Published in final edited form as:

J Mol Biol. 2009 June 19; 389(4): 661–673. doi:10.1016/j.jmb.2009.04.060.

Structural and functional characterisation of a conserved archaeal RadA paralog with antirecombinase activity

Anne-Marie McRobbie[†], Lester G. Carter^{†,§}, Melina Kerou, Huanting Liu, Stephen A. McMahon, Kenneth A. Johnson, Muse Oke, James H. Naismith, and Malcolm F. White^{*}
Centre for Biomolecular Sciences, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9ST, UK

Summary

DNA recombinases (RecA in bacteria, Rad51 in eukarya and RadA in archaea) catalyse strand-exchange between homologous DNA molecules, the central reaction of homologous recombination, and are among the most conserved DNA repair proteins known. In bacteria, RecA is the sole protein responsible for this reaction, whereas, in eukaryotes, there are several RAD51 paralogs that cooperate to catalyse strand exchange. All archaea have at least one (and as many as four) RadA paralogs, but their function remains unclear. Here we show the three RadA paralogs encoded by the *Sulfolobus solfataricus* genome are expressed under normal growth conditions, and are not UV-inducible. We demonstrate that one of these proteins, Sso2452, which is representative of the large aRadC sub-family of archaeal RadA paralogs, functions as an ATPase that binds tightly to ssDNA. However, Sso2452 is not an active recombinase *in vitro*, and inhibits D-loop formation by RadA. We present the high-resolution crystal structure of Sso2452, which reveals key structural differences from the canonical RecA family recombinases that may explain its functional properties. The possible roles of the archaeal RadA paralogs *in vivo* are discussed.

Keywords

Archaea; Recombinase; RadA; Homologous Recombination; Strand Exchange

Introduction

The RecA protein family, comprising Rad51 and its paralogs in eukarya, RadA in archaea and RecA in bacteria, is one of the few universally conserved DNA repair proteins. RecA family members are DNA recombinases catalysing strand-exchange reactions that are central to homologous recombination (HR) and double-strand break repair (DSBR)¹. They bind single-stranded DNA (ssDNA), forming a nucleoprotein filament that can invade duplex DNA with a cognate sequence, leading to the formation of recombination intermediates such as heteroduplexes, D-loops and Holliday junctions². Disruption of RecA function in bacteria, or Rad51 in yeast, is highly deleterious but not fatal to the cell, whereas in metazoa Rad51 is an essential protein³. This may reflect the fact that HR/DSBR is the primary pathway for the rescue of stalled or collapsed replication forks⁴, a phenomenon known as Recombination Dependent Replication (RDR).

In contrast to bacteria where only a single RecA protein suffices, eukarya tend to encode a number of Rad51 paralogs in addition to Rad51 itself⁵. There are seven RAD51-like genes

^{*}for correspondence: tel: +44-1334-463432; mfw2@st-and.ac.uk.

[§]current address: Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[†]These authors contributed equally to the work.

in humans, comprising *RAD51A*, *RAD51B*, *RAD51C*, *RAD51D*, *XRCC2*, *XRCC3* and the meiosis-specific *DMC1*⁵. Five of these paralogs exist in two complexes *in vivo*: the BCDX2 complex (*RAD51B*, *RAD51C*, *RAD51D*, and *XRCC2*) and the *RAD51C*-*XRCC3* complex^{6,7}. These genes are all essential in mice, the proteins cooperate with Rad51 in strand-exchange reactions *in vitro*, and are required for damage-specific Rad51 repair foci *in vivo*^{8,9}. However, there is no precise understanding of the molecular roles of the Rad51 paralogs, and the *RAD51C*-*XRCC3* complex has also been implicated in the latter stages of the HR pathway⁹. *Saccharomyces cerevisiae* has only two Rad51 paralogs, Rad55/Rad57, which form a heterodimer and stimulate Rad51-mediated strand exchange *in vitro*¹⁰.

The archaea, although lacking a nucleus and bearing a superficial resemblance to bacteria, are more closely related to eukaryotes with respect to their informational processes, including DNA replication, recombination and repair¹¹, transcription¹², and translation¹³. Structural studies of HR proteins from archaea, including Rad50, Mre11 and RadA, have supplied a great deal of very useful information relevant to their eukaryal equivalents^{14,15}. Archaeal RadA is much more similar to Rad51 than to RecA, sharing the dsDNA binding N-terminal domain (NTD), and lacking the RecA-specific C-terminal domain (CTD)¹⁶. A second Rad51 paralog, RadB, has been described in some euryarchaea. RadB lacks the NTD present in RadA but has the core ATPase domain of the RecA family¹⁷. RadB appears to lack the strand-exchange activity of RadA, and turns over ATP very slowly^{18,19}.

RadA paralogs (with Blast E-values less than 1×10^{-25}) are found in all crenarchaeal genomes, many euryarchaeal genomes and even in the genome of *Nanoarchaeum equitans*, which is one of the most streamlined genomes known²⁰, suggesting an important role for these proteins in the archaeal life cycle. Several crenarchaeal genomes encode multiple RadA paralogs, including three in *S. solfataricus* (Sso0777, Sso1861, Sso2452) and four in *Pyrobaculum aerophilum*. A phylogenetic analysis reveals that these are monophyletic, with robust bootstrap values supporting the differentiation of this family of RadA paralogs from both euryarchaeal RadB and the archaeal RadA proteins (Figure 1, reviewed in²¹). The collective name “aRadC” (archaeal RadC) has recently been suggested for this group of RadA paralogs²¹. The aRadC family are differentiated from the canonical archaeal recombinase RadA as, like RadB, they lack the NTD and are restricted to the core ATP binding domain. There is also some similarity to the cyanobacterial circadian clock protein KaiC, the N-terminal domain of which is known to bind ATP and adopt a hexameric ring structure²². However, KaiC has a duplicated RecA domain and the partner protein KaiA is not present in any archaea, ruling out a role in circadian rhythm in the archaea.

These bioinformatics analyses reinforced the expectation aRadC proteins play a role in some aspect of DNA recombination in conjunction with RadA. Studies of *S. tokodaii* Sto0579, which has 84% sequence identity to Sso2452, suggested a role in regulating RadA strand exchange activity by catalysing SSB displacement from DNA²³. In this study we investigated the expression levels and transcriptional response to UV damage for RadA and all three aRadC proteins from *S. solfataricus*. We cloned and expressed *sso2452* from *S. solfataricus*, and characterised its activity *in vitro*. Sso2452 displays ssDNA-stimulated ATPase activity and binds ssDNA with nanomolar affinity. Although the protein can support the formation of small DNA heteroduplexes in an ATP-independent manner, it fails to catalyse the formation of D-loops, a key intermediate of homologous recombination, and inhibits RadA in these assays. The structure of Sso2452 reveals a canonical RecA-type core fold, but there are key differences in the DNA binding regions compared to RadA. The potential roles of RadA paralogs in archaea are discussed.

Results

Expression and UV inducibility of RadA and paralogs in *S. solfataricus*

Whereas the *S. solfataricus* genome encodes three aRadC proteins, there is little information available on their function or even whether they are all expressed *in vivo*. Using quantitative one-step reverse transcriptase-PCR, we quantified the levels of mRNA transcripts encoding RadA and the three aRadC proteins in exponentially-growing *S. solfataricus* cells. The data, summarised in Table 1, revealed that mRNA for *radA* was present at the highest levels under these conditions. The most highly transcribed paralog was *sso1861*, with transcript levels only 5-fold lower than those for *radA*, suggesting that it may be present in appreciable quantities in the cell. *Sso2452* transcript levels were also significant, at 11% of *radA* transcript levels. This is consistent with previous observations that Sso2452 could be affinity purified from *S. solfataricus* cell extracts using a biotinylated oligonucleotide²⁴. Transcripts for the third paralog *sso0777* were present at the lowest levels (7% of *radA* levels). Although *S. solfataricus* transcript abundance does not necessarily correspond directly to protein concentration in the cell, there is often a correlation²⁵. These data suggest that all three *aradC* genes are transcribed under normal growth conditions.

