Co-expression and purification of the RadA recombinase with the RadB paralog from *Haloferax volcanii* yields heteromeric ring-like structures.

Authors:

Bushra B. Patoli[1, a,*], Jody A. Winter[1, b] Atif A. Patoli[1, a], Robin M. Delahay[2] and Karen A. Bunting[1, c]

Affiliations:

[1]School of Biology, Queens Medical Centre, University of Nottingham, Nottingham, UK, NG7 2UH.

[2]School of Medicine, University of Nottingham, Nottingham, UK, NG7 2UH. Email: rob.delahay@nottingham.ac.uk

[a]Institute of Microbiology, University of Sindh, Jamshoro, Pakistan (Present address). Email: bushrapatoli@gmail.com

[b]Nottingham Trent University, Nottingham, UK, NG11 8NS (Present address). Email: jody.winter@ntu.ac.uk

[c]Albumedix Ltd, Nottingham, UK, NG7 1FD (Present address). Email: karen.bunting@slidingclamp.com

[*]**Correspondence:** Bushra B. Patoli, Institute of Microbiology, University of Sindh, Jamshoro, Pakistan. Tel.: +92 3113005252. E.mail.bushrapatoli@gmail.com

Keywords: Homologous recombination, RadA, *Haloferax volcanii*, Co-expression, Halophilic protein purification.

Abbreviations: Immobilized metal affinity chromatography (IMAC); Discrete Optimised Protein Energy (DOPE); ATPase domain (AD); N-terminal domain (NTD); Size Exclusion Chromatography (SEC); Electron Microscopy (EM)

Abstract

The study of archaeal proteins and the processes they contribute to poses particular challenges due to the often extreme environments in which they function. DNA recombination, replication and repair proteins of the halophilic euryarchaeon, *Haloferax volcanii* (Hvo) are of particular interest as they tend to resemble eukaryotic counterparts in both structure and activity and genetic tools are available to facilitate their analysis. In the present study, we show using bioinformatics approaches that the Hvo RecA-like protein RadA is structurally similar to other recombinases although is distinguished by a unique acidic insertion loop. To facilitate expression of Hvo RadA a co-expression approach was used, providing its lone paralog, RadB, as a binding partner. At present, structural and biochemical characterization of Hvo RadA is lacking. Here, we describe for the first time co-expression of Hvo RadA with RadB and purification of these proteins as a complex under *in vitro* conditions. Purification procedures were performed under high salt concentration (>1 M sodium chloride) to maintain the solubility of the proteins. Quantitative densitometry analysis of the co-expressed and co-purified RadA-B complex estimated the ratio of RadA to RadB to be 4:1, which suggests that the proteins interact with a specific stoichiometry. Based on a combination of analyses, including size exclusion chromatography, Western blot and electron microscopy observations we suggest that RadA multimerizes into a ring-like structure in the absence of DNA and nucleoside cofactor.

Introduction

Homologous recombination is one of the principal pathways to repair deleterious DNA damage such as double-strand breaks. The basic mechanism of homologous recombination, comprising homologous base-pairing and a strand exchange reaction, is similar in all forms of life, however it varies in complexity and enzymology. Various enzymes participate in the recombination process, including the key recombinase which belongs to the RecA family of proteins (1). These enzymes are well conserved in all three domains of life and are termed RecA in bacteria (2), RadA in archaea (3), and Rad51 in eukaryotes (4). The strand exchange protein identified in archaeal RadA is more similar to the eukaryotic Rad51 protein (~40% amino acid identity) than to its prokaryotic counterpart RecA in eubacteria (~20% amino acid identity) (3, 5, 6). Crystallographic structure analysis of archaeal RadA has contributed significantly to our understanding of the structure and function of eukaryotic Rad51 (7, 8). Various accessory proteins also participate in the strand exchange reaction and these are usually termed recombination mediators. Many recombination mediators are encoded by genes that arise as a result of a duplication event within the genome, and are termed paralogs. In yeast, the paralogous proteins Rad55 and Rad57 share some similarity to Rad51 and function as a heterodimeric complex which may stabilize the Rad51-assembled nucleoprotein filament (9). Five Rad51 paralogs, XRCC2, XRCC3, RAD51B/RAD51L1, RAD51C/RAD51L2, and RAD51D/RAD51L3, are found in eukaryotes. These paralogs form various heteromeric complexes, mostly dimers such as RAD51B-RAD51C, RAD51D-XRCC2, RAD51C-XRCC3 or tetramers such as RAD51B-RAD51C-RAD51D-XRCC2 (10).

