 min; m = A and G size marker.

strand containing various types of DNA modification. A substrate with a single CPD in place of the thymine residues at positions 20 and 21 was digested by SsoXPF with kinetics comparable with the unmodified substrate (Figure 4b). Notably, the prominent cleavage observed at position T20 in the unmodified substrate was lost in the CPD-containing substrate, suggesting that the XPF nuclease skips this position and cleaves the DNA 5′ of the CPD lesion. Similarly, DNA containing a bulky fluorescein adduct at position T20 was also cleaved efficiently, but with a gap in the cleavage pattern close to the adduct (Figure 4c). Lastly, we tested the ability of SsoXPF to process substrates with a phosphate group at the 3′ end of the oligonucleotide. This mimics the types of blocked 3′-termini that are produced by certain base excision repair (BER) enzymes that are refractory to repair synthesis and must therefore be processed further by the repair apparatus (4). Again, we saw efficient digestion of the blocked substrate oligonucleotide (Figure 4d). Thus, XPF can digest DNA substrates containing a variety of types of DNA damage or modification, suggesting a potential role in the removal of these lesions *in vivo*.

XPF-PCNA can assemble at internal DNA sites *in vitro*

We have shown previously that the *Sulfolobus* XPF enzyme has little or no activity in the absence of the sliding clamp PCNA (12). For cleavage of long undamaged DNA substrates PCNA may confer processive-type nuclease activity on XPF, as discussed earlier. We predict that PCNA is most likely to be present on the upstream duplex, as this is the only DNA duplex close to the enzyme in gapped duplex substrates. However, bulky DNA lesions such as fluorescein adducts and photoproducts present a potential problem, as they will represent a physical barrier to a sliding clamp. Nevertheless DNA containing these lesions are cleaved efficiently by XPF. To explain this phenomenon, we tested the ability of XPF-PCNA to assemble on DNA substrates where no free DNA ends were available by blocking both ends of a 3′-overhang DNA substrate with a biotin–streptavidin conjugate (Figure 5). XPF was observed to cleave this blocked substrate efficiently, suggesting that the PCNA clamp can assemble around the DNA duplex without the requirement for a free DNA end to slide on to. The control lanes using the XPF mutant with a deletion of the six C-terminal residues that interact with PCNA showed no activity, confirming that the activity detected with the wild-type enzyme was dependent on PCNA.

Sulfolobus PCNA subunits 1 and 2 are known to form a very tight heterodimer in solution (K_D 1.2×10^{-11} M) whereas PCNA3 does not bind appreciably to either of the other 2 subunits, but binds the PCNA1-2 heterodimer with a dissociation constant of 2.7×10^{-7} M – four orders of magnitude more weakly than the 1–2 interaction (25). Thus *in vivo* the *Sulfolobus* PCNA molecule can be considered to be a 1–2 heterodimer with a weakly associated subunit 3. This implies that clamp loading in the absence of replication factor C may act through addition of PCNA3 to a PCNA1-2 heterodimer on the DNA. XPF forms direct interactions with PCNA subunits 1 and 3 in solution *in vitro* (12). *In vivo*, this is likely to mean that XPF binds a PCNA1-2 heterodimer through interaction with PCNA1. As *Sulfolobus* XPF is a dimer with two PCNA interaction motifs, it is possible that the second motif binds to PCNA3 and promotes ring closure around the DNA. In