We and others have shown previously by global microarray analysis and Western blotting that *radA* expression is not induced appreciably by UV irradiation^{25,26}. However, Sheng and coworkers²³ reported a two-fold increase in *S. tokodaii* RadA protein levels in response to 200 J/m² UV. Both RadA and Sso0777 have been reported as up-regulated in response to DNA damage by actinomycin D in *S. solfataricus*²⁷. Transcript abundance in response to sub-lethal doses (200 J/m²) of UV radiation was measured for *radA* and all three paralogs by quantitative RT-PCR (Table 1). The expression of all four genes remained relatively stable up to 120 min after UV treatment, in agreement with microarray data for *S. solfataricus*²⁵. It is possible that protein levels are influenced at a post-translational stage, explaining the modest induction in RadA and Sto0579 observed by Sheng and co-workers²³. Our findings suggest that UV radiation and actinomycin D cause distinct transcriptional responses in *S. solfataricus*.

Gene Cloning and Protein Expression

The *sso2452* gene from *S. solfataricus* was amplified by PCR and cloned into the expression vector pDEST14 for expression in *E. coli*. The recombinant protein was purified to homogeneity by immobilized metal affinity and gel filtration chromatography as described in Materials & Methods and analysed by SDS- PAGE. The N-terminal poly-histidine tag was cleaved by incubation with the Tobacco Etch Virus protease during protein purification. Mass spectrometry determined an intact mass of 30413.7 daltons for Sso2452 after removal of the polyhistidine tag, which was in close agreement with the calculated mass of 30,416.3 daltons.

Sso2452 is a DNA-dependent ATPase

A characteristic feature of the bacterial, archaeal and eukaryotic recombinases RecA, RadA and Rad51, respectively, is their ability to catalyse DNA-dependent ATP hydrolysis. Therefore, the Sso2452 protein was tested for ATP hydrolysis activity in the presence and absence of single- and double-stranded DNA. The rate of Sso2452-mediated ATP hydrolysis was calculated to be approximately 0.17 ATP/min in the presence of ssDNA at 60 °C. This was comparable to the rate observed for *S. solfataricus* RadA (0.1-0.2 ATP/min) in the presence of ssDNA monitored at 65 °C²⁸. ATP hydrolysis by Sso2452 was less efficient in the presence of dsDNA and in the absence of DNA. Sso2452 ATPase activity was optimal at temperatures between 70 °C and 80 °C (data not shown), consistent with the optimal growth temperature of the organism. ATP hydrolysis is not an inherent mechanistic requirement for

the strand exchange process catalysed by RecA family members, but is thought to provide directionality and stability and allow more extensive strand exchange reactions to proceed²⁹. In contrast, RadB from *P. furiosus* does not catalyse multiple turnover of ATP and is not DNA stimulated¹⁸.

Sso2452 binds tightly to ssDNA

The affinities of Sso2452 for ssDNA and dsDNA were determined by fluorescence anisotropy experiments in which protein was titrated into a solution containing fluorescein-labelled single- or double-stranded DNA (45 nucleotides or base pairs, respectively). Apparent equilibrium dissociation constants (K_{DS}) were calculated from binding curves plotted using fluorescence anisotropy against protein concentration. Sso2452 bound ssDNA and dsDNA with apparent K_{DS} of 120 nM and 2.4 μ M, respectively (Figure 2B). By contrast, RadA bound ssDNA and dsDNA with apparent K_{DS} of 2.6 μ M and 5.3 μ M, respectively. These data suggest that Sso2452 binds much more tightly to ssDNA than does RadA. The identification of Sso2452 as one of a handful of proteins that can be affinity purified from cell extracts of *S. solfataricus* using a biotinylated ssDNA oligonucleotide emphasises the potential for these RadA paralogs to be instrumental in coating ssDNA *in vivo*, perhaps as a prelude to HR/DSBR²⁴. Size exclusion chromatography using an analytical Superdex 200 10-300 column (GE Healthcare) revealed that Sso2452 was monomeric in solution in the absence of DNA, but formed large molecular weight complexes with predominant peaks at approximately 250 and 600 kDa when incubated with a 34mer oligonucleotide (Figure 2C). These results were consistent with DNA-mediated assembly of the protein into a nucleoprotein complex as observed for other RecA family members including the ortholog from *S. tokodaii*²³.

Strand exchange and D-loop formation by Sso2452

To assess the ability of Sso2452 to catalyse strand exchange, two assays were performed. Firstly, Sso2452 was assayed for heteroduplex formation using a [³²P]-radiolabelled 50-nt DNA and an unlabelled duplex of 25 base pairs (modified from reference³⁰) (Figure 3A). Assays were carried out at 60 °C over a 10 min time course and analysed by native acrylamide gel electrophoresis. Both RadA and Sso2452 catalysed strand exchange at comparable rates (Figure 3B), and this activity was also observed when the two proteins were added consecutively, in either order, without any significant change in rate. Neither reaction was completely ATP-dependent, an observation that is consistent with the behaviour of RecA family proteins when catalysing strand exchange reactions between relatively short DNA sequences²⁹. As a further control, we tested the ability of the *S. solfataricus* Alba1 protein, which binds both ss- and ds-DNA³¹, to stimulate strand exchange, and found that efficient ATP-independent strand exchange could be observed for Alba1 (Figure 3D). In contrast, the single-stranded DNA binding protein SSB from *S. solfataricus* did not support strand exchange (Figure 3D). These data suggest that proteins such as Sso2452, RadA and Alba1 that can bind to both ssDNA and dsDNA can promote limited strand exchange, potentially by passive equilibrium binding of all the species present. In contrast, SSB proteins bind tightly to ssDNA and are inhibitory to strand exchange reactions.

A more stringent test of recombinase activity is an assay for D-loop formation, (Figure 4A), which monitors the ability of a ssDNA molecule to invade a duplex DNA plasmid. Briefly, the protein under test was pre-incubated at 60 °C with a [³²P]-labeled ssDNA 80mer oligonucleotide and Mg²⁺-ATP before reactions were initiated with double-stranded supercoiled plasmid pUC19. Samples were taken at regular time intervals and deproteinised followed by agarose gel electrophoresis and phosphorimaging. Under these conditions, both *E. coli* RecA and *S. solfataricus* RadA displayed robust ATP-dependent D-loop formation

activity (Figure 4B). However, no activity was observed for Sso2452 alone. Furthermore, Sso2452 inhibited the recombination activity of RadA (Figure 4C), regardless of the order of addition. This suggests that Sso2452 may displace RadA from RadA:ssDNA nucleoprotein filaments, either due simply to the higher ssDNA binding affinity of Sso2452 or due to an active nucleofilament disassembly process. To mimic the *in vivo* environment, saturating concentrations of *S. solfataricus* SSB were incubated with ssDNA prior to the addition of RadA. SSB inhibited D-loop formation by RadA (Figure 4D), consistent with studies in bacteria and eukarya^{10; 32; 33}. To test whether Sso2452 could play a role in displacing SSB to facilitate RadA access to DNA, we introduced pre-incubated DNA with SSB, followed by RadA, and finally Sso2452 (molar ratio of 1 SSB:5 RadA:5 Sso2452) (Figure 4D). However, no D-loops were formed, suggesting that under these conditions Sso2452 does not function as a mediator by displacing SSB to allow RadA-catalysed strand exchange. These data are in contrast to those of Sheng and co-workers, who suggested such a mediator function for the equivalent protein from *Sulfolobus tokodaii*²³.