The identification and characterisation of RecA-like proteins in both crenarchaea and euryarchaea shows that, in common with eukaryotic Rad51, archaeal RadA proteins exhibit a high level of

86 diversity. This may reflect that the archaea demonstrate an evolutionary mixture of replication,
87 repair and recombination functions in between simple bacterial and more complex eukaryotic
88 forms of life.
89 RadB was the first RadA paralog to be characterised in archaea and its presence is confined to the
90 Euryarchaeota. RadB lacks the N-terminal domain of RadA (11) and strand exchange activity (12).
91 Interaction of RadB with RadA has been demonstrated in *Pyrococcus furiosus* (Pfu) and it was
92 proposed that RadB functions in a manner analogous to the yeast Rad55-57 proteins in the strand
93 exchange reaction (12). Genetic analysis of *radB* from the euryarchaeon *Haloferax volcanii* (Hvo)
94 demonstrates that RadB functions in the homologous recombination pathways in concert with
95 RadA (13, 14).
96 Hvo RadB has been successfully over-expressed in *Escherichia coli* (15) and both Hvo RadA and
97 RadB proteins have been purified after conditional over-expression in their native host (16, 17). A
98 recent study has confirmed the *in vivo* interaction of Hvo RadA and RadB, exploiting mass
99 spectrometry to identify co-purifying proteins following IMAC chromatography (18). However,
100 both quantity and purity of the resulting proteins has been insufficient for structural and
101 biochemical analyses previously described for other DNA-associated halophilic proteins such as
102 Hvo PCNA and RPA3 (19, 20). Additionally, when over-expressed in *E. coli*, recombinant Hvo
103 RadA has also been found to co-purify with DNA, providing further impediments to downstream
104 analyses since conventional approaches to remove contaminating DNA, such as denaturation and
105 nuclease treatment, are likely to be detrimental to subsequent characterization [17].
106 The study of recombination proteins in the hyperthermophilic crenarchaeon, *Sulfolobus tokodaii*
107 (St), has shown that over-expression and characterization of StRad55B, a paralog of StRadA, was
108 possible only when it was co-expressed with StRadA protein (21). We therefore explored co-

expression of Hvo RadA with RadB as both a novel strategy for maintenance and optimisation of RadA protein stability in the non-halophilic host, *E. coli*, and as a potential means to limit RadA interaction with host DNA to facilitate its purification. Using this approach we successfully demonstrate soluble over-expression of the Hvo RadA-RadB complex, provide clear evidence for RadA-RadB protein-protein interaction and present robust methodology to enable their further characterization and study.

Materials and Methods

Bacterial strains, plasmids and growth conditions.

E. coli chemically competent strains DH5α™ (Invitrogen) and Rosetta 2 (DE3) (Novagen) were used for cloning and gene expression, respectively. The PCR-amplified target gene was cloned first into the Zero Blunt PCR vector (Invitrogen) and sequenced to exclude amplification errors. The pET11 expression vector (Novagen) was used when single overexpression of *radA* or *radB* was desired. For co-expression and purification of RadA and RadB, the *E. coli* Rosetta 2 (DE3) strain was transformed with pET-Duet-1-based plasmid constructs (Novagen).

Bacterial cultures were grown at 37°C in Luria-Bertani (LB) broth in a shaking incubator or on LB agar plates supplemented with ampicillin (100 μg ml^{-1}), chloramphenicol (34 μg ml^{-1}) or kanamycin (34 μg ml^{-1}) as appropriate.

Knock out of the *arnA* gene from *E. coli* Rosetta2 (DE3) expression strain

A mutant of the Rosetta2 (DE3) expression strain carrying a kanamycin resistance cassette (kan^R) in place of the *arnA* gene (arnA736 (del):: kan^R) was generated by P1 phage-mediated transduction

(22). Donor strain, obtained from the *E. coli* Genetic Stock center (CGSC), was grown at 37°C to an OD_{650} of 0.3-0.4 to add the P1 phage stock (~10^7-10^8 PFU). After 3-4 h incubation, purified lysate was used to transduce the kan^R marker into recipient Rosetta2 (DE3) cells. Recipient strain was grown at 37°C to OD_{650} of 0.8, pelleted, re-suspended in 0.1 M $MgSO_4$, 5 mM $CaCl_2$ and incubated with purified lysate for 25 minutes at 37°C. After addition of sodium citrate to a final concentration of 0.5 M to chelate calcium ions and prevent further total lysis of the recipient strain, transductants were plated and sub-cultured to purity on LB agar containing kanamycin (34 μg ml^{-1}) and chloramphenicol (34 μg ml^{-1}).