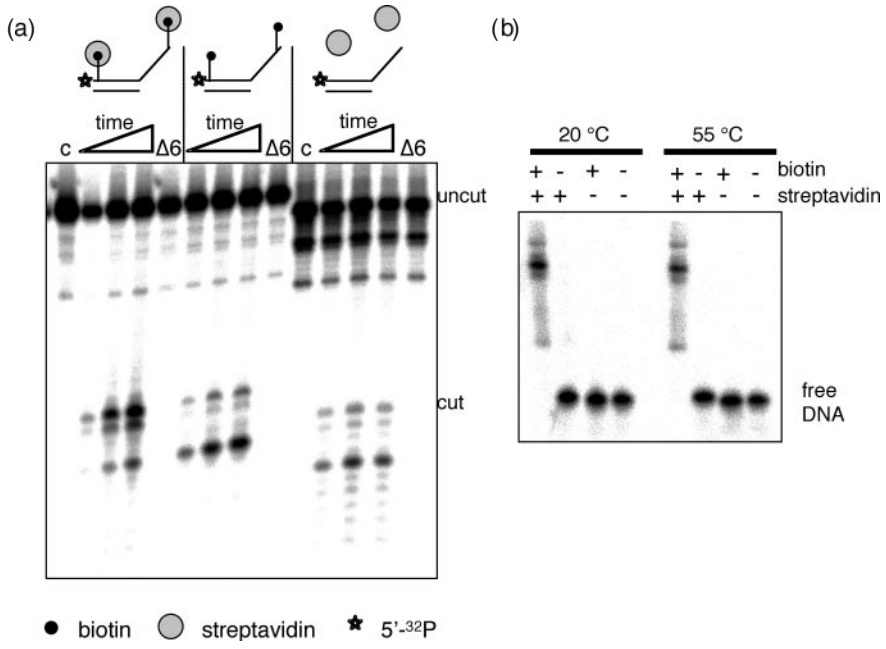

Figure 5. PCNA stimulates SsoXPF nuclease activity on biotin–streptavidin end-blocked DNA substrates. (**a**) A 50 nt oligonucleotide (b50-bio) with biotin-dT nucleotides at positions 3 and 45 was annealed to a complementary 25 nt oligonucleotide (x26–50) to make a 3′-overhang substrate. Addition of streptavidin resulted in biotin–streptavidin conjugate at both ends that would prevent PCNA sliding onto the DNA. Controls without streptavidin and without biotin modification were also assayed. DNA was incubated with 1 μM PCNA and 1 μM SsoXPF for 5, 10 and 15 min at 55°C. Control lanes are: c, DNA alone; Δ6, DNA incubated with PCNA and mutant XPF lacking the C-terminal six residues, which abolishes its interaction with PCNA. Weaker bands observed below the full-length substrate are present in all controls and do not represent products of the XPF nuclease activity. (**b**) Streptavidin specifically bound biotinylated DNA substrates. 80 nM ^{32}P-labelled biotinylated DNA substrate or unbiotinylated control was mixed with a 2.5-fold molar excess of streptavidin to biotin and incubated for 10 min at 20 or 55°C. Samples were run on a 6% native acrylamide gel and visualized by phosphoimaging. All the biotinylated DNA in the presence of streptavidin showed decreased electrophoretic mobility consistent with the formation of a streptavidin–biotin–DNA complex. Streptavidin did not bind the unbiotinylated DNA.

other words, XPF acts as a clamp loader in its own right. Alternatively, the PCNA heterotrimer may assemble spontaneously on duplex DNA. This has been observed previously for PCNA from *Pyrococcus furiosus* (26). In either case, disassembly and reassembly of PCNA on encountering bulky DNA lesions provides a plausible mechanism for the processive activity of XPF that is observed.

The downstream duplex controls XPF cleavage site selection

As discussed earlier, the preferred substrates of SsoXPF share the common property of a downstream duplex with a 5′ DNA end close to the branch point. The structure of *Aeropyrum pernix* XPF bound to a DNA duplex showed that the downstream duplex is bound specifically in the minor groove by the HhH_2 domain of one subunit of the enzyme, bringing the DNA end into close proximity with the nuclease domain (17). To test the importance of the downstream duplex for target site selection, we constructed a series of synthetic substrates, starting with a nicked duplex where there are no missing nucleotides between the downstream and upstream duplexes, and progressively increasing the size of the single-stranded DNA gap between the two duplexes by shortening the length of the r strand. In these substrates, we added 10 bp to the downstream duplex to increase the stability at elevated temperatures, so a full-length downstream duplex was 35 bp and for example substrates incorporating the r31 oligonucleotide have a gap of 4 nt. For a nicked duplex with a downstream duplex of 35 bp, cleavage of the substrate strand of the upstream duplex was centred on position G20, liberating 5 nt of DNA from the 3′ end of the substrate strand (Figure 6a). An overhang of 1 nt was also well tolerated by the enzyme (r36 substrate). As the gap between the 5′ end of the downstream duplex and the 3′ end of the upstream duplex was increased, a shift in the cleavage position in the substrate strand was observed, such that the gap between the two DNA ends was maintained at 5 nt. Thus, for example, a gap of 3 nt (r32 substrate) resulted in cleavage centred on position A23, 2 nt from the 3′ end of the substrate strand (Figure 6a). This suggests that the favoured cleavage site of SsoXPF is determined by the position of the end of the downstream duplex. This is a similar situation to that seen for Mus81 (15). Substrates with gaps from 0 to 3 nt in length were cut at very similar rates; however, a significant decrease in reaction rate was observed when the gap between the downstream and upstream duplexes was increased to 4 nt (Figure 6b). With larger gaps the stimulatory effect of the downstream duplex was largely lost, with reaction rates reverting to those observed for splayed duplex substrates (data not shown).