Structure of Sso2452

The crystal structure of recombinant Sso2452 was solved by molecular replacement using the structure deposited for protein Pho0284 (PDB code 2dr3, unpublished) (Figure 5A). The final model contains 220 protein residues, 98 water atoms, a pyrophosphate moiety, and four zinc atoms. Sso2452 contains the typical helicase domain of a large twisted central β -sheet, sandwiched by α -helices on both sides. The central β -sheet consists of nine strands – six parallel strands ($\beta 3$ to $\beta 8$), followed by three anti-parallel strands ($\beta 9$ to $\beta 11$). There are small regions of disorder between residues 167-177 (ssDNA binding loop 2) and residues 93 to 96. The last 27 residues at the C-terminus were not visible in the electron density. There is a single non-prolyl cis peptide, serine 131, which has been observed in other RecA superfamily members and appears to be a hallmark of the protein class (reviewed in³⁴). The Sso2452 structure (which contains a monomer in the asymmetric unit) does not form any obviously biologically relevant quaternary structure by crystal packing.

Structural comparison of Sso2452 to other RecA superfamily proteins

A structural similarity search conducted using SSM³⁵ confirmed the close homology of Sso2452 to the other members of the RadA/Rad51/RecA family. Sso2452 shares the highest structural similarity with the structure used for molecular replacement, the aRadC family member Pho0284 from *P. horikoshii* (*pdb:2dr3*) with a RMSD of 1.39 Å over 213 aligned C α atoms. The hexameric quaternary structure of Pho0284 is shown in Figure 5B, in which one subunit is overlaid with the Sso2452 structure. There are some minor structural differences in the interfaces between the Pho0284 and Sso2452 structures, with a 5 Å change in the location of the loop between $\beta 8$ and $\beta 9$. Contained on this loop are several residues involved in stabilizing the Pho0284 interface, including Glu198, Glu201 and Leu203, and Asp200, which forms a salt bridge with Arg241. Pho0284 contains ADP in the active site, in a position equivalent to pyrophosphate in the Sso2452 structure. This ADP binding site is positioned in the interface of the Pho0284 monomers, and ATP/ADP binding has been shown to play an important part in the RecA-family protein assembly³⁶, so it is possible the presence of ADP in the Pho0284 protein has stabilised the hexameric structure.

Sso2452 is a good structural match to the core domain of RadA from *S. solfataricus*, (*pdb:2bke*) with 184 C α atoms aligning with a RMSD of 1.84 Å (Figure 5C). Sso2452 and Pho0284, like RadB, lack the N-terminal helix-hairpin-helix domain that is implicated in dsDNA binding^{37; 38}, and the short beta-strand polymerisation motif including a conserved phenylalanine that forms a “ball and socket” joint with adjacent RadA monomers³⁹. The lack of these features probably corresponds to a fundamentally different role for the RadB and aRadC proteins, and an inability to function as recombinases on their own. Two loops,

named L1 and L2 are implicated in ssDNA binding in the RecA/Rad51 family. L2 is disordered in the absence of DNA, but the recent co-crystal structure of RecA with ssDNA shows that the L1 loop forms a short helix and a turn on DNA binding, whilst the L2 loop forms a β -hairpin structure⁴⁰. The L2 motif is not strongly conserved between RecA and Rad51/RadA, although two consecutive glycine residues at the C-terminal end of the L2 motif that are known to bind a phosphate group of ssDNA are conserved in RadA and Rad51 orthologs. However this motif is not present in the aRadC family (Figure 5C). The L1 motif contains three arginine residues conserved across the RadA/Rad51 family. Each of these residues has been shown to be important for ssDNA binding in *S. solfataricus* RadA⁴¹, however only one of the three is present in L1 from the aRadC proteins. In summary, the structures of Sso2452 and Pho0284 highlight the key differences between the aRadC family and canonical RadA that can be related to their differing functions. The ability of the aRadC proteins to bind tightly to ssDNA deserves further investigation. Since the L1 and L2 loops of the aRadC family seem to have a less basic character than those of RadA, the interaction may involve the intercalation of hydrophobic residues between DNA bases as is seen in SSB proteins as well as ionic interactions with the phosphodiester backbone.

Discussion

Unlike bacteria, which possess a sole recombinase, RecA, almost all archaeal genomes encode at least one RadA paralog, suggesting a fundamental role for these proteins alongside RadA in HR/DSBR. In *S. solfataricus*, mRNA for the three *aRadC* genes are present at levels roughly 10-20% of RadA, which is itself a highly abundant protein. Like RadA, none were induced by UV radiation. This seems to be a characteristic of DNA repair proteins in hyperthermophiles, where harsh environments require constitutive DNA repair²⁵. We have gone on to characterise one of these proteins, Sso2452, which is representative of the large family of aRadC paralogs²¹.

The structure of Sso2452 is likely to be representative of the aRadC family in general, and is closely related to the unpublished structure of Pho0284, an aRadC family member from *P. horikoshii*. The hexameric quaternary structure of Pho0284 is consistent with the known propensity of RecA-family recombinases to adopt 6-, 7- and 8-membered rings as well as a variety of helical forms⁴². Together with the observation of helical order in the RadB crystal structure¹⁷, this suggests that archaeal RadC proteins can multimerise despite the lack of the polymerisation motif and N-terminal domain that are present in RadA/ Rad51.

The biochemical properties of Sso2452 differentiate it from RadA. The former binds ssDNA around 30-fold more tightly than the latter. Although supporting limited ATP-independent strand exchange *in vitro*, Sso2452 cannot catalyse D-loop formation, suggesting that it does not function as a recombinase *in vivo*. This is consistent with the known properties of *P. furiosus* RadB¹⁸ and the yeast Rad55/Rad57 heterodimer¹⁰. Although Sso2452 binds tightly to ssDNA, its ability to promote limited strand exchange *in vitro* differentiate it clearly from the ssDNA binding protein SSB, which acts as a “trap” for ssDNA in these assays. A role for the archaeal RadA paralogs as “mediators” of RadA catalysed homologous recombination has been suggested, based on the observation that the Rad55/Rad57 heterodimer acts as a mediator of homologous recombination by stimulating strand exchange of RPA-associated DNA by Rad51¹⁰. Our results indicate that, *in vitro*, Sso2452 cannot overcome the inhibitory effect of SSB-coated DNA on strand exchange by RadA. This does not rule out such a function *in vivo*, and reactions could potentially depend upon DNA helicases to help displace SSB, or on other proteins not present in the D-loop assays. Furthermore, our data suggest that Sso2452 can prevent RadA-mediated D-loop formation when incubated with ssDNA either before or after the addition of RadA. Given the significantly higher ssDNA binding affinity of Sso2452 compared to RadA, it is

unsurprising that the former may sequester ssDNA and thus prevent the formation of a RadA nucleoprotein filament as does SSB. However, the observation that Sso2452 prevents a preformed ssDNA:RadA presynaptic filament from completing strand exchange suggests the possibility of an active disassembly process catalysed by Sso2452. Such an anti-recombination activity would be analogous to that observed for the yeast Srs2⁴³ and human Bloom's and Werner's syndrome helicases (reviewed in⁴⁴). The regulation of the initiation of homologous recombination in eukaryotes is important for genome stability and the aRadC paralogs may perform a similar function.