Molecular methods

Hvo *radA* gene (Accession number: U45311) was PCR-amplified from the *H. volcanii* wild-type DS2 strain (23). The forward primer (*radA*6HisF) contained an NdeI restriction site (underlined) and a 6xHis tag 5'GACCT**<u>CATATG</u>CATCACCATCACCATCAC**ATGGCAGAAGACGACCTC-3' and the reverse primer (*radA*BamR) encoded a *Bam*HI restriction site (underlined) 5'-GCAAT**<u>GGATCC</u>**TTATTACTCGGGCTTGAGACCGGCGTCCTG-3'. The His-tagged *radA* gene was cloned first into the Zero Blunt PCR vector for sequencing and then sub-cloned for overexpression either into pET11 (using NdeI and BamHI restriction sites) or pET-Duet1 plasmid at multiple cloning site 2 (MCS-2) (using the NdeI and EcoRV restriction sites). The Hvo *radB* coding sequence (690bp) was excised from plasmid pCPG42 using the NdeI and BamHI restriction sites (15) . After sequence verification in a sub-cloning vector, *radB* was cloned into MCS-1 of pET-Duet1 using the NcoI and HindIII restriction sites to yield the pBPRAD2 construct. The relevant properties of the strains are listed in Table 1.

Over-expression, purification and analysis of RadA and RadB proteins

Rosetta2 (DE3) or Rosetta2 (DE3) Δ*arnA E. coli* transformed with the pBPRAD2 construct were used for over-expression and purification of co-expressed His-tagged RadA and un-tagged RadB proteins. An overnight culture was inoculated into baffle flasks containing 2.4 L LB broth supplemented with antibiotics as appropriate, and grown at 37°C. Over-expression of *radA-B* was induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at an OD_{600} of 0.4 – 0.6 with incubation for a further 3 hrs in a shaking incubator at 30°C. The cell pellet was harvested by centrifugation and lysed by sonication on ice in buffer A (50 mM HEPES, pH7.0, 10 mM imidazole and 1 M NaCl) containing EDTA-free protease inhibitor cocktail (Roche). The cell lysate was clarified by centrifugation at 16,000 x *g* for 30 minutes at 4°C to remove insoluble debris.

Soluble proteins in the supernatant were used for downstream purification of the complex using Cobalt-based immobilized metal affinity chromatography (IMAC). A Liquid Chromatography Column (2.5 cm x 10 cm, Sigma-Aldrich) containing 5 mL of Talon® metal affinity resin (Clontech) was equilibrated with 20 column volumes of buffer A (50 mM HEPES pH 7.0, 1 M NaCl and 30 mM imidazole). The clarified soluble fraction was loaded and incubated with rolling for 10 minutes at room temperature. The flow-through and two 20 mL buffer A washes were retained separately in fresh tubes for SDS-PAGE gel analysis. Bound proteins were eluted with 13 mL of buffer B (50 mM HEPES pH 7.0, 1 M NaCl and 300 mM imidazole) following rolling for 5 minutes at room temperature, prior to collection of the eluate. The eluate was loaded onto a 26/60 Superdex 200 (S200) preparative column (GE Healthcare) using an ÄKTA Prime Plus system. The column was pre-equilibrated and run using size exclusion chromatography (SEC) buffer (50

mM HEPES pH 7.0 and 1 M NaCl). The proteins in each IMAC and SEC fraction were analyzed by SDS-PAGE gel. SEC fractions were also analysed by agarose gel electrophoresis to select fractions free from contaminating DNA. Samples containing low concentrations of protein were first concentrated using StrataClean resin (5 μl ml^{-1}) (Stratagene) as previously described (24) prior to gel loading. All samples were adjusted to maintain equivalent loading in all gels.