DISCUSSION

DNA end-directed activity of SsoXPF

Our data, coupled with the recent structural studies suggest two modes of action for *Sulfolobus* XPF: a downstream-duplex-directed endonuclease activity and a slower but processive

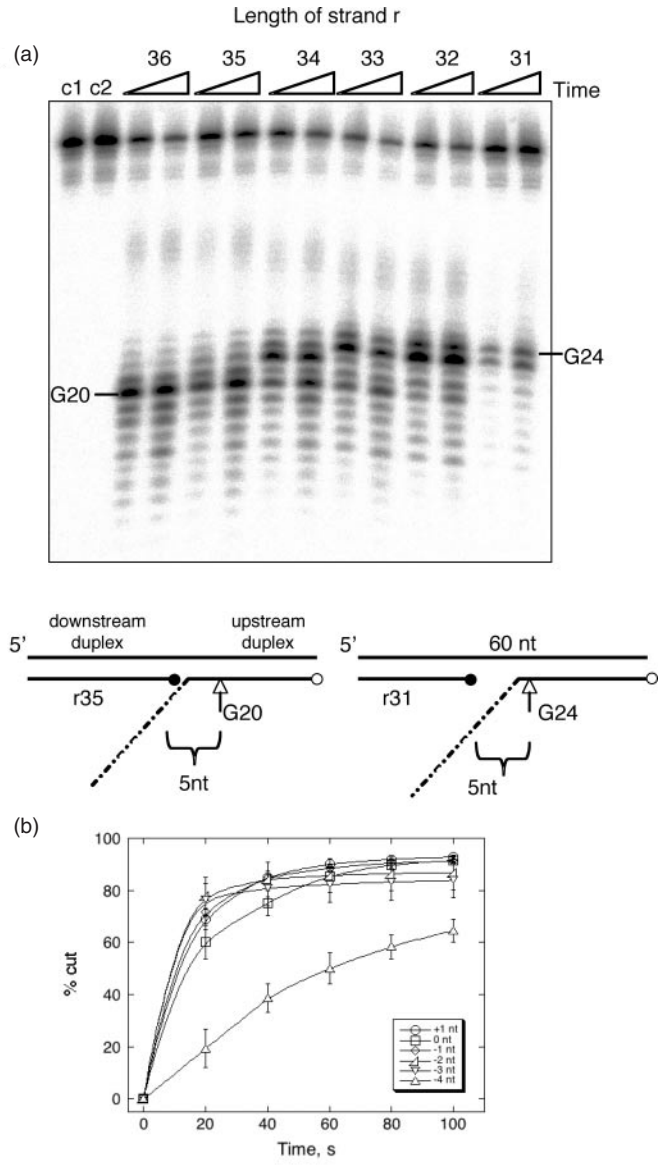

Figure 6. Effect of the downstream duplex position on SsoXPF cleavage. (**a**) The downstream duplex directs choice of cleavage site. b50 and x60 oligonucleotides were annealed with various lengths of r oligonucleotides (Table 1) to create 3′-flap structures with a 1 nt overlap with the 3′-flap (r36) or gaps of 1–4 nt between the 3′-flap and downstream duplex DNA (Table 2). DNA was incubated with 1 μM PCNA and 1 μM *Sso*XPF at 35°C for 20 and 40 s, and the uncut substrate separated from the cleavage product by denaturing polyacrylamide gel electrophoresis. Controls c1 had DNA and PCNA but XPF omitted; c2 had DNA and XPF but PCNA omitted. (**b**) Quantifying the effect of the position of the downstream duplex DNA on the rate of *Sso*XPF cleavage. The proportion of the substrate cleaved within a given time was calculated by phosphorimaging. The data points are the means of triplicate measurements, and standard errors are shown. The data were fitted with a smooth curve fit that does not represent a model for the kinetic progress of the reaction.

exonuclease-like activity. In the first, SsoXPF locates a branch point or discontinuity in a DNA duplex, such as a nick or 3′-flap, and binds the substrate with the downstream duplex engaged by one of HhH_2 domains and the 5′ end of the downstream duplex close to the nuclease domain, as observed in the crystal structure (17). This determines the position of cleavage by the enzyme on the substrate strand of the upstream duplex.