It is also informative to compare the properties of aRadC and the euryarchaeal-specific RadB protein. Both bind tightly to ssDNA and inhibit RadA-mediated strand exchange *in vitro*. Their overall structures are very similar. However, RadB is expressed at a very low level in the cell and does not turn over ATP *in vitro*¹⁸. In contrast, Sso2452 displays a robust ssDNA-dependent ATPase activity similar to that observed for RadA. These data suggest that aRadC and RadB may have different functions *in vivo*, and this is supported by the observation that many euryarchaea including *P. furiosus* have aRadC orthologs in addition to RadB²¹.

Protein interactions between eukaryotic recombination proteins are commonplace and may also be relevant in archaea. *P. furiosus* RadB has been reported to interact with both RadA and the Holliday junction resolving enzyme Hjc¹⁸. aRadC from *S. tokodaii* reportedly interacts with both SSB and RadA²³. This type of interaction assay is susceptible to the generation of false positives when proteins have a DNA binding activity in common. When the possibility for DNA bridging is excluded by the presence of ethidium bromide we have seen no evidence for a stable interaction between *S. solfataricus* Sso2452 and either SSB or RadA *in vitro* (data not shown). Further biochemical and genetic characterisation of the other archaeal RadA paralogs will be required to delineate their role in Homologous Recombination.

Methods & Materials

Preparation of DNA substrates

Oligonucleotides were purified by denaturing acrylamide gel electrophoresis prior to annealing by slow cooling from 95 °C. Substrates were subsequently purified by native acrylamide gel electrophoresis. For D-loop assays, the double-stranded supercoiled plasmid, pUC19, was purified by lysozyme/triton lysis followed by centrifugation on a caesium chloride/ethidium bromide density gradient.

Cloning and purification

The *sso2452* gene from *S. solfataricus* was amplified by the polymerase chain reaction using the following primers: 5' - ATGGTAAGTAGATTATCTACTGGAATGAGAATTGTCACCTTTTAAAAT - 3' 5' - GAACTTTATTTTCTCAACTTTAGTTTTCTGACTCCTCCTTAACCTTC-3' The amplified gene was cloned into a pDEST14 destination vector using the Gateway® cloning system (Invitrogen) for expression in *E. coli* with an N-terminal TEV-cleavable polyhistidine tag (Sso2452-pDEST14 clone generated by the Scottish Structural Proteomics Facility, University of St. Andrews). For protein expression, BL21 Rosetta cells containing the *sso2452* gene were grown at 37°C until an OD₆₀₀ of 1.0 was reached and expression was induced by the addition of 0.4 mM IPTG and incubation at 25°C for ~14 hours. Harvested cells were lysed by sonication in lysis buffer (20 mM Tris-HCl [pH 7.5], 500 mM NaCl, 1 mg/ml DNase, 1 mg/ml lysozyme, 1 mM benzamidine) and the lysate was centrifuged (50,000 x g, 30 min, 4°C), heat-treated (60°C, 20 min), and centrifuged for a further 30 min.

The protein was bound to a nickel-chelating column (HiTrap 5 ml Chelating HP, GE Healthcare) equilibrated with column buffer (20 mM Tris-HCl [pH 7.5], 500 mM NaCl, 10 mM imidazole) and eluted with a linear imidazole gradient (500 mM). Protein-containing fractions were identified by SDS-PAGE, pooled, and further purified on a HiLoad 26/60 Superdex 200 size exclusion column (GE Healthcare) equilibrated with gel filtration buffer (20 mM Tris [pH 7.5], 500 mM NaCl, 1 mM EDTA, and 1 mM DTT). Pure Sso2452 was incubated for ~14 hours at 22°C with 200 ng/ μ L TEV protease to remove the N-terminal polyhistidine tag. Pure protein was analysed by electrospray mass spectrometry to confirm the identity and integrity of protein.

S. solfataricus Alba1, RadA and SSB proteins were also expressed and purified as described previously^{31; 45; 46}.

Crystallization and Structure Solution

Crystallisation conditions were screened using a nano-drop crystallisation robot (Cartesian Honeybee, Genomic Solutions) as part of the Hamilton-Thermo Rhombix system, using commercially available sparse-matrix screens and one in-house screen. Sitting-drop vapour-diffusion was used, with a well volume of 100 μ L in 96 well plates (Griener; 3-square) at 20 °C. Drop sizes were 0.2 μ L (containing 0.1 μ L of protein and 0.1 μ L well solution) and 0.3 μ L (containing 0.2 μ L of protein and 0.1 μ L well solution). The initial hits were optimized using screens generated using an in-house stochastic optimization screen generator. The hanging-drop vapour-diffusion method (1 + 0.5 ml mixture of protein and crystallization solution equilibrated against 450 ml of the latter in 24-well plate) was used for these experiments in 24-well plates (EasyXtal; Qiagen). A cluster of crystals grew over a reservoir containing 9% PEG 8000, 0.1 M MES pH 5.5, and 0.14 M zinc-acetate. The cluster was broken up and two of the larger pieces were mounted in cryo-loops (Hampton Research) and cryo-protected by passing the crystals through a mixture of 10% PEG 8000, 0.1 M MES pH 5.5, 0.15 M zinc-acetate and 30% PEG 400. Crystals were then frozen by rapid immersion in liquid nitrogen, and transferred to sample changer baskets (Molecular Dimensions). Both crystals were screened at BM14UK (Grenoble; ESRF) using the sample changer, and a 2.0 Å data set was collected from a single crystal.

Data were indexed and scaled with XDS and XSCALE⁴⁷ (statistics are shown in Table 2) in space group $p2_12_12$. Analysis of the solvent content suggested that there was one molecule in the asymmetric unit (46% solvent; Matthew's coefficient⁴⁸ 2.26). Initial phases were determined by molecular replacement with Phaser^{49; 50} as implemented in the CCP4⁵¹ suite of programs, using the full length model RecA superfamily ATPase Pho0284 from *P. horikoshii* (pdb:2dr3). Using this initial solution, automated model building in ARP/wARP^{52; 53} successfully built 209 residues out of 262 expected residues, and produced an initial model with an *R*-factor of 23% and a *R*_{free} of 28%. This model was refined using Refmac5^{54; 55}, with manual readjustment using Coot⁵⁶. TLS refinement was also used in the later rounds of the refinement process^{57; 58}. Structure quality was checked using tools within Coot, and also by Molprobity⁵⁹. The final model statistics are shown in Table 1. The structure has been submitted to the Protein Data Bank with PDB code 2w0n.

Despite the fact that there was no pyrophosphate (or a compound with a similar atomic arrangement) in the crystallization or cryoprotection solution, there was a large area of elongated difference density with a clear tetragonal arrangement at one end, and a slightly less-well defined tetragonal arrangement at the other. This density is in the proximity of a Walker motif, in the area occupied by the phosphate backbone of ADP in the 2dr3 structure. This density has been modelled as a pyrophosphate moiety, on the assumption that the molecule is endogenous *E. coli* pyrophosphate that has been retained in the active site of protein during purification and crystallization.