The identity of RadA and RadB proteins was confirmed using MALDI-TOF-MS analysis of excised bands by the Biopolymer Synthesis and Analysis Unit (BSAU), University of Nottingham. Quantitative reflectance densitometric analysis of Coomassie Brilliant Blue-stained proteins was performed using a BioRad GS-800 calibrated densitometer and Quantity One® Software. Western blots were blocked with 5%, w/v, milk powder in phosphate buffered saline (PBS)-Tween 20, then probed with 10 μL (1/1000 diluted) of alkaline phosphatase conjugated-Mouse Anti-Hexa-His antibodies (Sigma) for 1 h and developed using BCIP/NBT substrate according to manufacturer's instructions (Sigma).

Electron microscopy

Co-expressed, purified RadA-B proteins were concentrated to 1 mg ml^{-1} using a Vivapore 10/20 7500 Da cutoff (Vivascience) and were maintained in 1 M NaCl buffer to preserve the soluble and native state of the proteins. Protein concentrations were determined using the Qubit protein assay kit (Invitrogen). Protein samples were applied to a carbon-Formvar grid (Agar Scientific) and allowed to settle for 10 minutes. The sample was then stained with either 0.5 or 1%, w/v, phosphotungstic acid (PTA; pH 7.0) for 1 minute. The grids were then imaged using a JEOL JEM1010 transmission electron microscope at 100, 000X or 200, 000X magnification.

Homology Modelling

Homology modelling of the Hvo RadA primary sequence was performed using the PyMod 2.0 plugin module for PyMol (25, 26) as a convenient interface to Modeller 9v4 (27). The Pfu RadA structure (1PZN) chain A was used as the highest scoring template and aligned against the Hvo RadA sequence using Clustal Omega. Modeller parameters were adjusted to accommodate automated building of disulfide bridges and the highest level of optimization and refinement, with additional energy minimisation performed on resulting models. Output models with corresponding low scoring Discrete Optimised Protein Energy (DOPE) profiles, indicative of limited modelling errors, were further inspected for agreement with secondary structural elements and diversity in loop disposition prior to final model selection.

Results

Conservation of the ATPase domain of RadA from both halophilic and non-halophilic species

Structural comparison of RadA from *P. furiosus* with RadA homologs in other domains of life reveals similarity in the ATPase domain (AD) although differences are apparent at amino and carboxyl terminal domains of the proteins (Fig. S1a). Conservation of amino acid residues at the Walker A, Walker B and DNA binding (L1 and L2) motifs is also apparent in a protein sequence alignment of Hvo RadA with RadA homologs (Fig. S1b), and the Hvo RadB protein. Aside from this conserved motif organization, Hvo RadA and RadB proteins are otherwise dissimilar, sharing only 18.5%/34.1% identity/similarity at the amino acid level. Hvo RadB is also a somewhat

smaller protein, due to lack of a RadA-equivalent N-terminal domain (NTD) (Fig. S1c).

Along with a characteristic negative charge distribution on the protein surface, a distinct preference for certain amino acids is also a hallmark of halophilic proteins. This is apparent in Hvo RadA as a comparative increase in the acidic residue, aspartate, which is the most common adaptation relative to mesophilic and thermophilic counterparts (Fig. S2). However, whereas a marked reduction in the lysine content and increases in serine and alanine have been reported in other halophilic proteins (28) this is not similarly the case for Hvo RadA. This disparity reflects the lack of any currently known universal form of halophilic adaptation.

It is of particular interest to understand how halophilic proteins have adapted to interact with DNA, given the involvement of basic residues in DNA binding of the phosphate backbone. Guy and others (15) demonstrated that residues in a conserved basic patch (the KHR triplet) were crucial for DNA binding in Hvo RadB. To gain some insight into how preference for particular amino acid usage in Hvo RadA, would likely affect the surface character of Hvo RadA, we built a structural model in Modeller 9v4 for comparative analysis. Superposition of the Hvo RadA model with RadA homologs shows the expected correspondence with the principal structural elements in the N-terminal and ATPase domains (Fig. 1a) although, unlike other RadA proteins, an extensive insertion loop found only in the Hvo RadA sequence (159-182) is also apparent (Fig. 1a and Fig. S1b).