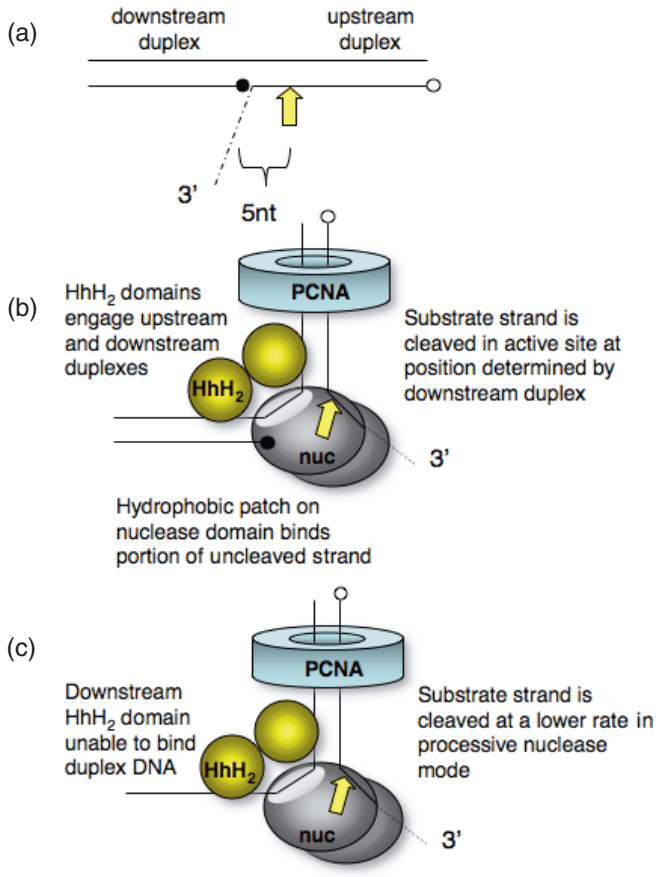

Figure 7. Model for DNA binding and cleavage by SsoXPF. The nuclease domain dimer, helix–hairpin–helix (HhH_2) domain dimer, and PCNA heterotrimer are indicated. Linkages between these species are omitted for clarity. (**a**) DNA substrates have a nick or 3′-flap that allows distortion on binding by the enzyme. The 5′ end of the downstream duplex (black circle) directs the cleavage position on the upstream duplex, with a spacing of 5 nt (yellow arrow). (**b**) When functioning as an endonuclease, the HhH_2 domains engage both the upstream and downstream duplexes in the minor groove. We predict that the uncleaved strand linking the two duplexes passes through the ssDNA-binding surface on the nuclease domain (light oval), thus generating a small 3′-flap in the substrate strand that can enter the active site, allowing cleavage at a defined position with respect to the downstream duplex. PCNA has been drawn on the upstream duplex in accordance with Newman *et al*. (17). (**c**) In processive nuclease mode, the HhH_2 interaction with the downstream duplex is lost. The enzyme processively unwinds and cleaves small flaps of ssDNA in a 3′–5′ direction, with the ability to bypass and thus remove a variety of DNA lesions and blocked 3′-termini. PCNA can disassemble and reassemble at bulky DNA lesions.

Moving the position of the downstream duplex results in correlated movement of the cleavage site. The relative importance of downstream duplex binding by the HhH_2 domain versus engagement of the 5′ end of the downstream duplex by the nuclease domain is not yet clear, but the observation that phosphorylation of the 5′ DNA end does not affect the cleavage rate significantly (16) suggests that the interaction of the HhH_2 domain with the duplex may be more important. The structure of *Aeropyrum* XPF highlighted a prominent hydrophobic strip on the surface of the nuclease domain, which was predicted to function as a binding site for the ssDNA linking the downstream and upstream duplexes (17). The model for DNA binding and cleavage by the XPF family of proteins suggested by structural studies (17) is shown in Figure 7.