Quantitative RT-PCR

RNA was prepared from *S. solfataricus* cells and quantitative RT-PCR was carried out using the BioRad iQ5 RT-PCR system as described previously⁶⁰. In brief, *S. solfataricus* cells were grown to an OD₆₀₀ of 0.1-0.2 and RNA was extracted using the RNeasy mini kit (Qiagen). Quantitative PCR was carried out in a Bio-Rad iQ5 thermocycler using the Bio-Rad iScript One-Step RT-PCR with SYBR Green 1 kit (BioRad) in accordance with the manufacturer's directions. Reactions were carried out in triplicate. The gene-specific primers used for amplification were as follows:

radA 5' primer: AGCAGCTGGCATTCCATTAT

radA 3' primer: GACCCAAACTCACCGAAGAA

sso0777 5' primer: GGACTTCCGTTTTTCATCTCG

sso0777 3' primer: GGCGATCGACCTCAAAATAA

sso2452 5' primer: TGTGGCAGATGGGATAATCA

sso2452 3' primer: TGCTTATCGTGATCGGTTTG

sso1861 5' primer: GACCGGGAAGTGGTAAATCA

sso1861 3' primer: CTTCTCCCTTTGTGCGACAT

The amplification efficiency of these primers was determined by gene amplification from neat DNA, 1/10, 1/100 and 1/1000 genomic DNA dilutions. Primer efficiencies of 1.06, 1.00, 1.04, and 1.08 were obtained from the *radA*, *sso0777*, *sso2452*, and *sso1861* primer sets, respectively. Ct's (cross-over points) were measured and the ratio of gene expression in UV-irradiated and control samples was quantified as described by Pfaffl⁶¹ using gene *sso0961* as a control as its expression has been shown to remain unchanged after UV irradiation²⁵.

ATPase activity

ATPase assays were performed in a final volume of 300 μ l containing 20 mM MES [pH 6.5], 1 mM DTT, 0.1 mg/mL BSA, 100 mM KCl, 1 μ M protein, and 10 nM DNA (Φ X174 virion or RFI DNA (New England Biolabs)). Reactions were incubated at 60 °C for 1 min and initiated by 1 mM ATP/MgCl₂. At indicated time points, 40 μ l samples were taken and immediately added to 40 μ l 0.3 M chilled perchloric acid on a 96-well plate. Samples were equilibrated to room temperature prior to the addition of malachite green (20 μ l) and, following a 12 min incubation at room temperature, the absorbance at 650 nm was measured on a SpectraMAX 250 Microplate Reader (Molecular Devices). For each reaction, a blank without protein was quantified and subtracted as background from sample reactions. All experiments were carried out in triplicate.

Fluorescence anisotropy

Sso2452-DNA binding was measured by fluorescence anisotropy using a Varian Cary Eclipse fluorimeter equipped with automatic polarisers. 5'-fluorescein labelled single- or double-stranded (15 or 45 nucleotides/base pairs) DNA (final concentration 20 nM) was incubated in anisotropy buffer (20 mM HEPES [pH 7.5], 100 mM NaCl, 1 mM DTT, 0.01% Triton-X100) at 20°C. Following 5-minute equilibration, Sso2452 was titrated into the DNA

solution and anisotropy and total fluorescence intensity (using “magic angle” conditions) were measured after each protein addition. Anisotropy in the absence of protein was subtracted from each data point and all experiments were performed in triplicate. Data were fitted, using Kaleidagraph, to the following equation:

$$A = A_{min} + \left[(D+E+K_D) - \left((D+E+K_D)^2 - (4DE) \right)^{1/2} \right] (A_{max} - A_{min}) / (2D)$$

(A, measured anisotropy; E, variable protein concentration; D, total DNA concentration; A_{min} , anisotropy of free DNA; A_{max} , anisotropy of DNA-protein complex; K_D , dissociation constant)⁶². Since DNA binding is likely to involve some degree of cooperativity and binding of multiple protein molecules to each DNA species, dissociation constants are noted as “apparent”. The data was used to support comparative estimations of DNA binding, not absolute DNA binding affinities.

Analytical size exclusion chromatography

Analytical size exclusion chromatography was performed using a Superdex 200 10/300 column (GE Healthcare) equilibrated with 50 mM Tris-HCl (pH 7.2), 150 mM NaCl, 1 mM EDTA and 1 mM DTT. The column was calibrated with standard proteins of known molecular weight (blue dextran, β -amylase, alcohol dehydrogenase, albumin, carbonic anhydrase and cytochrome C). Sso2452 (149 μ M) was passed over the column either alone or in complex with DNA (80 μ M). In the latter case, Sso2452 was pre-incubated with DNA for 1 hr at room-temperature. Elution volumes were determined from the absorbance at 280 nm and molecular weights were calculated using the standard curve generated from the elution volumes of standard proteins (above). In all cases, samples from each eluted peak were analysed by SDS-PAGE to confirm protein identity (not shown).

Strand-exchange activity

Strand-exchange reactions (80 μ l) were performed in 50 mM Hepes-HCl [pH 7.4], 0.1 mg/ml BSA, 100 mM NaCl, 100 mM KCl, 5 mM ATP/MgCl₂, 0.75 μ M (in nucleotides) [³²P]-radiolabelled ssDNA (50 nucleotides) and 4 μ M protein. Following a 3 min incubation at 60 °C, dsDNA of 25 base pairs (3.75 μ M in nucleotides) was added to initiate the reaction. At indicated time points, samples (10 μ l) were added to chilled stop solution (10 mM Tris [pH 7.5], 20 mM EDTA, 100 mM NaCl, 0.5% SDS, 1 mg/ml proteinase K) and incubated at room temperature for 15 min. Labelled DNA products were separated on a native 12% polyacrylamide:TBE gel at 130 V for 3 hours and analysed by phosphorimaging.

To study the effect of additional proteins, the [³²P]-radiolabelled ssDNA (50 nucleotides) was incubated at 60 °C for 3 min with SSB (10 μ M), Albal (10 μ M), or RadA (4 μ M) prior to the addition of Sso2452 (4 μ M). Following a further 3-minute incubation, dsDNA was added to initiate the reaction³⁰. Appropriate controls were included for each assay, including two end-point assays (reactions stopped 10 min after initiation), one performed in the absence of protein and a second performed in the absence of ATP/MgCl₂.

D-loop formation

D-loop reactions (50 μ l) contained 5' [³²P]-labelled 80-mer ssDNA (3 μ M in nucleotides) and 5 μ M SsoRadA, RecA or Sso2452 in D-loop buffer (50 mM HEPES-HCl [pH 7.4], 1 mM DTT, 0.1 mg/ml BSA, 100 mM NaCl, 100 mM KCl, 2 mM ATP, 15 mM MgCl₂). After 5 min at 60 °C, the reaction was initiated by the addition of supercoiled pUC19 dsDNA (300 μ M in nucleotides) and 10 μ l samples were added to ½ vol. chilled stop solution (10 mM Tris [pH 7.4], 20 mM EDTA, 100 mM NaCl, 3% SDS, 1 mg/ml proteinase K, 100 mM

MgCl₂) at specified time points (time points: 1, 5, 10, 20, 30 min). Following a 15 min incubation at room temperature to allow for proteinase K digestion, samples were mixed with 1/5 vol. loading dye (70% glycerol, 0.1% bromophenol blue) and analysed by electrophoresis through 0.8% agarose:TBE gels at 4V/cm for 3 hours. Gels were dried onto 3 mm Whatman paper at 50°C for 2 hours and visualised by autoradiography.

In order of addition experiments, ssDNA was pre-incubated with protein 1 for 5 min prior to the addition of protein 2. In the event of introducing protein 3, reactions were incubated for a further 5 min prior to addition of dsDNA to initiate the reaction. Reactions were stopped after: 1, 5, 10, 20, 30 min. Final protein concentrations were as follows: Sso2452 and RadA, 5 μ M; SSB, 1 μ M.