Fig. 1(b) shows the conserved elements of RadA in the Hvo RadA homology model in relation to lysine residues within the molecule. The L1 and L2 loops within the ATPase domain are involved in ssDNA binding. Chen and others (29) identified a number of key positively charged residues in Sso RadA L1 (R217, R223 and R229) that are also conserved in Hvo RadA. Conversely, of the two lysine residues in the N-terminal domain of Sso RadA involved in dsDNA binding, only the

equivalent of K27 is retained in Hvo RadA, with K60 being substituted with as aspartate residue (D52). However, Hvo RadA possesses an alternative lysine at position 42. The majority of the lysine residues within Hvo RadA are located on the face of the molecule containing the NTD, L1 and L2 motifs, suggesting their retention may be related to involvement in DNA-binding.

The large insertion sequence seemingly unique to Hvo RadA (residues 159-182) is of potential interest (Fig. 1 and Fig S1b). Given the lack of homology to other family members the conformation of the modelled loop remains speculative but it can be assumed that it is positioned on the face of the molecule as indicated in Fig. 1(b). The loop region comprises a large number of acidic residues, both aspartate and glutamate, which, in the context of the uncommonly higher proportion of lysines in Hvo RadA relative to other halophilic proteins, may function to maintain solubility of the molecule or promote/stabilize RadA interactions with other proteins.

Differential co-expression and purification of soluble RadA-RadB proteins

His-tagged Hvo *radA* and un-tagged *radB* were co-expressed from pBPRAD2 in the Rosetta 2 (DE3) *E. coli* strain in order to minimise issues with codon usage that can occur when expressing archaeal proteins in a heterologous host (30). SDS-PAGE gel analysis of clarified lysates of induced cultures identified overexpressed proteins at the anticipated molecular mass for 6xHis-RadA (~ 40 kDa) and RadB (~25 kDa) (Fig. 2). However, by comparison with the broadly equivalent amounts of RadA and RadB observed when each protein was expressed from pET11 individually (Fig 2. lanes 2 & 4 respectively), RadB, when co-expressed with RadA, appeared substantially diminished in abundance. This may suggest a possible issue with bicistronic expression, although DUET vectors are engineered for independent promoter control of both multiple cloning sites to avoid such problems. Alternatively, it may be the case that RadA affects

RadB expression or stability; a previous analysis of the RadA:RadB ratio in *P. furiosus* cell extract showed the cellular amount of RadB protein to be ~200 times lower than that of RadA (12). Consequently, the lack of equimolar expression observed may better reflect a more physiological ratio of RadA and RadB proteins and, as such, was not considered to be an impediment to further analysis.

H. volcanii requires 2 M sodium chloride to maintain internal osmotic balance in artificial media (23). Consideration is therefore required around buffer selection to maintain salt levels to promote solubility of halophilic proteins and avoid aggregation. Structural characterization of Hvo PCNA at 2 Å resolution demonstrated the correct folding of the trimeric protein with a root mean square deviation of 1.3-1.7 Å compared with other archaeal and eukaryotic PCNA molecules (19, 20). Similarly, the Hvo RPA3 DNA-binding protein was shown to be competent to bind DNA in buffers containing 1 M KCl, but not at the lower concentration tested (0.2 M KCl) (19, 20). Both proteins were expressed in *E. coli* and were principally purified in buffers containing 1M NaCl, therefore this concentration was maintained in this study.

IMAC was used as a first step for the purification of His-tagged RadA and un-tagged RadB using a cobalt-based metal affinity resin. The cobalt-based resin was observed to bind His-tagged halophilic RadA protein with a higher specificity compared to nickel-based resins, reducing non-specific binding of host proteins (data not shown).

Co-eluted His-tagged RadA and un-tagged RadB proteins were observed at a similar ratio as before at their expected molecular mass positions by SDS-PAGE analysis (Fig. 2b), indicating specific interaction. Eluted fractions were subsequently subjected to SEC for further purification and to provide an initial indication of the stoichiometry of RadA-RadB interaction. Both proteins eluted together as a complex of ~370 kDa across the elution range of 120-170 mL (Fig. 3a). No difference

in protein ratio was apparent in any elution fraction suggesting that complex formation may be constrained by a particular stoichiometric configuration (Fig. 3b). Absence of contaminating DNA was confirmed by agarose gel electrophoresis.