In this model, the downstream duplex is engaged by one HhH_2 domain and the upstream duplex by the other. The strand linking the two duplexes engages the hydrophobic strip, resulting in generation of a small stretch of unpaired ssDNA in the substrate strand, which is cleaved in the nuclease active site.

This model is consistent with footprinting data that demonstrate opening of a DNA duplex near the point of cleavage in the Hef:DNA complex (14) and with other published observations regarding the structure and activity of the enzyme. In particular, the spacing of 5 nt between the downstream duplex and the cleavage site is explained by the length of ssDNA in the uncleaved strand that is required to link the upstream and downstream duplexes. This must be at least 5 nt in length, but may be slightly longer if the substrate strand is not cleaved immediately next to the ds–ssDNA junction. The model also explains why nicked duplex DNA is cut almost as quickly as 3′-flap substrates, as a 3′-flap is generated upon substrate binding. The same 5′ end-directed endonuclease activity has previously been observed for the Mus81–Mms4 enzyme (15), where the spacing between the 5′ end and the cleavage site is also 5 nt. This may reflect a fundamental conservation of DNA substrate recognition between the two enzymes, which is interesting given the differences in domain organization of the proteins. The nuclease domain in Mus81 is flanked by two predicted HhH motifs, rather than an arrangement with a double motif at the C-terminus as seen in XPF (17). It is conceivable that these two motifs associate in the tertiary structure of the enzyme, but there is no data to support this presently. It is also unclear whether the binding partner of Mus81 (Mms4 in *Saccharomyces cerevisiae*) contributes towards DNA substrate recognition.

Processive activity of SsoXPF

Once the gap between the downstream and upstream duplexes is greater than 3 nt, the enzyme cannot simultaneously engage the downstream and upstream duplexes and the 3′-flap substrate strand. This is reflected in the lower rates of cleavage of gapped duplex and splayed duplex substrates. Unlike Mus81, however, SsoXPF does have an appreciable activity against substrates where the downstream DNA is single stranded (16). Mus81 shows no processive cleavage of DNA substrates, whereas SsoXPF clearly has the ability to cut gapped duplex substrates repeatedly. We therefore propose a second mode of action for SsoXPF, where the upstream duplex is bound by one HhH_2 domain and a stretch of ssDNA is engaged by the hydrophobic strip as before, but the interaction with the downstream duplex is absent (Figure 7). Eukaryal XPF-ERCC1 has apparently evolved to favour this type of substrate, and Newman *et al*. (17) have suggested this relates to changes in the 'downstream' HhH_2 domain that have altered its binding preference from ds to ssDNA. Recent studies suggest the N-terminal domain of ERCC1 also binds ssDNA (27), which may also account for the altered substrate specificity of this enzyme. Although the activity of the archaeal enzyme on gapped duplexes is significantly slower than the end-directed, Mus81-type mode, because it has no requirement for a downstream duplex the enzyme can function in a processive manner, repeatedly engaging an ssDNA stretch to unwind the upstream duplex and cleave short ssDNA flaps from the substrate strand with a 3′–5′ polarity. This is facilitated by an interaction with PCNA, which may anchor the XPF nuclease on the upstream DNA duplex and which can itself assemble at internal sites on DNA.

A role for SsoXPF in the digestion of damaged DNA?

This mode of action has two other potentially important consequences: the ability to bypass blocked 3′-termini and DNA lesions. The presence of a phosphate at the 3′ end of the substrate strand did not affect the exonuclease activity of SsoXPF. This is easy to understand, as the enzyme apparently does not interact with the 3′ end of the substrate strand, cleaving it a few nt upstream of the branch point. This property is also found for the yeast Rad1–Rad10 enzyme, which can remove 3′ blocking groups including a 3′-phosphate and 3-phosphoglycolate by 3′–5′exonucleolytic digestion, generating 3–5 nt products (4). In yeast, this activity appears important for the repair of 3′ blocked termini arising from damage by reactive oxygen species and as a product of the action of DNA glycosylases, such as OGG1, Ntg1 and Ntg2 (4). The Rad1–Rad10 enzyme has also been implicated in the removal of topoisomerase 1 covalently linked to DNA 3′ ends (28,29), and it is likely that the enzyme simply removes a small piece of DNA that includes the protein from the 3′ DNA end, producing a ligatable 3′ hydroxyl end. If that is the case, then even large protein adducts at the 3′ ends of DNA may not present an obstacle to the XPF family of enzymes.