To investigate the concentration-dependent effect of Sso2452 inhibition of RadA-catalysed D-loop formation, samples were incubated with increasing concentrations of Sso2452 (0.5, 1, 2, 4, 6 μ M). Reactions were initiated by dsDNA and incubated for 30 min prior to the addition of chilled stop solution.

Acknowledgments

We thank Michael McIlwraith (Cancer Research UK) for help developing the D-loop assay. Thanks to the mass spectrometry facility and to the Scottish Structural Proteomics Facility at the University of St Andrews. Thanks to Stuart MacNeill for critical reading of this manuscript, and to the Biotechnology and Biological Sciences Research Council for financial support.

References

1. Cox MM. The bacterial RecA protein as a motor protein. *Annu Rev Microbiol.* 2003; 57:551–77. [PubMed: 14527291]
2. Kowalczykowski SC, Eggleston AK. Homologous pairing and DNA strand-exchange proteins. *Ann. Rev. Biochem.* 1994; 63:991–1043. [PubMed: 7979259]
3. Sonoda E, Sasaki MS, Buerstedde JM, Bezzubova O, Shinohara A, Ogawa H, Takata M, Yamaguchi-Iwai Y, Takeda S. Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death. *Embo J.* 1998; 17:598–608. [PubMed: 9430650]
4. McGlynn P, Lloyd RG. Recombinational repair and restart of damaged replication forks. *Nat Rev Mol Cell Biol.* 2002; 3:859–70. [PubMed: 12415303]
5. Sung P, Krejci L, Van Komen S, Sehorn MG. Rad51 recombinase and recombination mediators. *J Biol Chem.* 2003; 278:42729–32. [PubMed: 12912992]
6. Masson JY, Tarsounas MC, Stasiak AZ, Stasiak A, Shah R, McIlwraith MJ, Benson FE, West SC. Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev.* 2001; 15:3296–307. [PubMed: 11751635]
7. Liu N, Schild D, Thelen MP, Thompson LH. Involvement of Rad51C in two distinct protein complexes of Rad51 paralogs in human cells. *Nucleic Acids Res.* 2002; 30:1009–15. [PubMed: 11842113]
8. Thacker J. The RAD51 gene family, genetic instability and cancer. *Cancer Lett.* 2005; 219:125–35. [PubMed: 15723711]
9. Liu Y, Tarsounas M, O'Regan P, West SC. Role of RAD51C and XRCC3 in genetic recombination and DNA repair. *J Biol Chem.* 2007; 282:1973–9. [PubMed: 17114795]
10. Sung P. Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein a to promote DNA strand exchange by Rad51 recombinase. *Genes Dev.* 1997; 11:1111–1121. [PubMed: 9159392]
11. Kelman Z, White MF. Archaeal DNA replication and repair. *Curr Opin Microbiol.* 2005; 8:669–76. [PubMed: 16242991]
12. Bell SD, Magill CP, Jackson SP. Basal and regulated transcription in Archaea. *Biochem Soc Trans.* 2001; 29:392–5. [PubMed: 11497995]

13. Bell SD, Jackson SP. Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. *Trends Microbiol.* 1998; 6:222–8. [PubMed: 9675798]
14. Hopfner KP, Karcher A, Craig L, Woo TT, Carney JP, Tainer JA. Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell.* 2001; 105:473–85. [PubMed: 11371344]
15. Shin DS, Chahwan C, Huffman JL, Tainer JA. Structure and function of the double-strand break repair machinery. *DNA Repair (Amst).* 2004; 3:863–73. [PubMed: 15279771]
16. Lin Z, Kong H, Nei M, Ma H. Origins and evolution of the recA/RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. *Proc Natl Acad Sci U S A.* 2006; 103:10328–33. [PubMed: 16798872]
17. Akiba T, Ishii N, Rashid N, Morikawa M, Imanaka T, Harata K. Structure of RadB recombinase from a hyperthermophilic archaeon, *Thermococcus kodakaraensis* KOD1: an implication for the formation of a near-7-fold helical assembly. *Nucleic Acids Res.* 2005; 33:3412–23. [PubMed: 15956102]
18. Komori K, Miyata T, DiRuggiero J, Holley-Shanks R, Hayashi I, Cann IK, Mayanagi K, Shinagawa H, Ishino Y. Both RadA and RadB are involved in homologous recombination in *Pyrococcus furiosus*. *J. Biol. Chem.* 2000; 275:33782–33790. [PubMed: 10903318]
19. Guy CP, Haldenby S, Brindley A, Walsh DA, Briggs GS, Warren MJ, Allers T, Bolt EL. Interactions of RadB, a DNA repair protein in archaea, with DNA and ATP. *J Mol Biol.* 2006; 358:46–56. [PubMed: 16516228]
20. Waters E, Hohn MJ, Ahel I, Graham DE, Adams MD, Barnstead M, Beeson KY, Bibbs L, Bolanos R, Keller M, Kretz K, Lin X, Mathur E, Ni J, Podar M, Richardson T, Sutton GG, Simon M, Soll D, Stetter KO, Short JM, Noordewier M. The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc Natl Acad Sci U S A.* 2003; 100:12984–8. [PubMed: 14566062]
21. Haldenby S, White MF, Allers T. RecA family proteins in archaea: RadA and its cousins. *Biochem Soc Trans.* 2009; 37:102–7. [PubMed: 19143611]
22. Wang J. Recent cyanobacterial Kai protein structures suggest a rotary clock. *Structure.* 2005; 13:735–41. [PubMed: 15893664]
23. Sheng D, Zhu S, Wei T, Ni J, Shen Y. The in vitro activity of a Rad55 homologue from *Sulfolobus tokodaii*, a candidate mediator in RadA-catalyzed homologous recombination. *Extremophiles.* 2008; 12:147–57. [PubMed: 17938853]
24. Cubeddu L, White MF. DNA damage detection by an archaeal single-stranded DNA-binding protein. *J. Mol. Biol.* 2005; 353:507–516. [PubMed: 16181640]
25. Götz D, Paytubi S, Munro S, Lundgren M, Bernander R, White MF. Responses of hyperthermophilic crenarchaea to UV irradiation. *Genome Biol.* 2007; 8:R220. [PubMed: 17931420]
26. Fröls S, Gordon PM, Panlilio MA, Duggin IG, Bell SD, Sensen CW, Schleper C. Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV damage. *J Bacteriol.* 2007; 189:8708–8718. [PubMed: 17905990]
27. Abella M, Rodriguez S, Paytubi S, Campoy S, White MF, Barbe J. The *Sulfolobus solfataricus* radA paralogue sso0777 is DNA damage inducible and positively regulated by the Sta1 protein. *Nucl. Acids Res.* 2007; 35:6788–6797. [PubMed: 17921500]
28. Seitz EM, Brockman JP, Sandler SJ, Clark AJ, Kowalczykowski SC. RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev.* 1998; 12:1248–53. [PubMed: 9573041]
29. Jain SK, Cox MM, Inman RB. On the role of ATP hydrolysis in RecA protein-mediated DNA strand exchange. III. Unidirectional branch migration and extensive hybrid DNA formation. *J Biol Chem.* 1994; 269:20653–61. [PubMed: 8051165]
30. Bugreev DV, Mazin AV. Ca²⁺ activates human homologous recombination protein Rad51 by modulating its ATPase activity. *Proc Natl Acad Sci U S A.* 2004; 101:9988–93. [PubMed: 15226506]