Optimisation of the heterologous expression system

Whereas previous experiments demonstrated successful purification of a stable RadA-RadB complex, subsequent analysis was hampered by a limitation of the purification protocol to remove all non-specific host proteins. One major persistent contaminant (~64 kDa) that could not be removed by standard optimisation approaches without significant concomitant decrease in yield of the RadA-RadB complex was identified by MALDI analysis to be the product of the *arnA* gene (bi-functional polymyxin resistance protein, ArnA). The primary sequence of ArnA demonstrates an abundance of histidine residues (31) likely accounting for its capture by IMAC resins. Similarly, the ~74 kDa ArnA protein is reported to form a hexameric structure consisting of a dimer of trimers (31, 32) with a predicted molecular mass of ~440 kDa, close to the observed ~370 kDa RadA-B complex.

In view of this, bacteriophage-mediated P1 transduction was performed to delete the *arnA* gene from the Rosetta 2 (DE3) expression strain. The resulting Rosetta 2 (DE3) Δ*arnA* mutant strain was found to be equivalent to the parent strain in respect of growth characteristics and over-expression of RadA-RadB. Subsequently, RadA-RadB complex expressed from the modified strain was purified to ~95% homogeneity following the previous two-step purification protocol (data not shown).

RadA oligomerizes into ring like structures

Recombinase proteins characteristically oligomerise into ring structures or helical nucleoprotein filaments. Electron microscopy of *P. furiosus* RadA revealed that the protein forms dimers of heptameric ring structures in solution in the absence of DNA (8) whereas, in *S. solfataricus*, RadA forms an octameric ring bound to DNA (33). Since Hvo RadA-RadB is consistently purified as a large ~370 kDa stable complex, some form of RadA oligomerization is clearly indicated. Preliminary densitometric analysis of co-purified RadA and RadB protein bands estimates the ratio of RadA co-purified with RadB to be 4:1. Therefore the observed ~370 kDa RadA-RadB SEC complex would most closely correspond to eight molecules of RadA in complex with two molecules of RadB although this estimation is limited by the resolution of the S200 SEC column employed in this study.

Subsequent electron microscopy observation of our purified Hvo RadA-RadB protein complex confirms that Hvo RadA-RadB complex also possesses a tendency to self-associate into multimeric structures in the absence of DNA (Fig. 4). In order to distinguish protein structures from electron microscopy (EM) artifacts resulting as a consequence of high salt levels, several control grids were prepared and imaged with either buffer only or buffer plus stain at 100,000X magnification. The images captured showed ring-like structures throughout the field which suggests that Hvo RadA-RadB complex in its native state similarly exists as a ring-like structure (Fig. 4).

Discussion

The expression and purification of halophilic proteins, particularly from heterologous host overexpression systems presents several challenges, principally due to the particular physicochemical properties of these proteins which constrain conventional approaches.

Halophilic proteins are adapted to maintain solubility and function in hyper saline (1-4 M) conditions (34). The presence of high salt in halophilic organisms affects the gross molecular conformation of both proteins (35) and DNA (36). Mesophilic DNA-binding proteins typically make heavy use of positively charged residues to mediate electrostatic interactions with DNA (15, 29, 37). In the majority of organisms the presence of higher concentrations of salt interferes with DNA-binding and leads to aggregate formation in mesophilic proteins. The situation is reversed in halophilic counterparts. Halophilic proteins are generally decorated with negatively charged amino acid residues which allow the binding of surplus water and salt to build up a hydrated solvent network on the surface of proteins. Structural studies have shown that binding of hydrated cations (provided by excess salt and water molecules) around the negatively charged residues on the protein surface reduces the electrostatic repulsive forces between polyanionic DNA and protein molecules (19, 38). At lower concentrations of salt, the protective effect is lost and the repulsive forces between the acidic residues lead to unfolding and inactivation of the protein (28).

Consistently, *H. volcanii* RadA has an overall abundance of acidic amino acids, particularly aspartate which, based on homology modelling of RadA, contributes to the overall predicted negative surface charge of the protein. Accounting for this, we successfully adapted a protocol for purification of soluble RadA under moderate salt conditions. Further protocol modifications were introduced to enable purification from persistent contaminants including both DNA and particular host proteins. Most notably, we overcame the previously observed deleterious effect of RadA over-expression in *E. coli* by co-expression of RadA with RadB. These adaptations enabled RadA

purification as a DNA-free complex with RadB using a relatively facile two-step purification protocol. Based on our observations we speculate that RadA interacts with a small proportion of RadB in a particular stoichiometry to stabilize it as a binding partner.