Because SsoXPF successively unwinds and cleaves DNA in steps of a few nucleotides, it is not inhibited by bulky and helix distorting lesions such as photoproducts. These are presumably bypassed by extruding them past the enzyme active site. There is no obvious reason why substrates containing DNA mismatches would not also be processed by the enzyme. As we have shown, this processive activity can persist over at least 60 nt of DNA, and probably much further. Therefore, any DNA molecule with a nick could be processed, with the DNA strand on the 5′ side of the nick degraded to produce a gapped duplex molecule. SsoXPF could therefore act as flexible nuclease for the removal of a variety of types of DNA damage. The requirement for a nick in the damaged strand 3′ of a lesion places obvious constraints on this activity. For example, SsoXPF could function along with DNA glycosylases and AP endonucleases in BER as has been shown for eukaryal Rad1–Rad10. Alternatively, SsoXPF could function with a UV damage endonuclease to remove photoproducts. Presently, no such enzyme is annotated in *S.solfataricus*, and although a UVSE family endonuclease has been annotated in the related organism *Sulfolobus acidocaldarius* (30) this enzyme cuts on the 5′ side of photoproducts (31), which would not allow removal of the lesion by SsoXPF. This would represent a long-patch BER-type pathway that could potentially provide an alternative to NER in archaea. Lastly, XPF could play a role in an archaeal-specific mismatch detection system to degrade mismatched DNA. The latter possibility is reminiscent of the mismatch repair activity of human exonuclease I, which has a 3′–5′ exonuclease activity that is also dependent on PCNA (32). Similar to SsoXPF, exonuclease I activity is modulated by the single-stranded DNA-binding protein RPA, which may control the processivity of the enzyme (23). Given the absence of any clear mismatch

repair system in archaea this possibility warrants further consideration.

ACKNOWLEDGEMENTS

Thanks to the BBSRC and AICR for financial support. M.F.W. is a Royal Society URF. Thanks to Jana Rudolf and Neil McDonald for valuable discussions. Funding to pay the Open Access publication charges for this article was provided by the BBSRC.

Conflict of interest statement. None declared.