31. Jelinska C, Conroy MJ, Craven CJ, Hounslow AM, Bullough PA, Waltho JP, Taylor GL, White MF. Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. *Structure*. 2005; 13:963–71. [PubMed: 16004869]
32. Morrical SW, Cox MM. Stabilization of recA protein-ssDNA complexes by the single-stranded DNA binding protein of *Escherichia coli*. *Biochemistry*. 1990; 29:837–43. [PubMed: 2186808]
33. Sigurdsson S, Van Komen S, Bussen W, Schild D, Albala JS, Sung P. Mediator function of the human Rad51B-Rad51C complex in Rad51/RPA-catalyzed DNA strand exchange. *Genes Dev*. 2001; 15:3308–18. [PubMed: 11751636]
34. Lusetti SL, Cox MM. The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu Rev Biochem*. 2002; 71:71–100. [PubMed: 12045091]
35. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr*. 2004; 60:2256–68. [PubMed: 15572779]
36. Logan KM, Forget AL, Verderese JP, Knight KL. ATP-mediated changes in cross-subunit interactions in the RecA protein. *Biochemistry*. 2001; 40:11382–9. [PubMed: 11560486]
37. Aihara H, Ito Y, Kurumizaka H, Yokoyama S, Shibata T. The N-terminal domain of the human Rad51 protein binds DNA: structure and a DNA binding surface as revealed by NMR. *J Mol Biol*. 1999; 290:495–504. [PubMed: 10390347]
38. Kinebuchi T, Kagawa W, Kurumizaka H, Yokoyama S. Role of the N-terminal domain of the human DMC1 protein in octamer formation and DNA binding. *J Biol Chem*. 2005; 280:28382–7. [PubMed: 15917243]
39. Shin DS, Pellegrini L, Daniels DS, Yelent B, Craig L, Bates D, Yu DS, Shivji MK, Hitomi C, Arvai AS, Volkmann N, Tsuruta H, Blundell TL, Venkitaraman AR, Tainer JA. Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *Embo J*. 2003; 22:4566–76. [PubMed: 12941707]
40. Chen Z, Yang H, Pavletich NP. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature*. 2008; 453:489–4. [PubMed: 18497818]
41. Chen LT, Ko TP, Chang YW, Lin KA, Wang AH, Wang TF. Structural and functional analyses of five conserved positively charged residues in the L1 and N-terminal DNA binding motifs of archaeal RADA protein. *PLoS ONE*. 2007; 2:e858. [PubMed: 17848989]
42. Wang TF, Chen LT, Wang AH. Right or left turn? RecA family protein filaments promote homologous recombination through clockwise axial rotation. *Bioessays*. 2008; 30:48–56. [PubMed: 18081011]
43. Krejci L, Van Komen S, Li Y, Villemain J, Reddy MS, Klein H, Ellenberger T, Sung P. DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature*. 2003; 423:305–9. [PubMed: 12748644]
44. Bohr VA. Rising from the RecQ-age: the role of human RecQ helicases in genome maintenance. *Trends Biochem Sci*. 2008; 33:609–20. [PubMed: 18926708]
45. Ariza A, Richard DJ, White MF, Bond CS. Conformational flexibility revealed by the crystal structure of a crenarchaeal RadA. *Nucleic Acids Res*. 2005; 33:1465–73. [PubMed: 15755748]
46. Wadsworth RI, White MF. Identification and properties of the crenarchaeal single-stranded DNA binding protein from *Sulfolobus solfataricus*. *Nucleic Acids Res*. 2001; 29:914–920. [PubMed: 11160923]
47. Kabsch, w. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. App. Cryst*. 1993; 26:795–800.
48. Matthews BW. Solvent content of protein crystals. *Journal of Molecular Biology*. 1968; 33:491–497. [PubMed: 5700707]
49. Storoni LC, McCoy AJ, Read RJ. Likelihood-enhanced fast rotation functions. *Acta Crystallographica Section D-Biological Crystallography*. 2004; 60:432–438.
50. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. *Acta Crystallographica Section D-Biological Crystallography*. 2005; 61:458–464.
51. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr*. 1994; 50:760–3. [PubMed: 15299374]

52. Morris RJ, Perrakis A, Lamzin VS. ARP/wARP's model-building algorithms. I. The main chain. *Acta Crystallographica Section D-Biological Crystallography*. 2002; 58:968–975.
53. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. *Nat Struct Biol*. 1999; 6:458–63. [PubMed: 10331874]
54. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*. 1997; 53:240–55. [PubMed: 15299926]
55. Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallographica Section D-Biological Crystallography*. 1999; 55:247–255.
56. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 2004; 60:2126–32. [PubMed: 15572765]
57. Winn MD, Murshudov GN, Papiz MZ. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Macromolecular Crystallography, Pt D*. 2003; 374:300–321.
58. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallographica Section D-Biological Crystallography*. 2001; 57:122–133.
59. Davis IW, Murray LW, Richardson JS, Richardson DC. MolProbity: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Research*. 2004; 32:W615–W619. [PubMed: 15215462]
60. Dorazi R, Götz D, Munro S, Bernander R, White MF. Equal rates of repair of DNA photoproducts in transcribed and non-transcribed strands in *Sulfolobus solfataricus*. *Mol Microbiol*. 2007; 63:521–9. [PubMed: 17163966]
61. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001; 29:e45. [PubMed: 11328886]
62. Reid SL, Parry D, Liu HH, Connolly BA. Binding and recognition of GATATC target sequences by the EcoRV restriction endonuclease: a study using fluorescent oligonucleotides and fluorescence polarization. *Biochemistry*. 2001; 40:2484–94. [PubMed: 11327870]
63. Salerno V, Napoli A, White MF, Rossi M, Ciaramella M. Transcriptional response to DNA damage in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res*. 2003; 31:6127–38. [PubMed: 14576299]

Figure 1. Phylogenetic analysis of archaeal RadA paralogs

This unrooted bootstrapped phylogenetic tree shows the monophyletic family of uncharacterised archaeal aRadC proteins, with RadB from *Pyrococcus horikoshii* and RadA from *S. solfataricus* as representative outgroups. Each protein is represented by a three-letter species code followed by the gene number from the respective genome sequences (Sso, *S. solfataricus*; Pho, *P. horikoshii*; Neq, *Nanoarchaeum equitans*; Nmar, *Nitrosopumilus maritimus*; Kcr, *Korarchaeum cryptofilum*). This neighbour-joining tree was generated from a ClustalW alignment of the proteins using the programme MacVector, with pairwise distances between sequences uncorrected. The bootstrap values shown at each node represent the percentage of all trees (1000 total) agreeing with this topology.

Figure 2. ATPase activity and DNA binding affinity of Sso2452

(A) Rate of ATP hydrolysis by Sso2452 in the presence and absence of single- and double-stranded DNA (ΦX174 virion and RFI, respectively). Data points represent the mean of triplicate measurements, with standard errors indicated.

(B) DNA binding affinity of Sso2452 measured by change in fluorescence anisotropy. Sso2452 was titrated into a solution containing a 5'-fluorescein labelled oligonucleotide (single or double-stranded). Fluorescence anisotropy increases as a function of protein-DNA binding enabling equilibrium dissociation constants to be calculated. The means of triplicate measurements were plotted, and standard errors are shown. Closed circles, ssDNA RadA; closed squares, dsDNA RadA; open circles, ssDNA Sso2452; open squares, dsDNA Sso2452.