H. volcanii RadA is similar to other characterised RadA homologues. Consistent with its function as a recombinase, ATP and DNA binding motifs are apparent which have a well-conserved sequence composition despite the common skews in overall amino acid usage typically observed in halophilic proteins (Fig. 1 and Fig. S1). The abundance of acidic amino acids such as aspartate in Hvo RadA could potentially result in electrostatic repulsion with the negatively charged DNA backbone under *in vitro* conditions (Fig. 1). The crystal structure analysis of Hvo PCNA indicates that this protein compensates for the reduction in positively charged surface residues by employing cation binding (19). However, the interaction in this instance is largely topological in nature rather than via interaction of specific residues. A conserved basic patch on the surface of RadB has been shown to be crucial in DNA binding (15) and number of similarly conserved positively charged residues are retained in Hvo RadA. The presence of a highly negatively charged loop uniquely present in Hvo RadA may compensate for the retention of lysine residues that presumably are maintained to enable direct DNA-binding (Fig. 1a).

The failure to observe DNA binding activity in Hvo RadA across a range of methods (refer to supplementary text, Fig. S3 and Table S1) suggests that, for this protein, the protein-DNA interaction may require additional, as yet unidentified, factors. It is also possible that under the conditions of our assays, RadA and RadB interaction occluded the DNA binding site or prevented an active conformational state of the proteins required for DNA binding and ATPase activity. It is also possible that additional co-factors are required for enzymatic activity of the complex.

A recently published study exploring the role of Hvo RadB in homologous recombination also confirmed interaction of RadA with RadB when the proteins were individually over-expressed in *H. volcanii*, identifying co-purifying proteins by mass spectrometry [39]. Genetic analyses identified point mutations in RadA that suppress the Δ*radB* phenotype, supporting the hypothesis that RadB functions as a recombination mediator, potentially by inducing conformational changes within RadA. Additionally, the authors noted that *in vitro* assays performed did not demonstrate functional DNA binding, as found in this study, also noting that additional co-factors may be required.

Recombinases have been generally observed to self-assemble into ring structures (six to eight protomers) in their native state and monomerize to interact with DNA to form the helical nucleoprotein filaments. So far the precise function of the ring structures is not clear in the process of homologous base pairing. The toroidal form in recombinases has been suggested to function as an inactive storage form of protein, which is likely utilized to occlude the polymerization motif, preventing unwanted interaction of the protein with DNA. EM analysis of the purified protein complex showed that RadA-RadB also multimerizes into ring-like structures, which we consider may be of the order of a heptamer or octamer (Fig. 4), similar to its RadA/Rad51 homologs. On the basis of our collective observations, further studies to understand the dynamics of halophilic protein-protein and protein-DNA interactions in *H. volcanii*, including gross structural determination and robust biophysical characterization can now be anticipated.

Funding

This work was supported by the University of Sindh, Jamshoro and the Higher Education Commission of Pakistan. K.B. and J.W. were supported by a Wellcome Trust RCDF award to K.B. (Grant no. 076556/Z/05/Z).

Conflicts of interest

None.

References

1. **Kowalczykowski SC, Eggleston AK.** Homologous pairing and DNA strand-exchange proteins. *Annu Rev Biochem.* 1994;63:991-1043.

2. **Story RM, Weber IT, Steitz TA.** The structure of the E. coli recA protein monomer and polymer. *Nature.* 1992;355(6358):318-25.

3. **Sandler SJ, Satin LH, Samra HS, Clark AJ.** recA-like genes from three archaean species with putative protein products similar to Rad51 and Dmc1 proteins of the yeast Saccharomyces cerevisiae. *Nucleic Acids Res.* 1996;24(11):2125-32.

4. **Ogawa T, Shinohara A, Nabetani A, Ikeya T, Yu X, et al**. RecA-like recombination proteins in eukaryotes: functions and structures of RAD51 genes. *Cold Spring Harb Symp Quant Biol*. 1993;58:567-76.

5. **Shinohara A, Ogawa H, Ogawa T**. Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. *Cell.* 1992;69(3):457-70.

6. **Komori K, Miyata T, Daiyasu H, Toh H, Shinagawa H, et al.** Domain analysis of an archaeal RadA protein for the strand exchange activity. *J Biol Chem.* 2000;275(43):33791-7.

7. **Ariza A, Richard DJ, White MF, Bond CS.** Conformational flexibility revealed by the crystal structure of a crenarchaeal RadA. *Nucleic Acids Res.* 2005;33(5):1465-73.