REFERENCES

1. Evans,E., Moggs,J.G., Hwang,J.R., Egly,J.M. and Wood,R.D. (1997) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J.*, **16**, 6559–6573.
2. Motycka,T.A., Bessho,T., Post,S.M., Sung,P. and Tomkinson,A.E. (2004) Physical and functional interaction between the XPF/ERCC1 endonuclease and hRad52. *J. Biol. Chem.*, **279**, 13634–13639.
3. Zhu,X.D., Niedernhofer,L., Kuster,B., Mann,M., Hoeijmakers,J.H. and de Lange,T. (2003) ERCC1/XPF removes the 3′ overhang from uncapped telomeres and represses formation of telomeric DNA-containing double minute chromosomes. *Mol. Cell*, **12**, 1489–1498.
4. Guzder,S.N., Torres-Ramos,C., Johnson,R.E., Haracska,L., Prakash,L. and Prakash,S. (2004) Requirement of yeast Rad1-Rad10 nuclease for the removal of 3′-blocked termini from DNA strand breaks induced by reactive oxygen species. *Genes Dev.*, **18**, 2283–2291.
5. de Laat,W.L., Appeldoorn,E., Jaspers,N.G. and Hoeijmakers,J.H. (1998) DNA structural elements required for ERCC1-XPF endonuclease activity. *J. Biol. Chem.*, **273**, 7835–7842.
6. Osman,F., Dixon,J., Doe,C.L. and Whitby,M.C. (2003) Generating crossovers by resolution of nicked Holliday junctions: a role for Mus81-Eme1 in meiosis. *Mol. Cell*, **12**, 761–774.
7. Hollingsworth,N.M. and Brill,S.J. (2004) The Mus81 solution to resolution: generating meiotic crossovers without Holliday junctions. *Genes Dev.*, **18**, 117–125.
8. Nishino,T., Komori,K., Ishino,Y. and Morikawa,K. (2003) X-ray and biochemical anatomy of an archaeal XPF/Rad1/Mus81 family nuclease: similarity between its endonuclease domain and restriction enzymes. *Structure (Camb.)*, **11**, 445–457.
9. Komori,K., Fujikane,R., Shinagawa,H. and Ishino,Y. (2002) Novel endonuclease in Archaea cleaving DNA with various branched structure. *Genes Genet. Syst.*, **77**, 227–241.
10. Nishino,T., Komori,K., Tsuchiya,D., Ishino,Y. and Morikawa,K. (2005) Crystal structure and functional implications of *Pyrococcus furiosus* hef helicase domain involved in branched DNA processing. *Structure (Camb.)*, **13**, 143–153.
11. 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.
12. 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.
13. Gaillard,P.H. and Wood,R.D. (2001) Activity of individual ERCC1 and XPF subunits in DNA nucleotide excision repair. *Nucleic Acids Res.*, **29**, 872–879.
14. Nishino,T., Komori,K., Ishino,Y. and Morikawa,K. (2005) Structural and functional analyses of an archaeal XPF/Rad1/Mus81 nuclease: asymmetric DNA binding and cleavage mechanisms. *Structure (Camb.)*, **13**, 1183–1192.
15. Bastin-Shanower,S.A., Fricke,W.M., Mullen,J.R. and Brill,S.J. (2003) The mechanism of Mus81-Mms4 cleavage site selection distinguishes it from the homologous endonuclease Rad1-Rad10. *Mol Cell. Biol.*, **23**, 3487–3496.
16. Roberts,J. and White,M.F. (2004) An archaeal endonuclease displays key properties of both eukaryal XPF-ERCC1 and Mus81. *J. Biol. Chem.*, **280**, 5924–5928.
17. 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.
18. Grogan,D.W. (2000) The question of DNA repair in hyperthermophilic archaea. *Trends Microbiol.*, **8**, 180–185.
19. White,M.F. (2003) Archaeal DNA repair: paradigms and puzzles. *Biochem. Soc. Trans.*, **31**, 690–693.
20. 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.
21. Wadsworth,R.I. and White,M.F. (2001) Identification and properties of the crenarchaeal single-stranded DNA binding protein from *Sulfolobus solfataricus*. *Nucleic Acids Res.*, **29**, 914–920.
22. Kvaratskhelia,M. and White,M.F. (2000) An archaeal Holliday junction resolving enzyme with unique properties. *J. Mol. Biol.*, **295**, 193–202.
23. Genschel,J. and Modrich,P. (2003) Mechanism of 5′-directed excision in human mismatch repair. *Mol. Cell*, **12**, 1077–1086.
24. Cubbedu,L. and White,M.F. (2005) DNA damage detection by an archaeal single stranded DNA binding protein. *J. Mol. Biol.*, **353**, 507–516.
25. Dionne,I., Nookala,R.K., Jackson,S.P., Doherty,A.J. and Bell,S.D. (2003) A Heterotrimeric PCNA in the Hyperthermophilic Archaeon *Sulfolobus solfataricus*. *Mol. Cell*, **11**, 275–282.
26. Cann,I.K., Ishino,S., Hayashi,I., Komori,K., Toh,H., Morikawa,K. and Ishino,Y. (1999) Functional interactions of a homolog of proliferating cell nuclear antigen with DNA polymerases in Archaea. *J. Bacteriol.*, **181**, 6591–6599.
27. 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. USA*, **102**, 11236–11241.
28. Vance,J.R. and Wilson,T.E. (2002) Yeast Tdp1 and Rad1-Rad10 function as redundant pathways for repairing Top1 replicative damage. *Proc. Natl Acad. Sci. USA*, **99**, 13669–13674.
29. Pouliot,J.J., Robertson,C.A. and Nash,H.A. (2001) Pathways for repair of topoisomerase I covalent complexes in *Saccharomyces cerevisiae*. *Genes Cells*, **6**, 677–687.
30. Chen,L., Brugger,K., Skovgaard,M., Redder,P., She,Q., Torarinsson,E., Greve,B., Awayez,M., Zibat,A., Klenk,H.P. *et al.* (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. *J. Bacteriol.*, **187**, 4992–4999.
31. Takao,M., Yonemasu,R., Yamamoto,K. and Yasui,A. (1996) Characterization of a UV endonuclease gene from the fission yeast *Schizosaccharomyces pombe* and its bacterial homolog. *Nucleic Acids Res.*, **24**, 1267–1271.
32. Dzantiev,L., Constantin,N., Genschel,J., Iyer,R.R., Burgers,P.M. and Modrich,P. (2004) A defined human system that supports bidirectional mismatch-provoked excision. *Mol. Cell*, **15**, 31–41.