(C) Analytical size exclusion chromatography of Sso2452:DNA complex. A Superdex 200 10-300 gel filtration column (GE Healthcare) was equilibrated with 50 mM Tris-HCl [pH 7.2], 0.15 M NaCl and calibrated with proteins of known molecular weight, whose elution

positions are indicated at the top of the graph. Sso2452 (149 μM) was passed through the column alone (black trace) and in complex with a 34-nt DNA oligonucleotide (80 μM) (grey line). The maximum absorbance (280 nm) was normalised to 1. Original absorbance values (280 nm): 600 kDa peak, 0.31; 250 kDa peak, 0.32; 30 kDa peak, 0.54. The presence of Sso2452 in eluted fractions was confirmed by SDS-PAGE.

Figure 3. Strand-exchange Activity of Sso2452 and RadA

(A) Schematic of the strand-exchange assay and substrate design. Length (nucleotides) of oligonucleotide is indicated above the strand. Black circles denote the ^{32}P radiolabel. (B) Sso2452 ($4\ \mu\text{M}$) or RadA ($4\ \mu\text{M}$) was pre-incubated with [^{32}P]-radiolabelled ssDNA (50-mer) for 3 min prior to addition of dsDNA (25-mer) and incubation for up to 10 min at $60\ ^\circ\text{C}$. Both show strand exchange activity (C) Order of addition experiments. Protein (1) was pre-incubated with ssDNA for 3 min prior to the addition of the second protein. Following a further 3 min incubation, dsDNA was added to initiate the reaction. (D) Strand exchange reaction with Alba1 ($10\ \mu\text{M}$) and SSB ($10\ \mu\text{M}$). Reaction performed as described above.

Alba1 supports strand exchange activity but SSB does not. For all panels, time points shown are: 0.5, 1, 3, 5, 8, 10 min. Controls used were: c1, DNA species in the absence of protein after 10 min at 60 °C; c2, size marker for the strand exchange product; c3, reactions lacking ATP/MgCl₂ after 10 min.

Figure 4. D-loop formation assays for RadA and Sso2452

D-loop reactions were performed over a 30 min time course at 60 °C. [³²P]-labelled ssDNA (80 nucleotides) was incubated with protein for 5 min prior to the addition of dsDNA to initiate the reaction. Where additional proteins were added, the order of addition is indicated. Each protein addition was followed by a 5 min incubation. In all cases, reactions were initiated by dsDNA. (A) Schematic of D-loop reaction. (B) D-loop formation by RecA at 37 °C (time course: 1, 2, 4, 8, 16 min), RadA (time course: 1, 5, 10, 20, 30 min), and Sso2452 (time course as for RadA). (C) Left hand: Order of addition experiments where Sso2452 was pre-incubated with ssDNA prior to the addition of RadA (or vice versa). Time points: 1, 5, 10, 20, 30 min. Right-hand: Inhibition of RadA as a function Sso2452 concentration (Sso2452 titration: 0.5, 1, 2, 4, 6 μM) Reactions were initiated by dsDNA and were incubated for 30 min prior to addition of chilled stop solution. (D) Order of addition experiments with SSB, RadA, and Sso2452. Time points: 1, 5, 10, 20, 30 min. Controls in

panels B-D: c1, no protein (reaction stopped after 30 min); c2, no ATP/MgCl₂ (reaction stopped after 30 min); c3, RecA-catalysed D-loop formation after 1 min incubation at 37 °C.

Figure 5. Structural biology of SSo2452

(A) A ribbon representation of the monomer of SSo2452 colored in slate blue. The four Zn^{2+} ions which are believed to be an artifact of crystallization are shown as grey spheres. The termini of the disordered L2 DNA binding loop from residues Q166 to G178 are shown as yellow spheres. The pyrophosphate modeled into the electron density is shown as spheres.

(B) Superposition of Sso2452 (shown as slate blue wire) with the hexameric Pho0284 (shown as wire; colored differently for each monomer with the superimposing monomer shown in red). In Pho0284 ADP (shown in space filling) binds at the interface of the hexamer. Two of the Zn^{2+} ions in our crystal (shown as grey spheres) disrupt this crystal packing arrangement and thus ATP binding. The PPi molecule shown in 5A overlaps with the phosphates of the ADP molecule in Pho0284.

(C) Structural superposition of the *S. solfataricus* RadA (yellow ribbon; PDB code 2bke) and Sso2452 (colored as above). RadA has an extra N-terminal domain implicated in multimerisation and dsDNA binding, linked to the core domain by a short polymerization motif (PM) that includes a conserved phenylalanine residue (F73) that acts as the “ball” in the “ball and socket” joint formed in RadA filaments. The ATPase domain which consists of central beta sheet and flanking alpha helices is common to both structures.

(D) Sequence comparison for the L1 and L2 ssDNA binding loops in RadA and aRadC. Residues in bold are discussed in the text.

Table 1

Relative abundance of mRNA transcripts encoding RadA and paralogs in *S. solfataricus* grown in the presence or absence of UV damage

Gene	Ct value (mean $\pm$ SE)	mRNA level relative to RadA %	Relative gene expression (UV/control) after UV treatment ^a			
			30 min	60 min	90 min	120 min
<i>radA</i>	13.7 $\pm$ 0.1	100	1.00	1.04	1.01	1.00
<i>ss00777</i>	17.5 $\pm$ 0.3	7	1.00	0.99	1.00	1.00
<i>ss01861</i>	16.3 $\pm$ 0.2	17	1.00	0.94	0.96	0.94
<i>ss02452</i>	16.8 $\pm$ 0.2	11	0.97	0.97	0.97	0.93
<i>ss00961</i>	22.3 $\pm$ 0.1	N/A	1	1	1	1

The relative levels of mRNA encoding RadA and paralogs in *S. solfataricus* were analysed by quantitative one-step RT-PCR. A total of 100 μ g of mRNA was used in each reaction, and measurements were performed in triplicate, with standard errors shown. The Ct value is the cycle at which levels of PCR product cross a predefined threshold, with lower values indicating a more abundant transcript.

Genespecific RT-PCR primer efficiencies (*E*) were calculated from standard curves, according to Pfaffl²⁴

^aThe ratio of gene expression refers to the relative levels of mRNA transcript present in cultures grown in the absence of UV irradiation or following exposure to 200 J/m² of UV radiation. A value of 1 represents no change in transcript level. Ratios were calculated according to Pfaffl²⁴ using the *ss00961* transcript as a reference, as it is unaffected by UV radiation. [25] and [26]

Table 2**Crystallographic data and refinement**

<i>Data statistics</i>	
Wavelength (Å)	1.0
Resolution (highest-shell Å)	30.0–2.0 (2.10–2.00)
Space group	$P2_12_12$
Temperature (K)	100
Detector	MAR 225
Unit-cell parameters	$a = 41.1 \text{ Å}$, $b = 169.6 \text{ Å}$, $c = 39.4 \text{ Å}$, $\alpha = \beta = \gamma = 90^\circ$
$V_m (\text{Å}^3/\text{Da})$	2.26
Solvent content (%)	46
Total number of reflections	93,976 (12,776)
Unique reflections	19,173 (2588)
$I/\sigma(I)$	14.6 (3.3)
Average redundancy	4.9 (4.9)
Completeness (%)	98.3 (99.2)
$R_{\text{merge}} (\%)$	6.8 (48.7)
<i>Refinement</i>	
Resolution (highest-shell Å)	29.51–2.0 (2.05–2.00)
$R (\%)$	20 (26)
$R_{\text{free}} (\%)$ (5% of reflections)	25 (47)
Overall B -factor (Å^2)	29.125
RMSD bonds (Å)/angles ($^\circ$)	0.007/1.017
Protein atoms	1787
Water atoms	98
Zinc atoms	4
Pyrophosphate atoms	9