8. **Shin DS, Pellegrini L, Daniels DS, Yelent B, Craig L, et al.** Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *EMBO J.* 2003;22(17):4566-76.

9. **Sung P, Klein H.** Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol*. 2006;7(10):739-50.

10. **Symington LS.** Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol Mol Biol Rev.* 2002;66(4):630-70.

11. **Sandler SJ, Hugenholtz P, Schleper C, DeLong EF, Pace NR, et al.** Diversity of radA genes from cultured and uncultured archaea: comparative analysis of putative RadA proteins and their use as a phylogenetic marker. *J Bacteriol.* 1999;181(3):907-15.

12. **Komori K, Miyata T, DiRuggiero J, Holley-Shanks R, Hayashi I, et al.** Both RadA and RadB are involved in homologous recombination in Pyrococcus furiosus. *J Biol Chem.* 2000;275(43):33782-90.

13. **Allers T, Ngo HP.** Genetic analysis of homologous recombination in Archaea: Haloferax volcanii as a model organism. *Biochem Soc Trans.* 2003;31(Pt 3):706-10.

14. **Haldenby S, White MF, Allers T.** RecA family proteins in archaea: RadA and its cousins. *Biochem Soc Trans*. 2009;37(Pt 1):102-7.

15. **Guy CP, Haldenby S, Brindley A, Walsh DA, Briggs GS, et al.** Interactions of RadB, a DNA repair protein in archaea, with DNA and ATP. *J Mol Biol.* 2006;358(1):46-56.

16. **Allers T.** Overexpression and purification of halophilic proteins in Haloferax volcanii. *Bioeng Bugs.* 2010;1(4):288-90.

17. **Allers T, Barak S, Liddell S, Wardell K, Mevarech M.** Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in Haloferax volcanii. *Appl Environ Microbiol.* 2010;76(6):1759-69.

18. **Wardell K, Haldenby S, Jones N, Liddell S, Ngo GHP, et al.** RadB acts in homologous recombination in the archaeon Haloferax volcanii, consistent with a role as recombination mediator. *DNA Repair* (Amst). 2017;55:7-16.

19. **Winter JA, Christofi P, Morroll S, Bunting KA.** The crystal structure of Haloferax volcanii proliferating cell nuclear antigen reveals unique surface charge characteristics due to halophilic adaptation. *BMC Struct Biol.* 2009;9:55.

20. **Winter JA, Patoli B, Bunting KA.** DNA binding in high salt: analysing the salt dependence of replication protein A3 from the halophile Haloferax volcanii. *Archaea.* 2012;2012:719092.

21. **Sheng D, Li M, Jiao J, Ni J, Shen Y.** Co-expression with RadA and the characterization of stRad55B, a RadA paralog from the hyperthermophilic crenarchaea Sulfolobus tokodaii. *Sci China C Life Sci.* 2008;51(1):60-5.

22. **Sternberg N.** Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. *Proc Natl Acad Sci U S A.* 1990;87(1):103-7.

23. **Mullakhanbhai MF, Larsen H.** Halobacterium volcanii spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch Microbiol.* 1975;104(3):207-14.

24. **Ziegler J, Vogt T, Miersch O, Strack D.** Concentration of dilute protein solutions prior to sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Anal Biochem.* 1997;250(2):257-60.

25. **Janson G, Zhang C, Prado MG, Paiardini A**. PyMod 2.0: improvements in protein sequence-structure analysis and homology modeling within PyMOL. *Bioinformatics*. 2016.

26. **DeLano WL.** The PyMol Molecular Graphics System. *DeLano Scientific LLC, Palo Alto, Calif, USA.* 2008.

27. **Webb B, Sali A.** Comparative Protein Structure Modeling Using MODELLER. Curr Protoc *Bioinformatics.* 2014;47:5 6 1-32.

28. **Lanyi JK.** Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriol Rev.* 1974;38(3):272-90.

29. **Chen LT, Ko TP, Chang YW, Lin KA, Wang AH, et al.** 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(9):e858.

30. **Poidevin L, MacNeill SA.** Biochemical characterisation of LigN, an NAD+-dependent DNA ligase from the halophilic euryarchaeon Haloferax volcanii that displays maximal in vitro activity at high salt concentrations. *BMC Mol Biol.* 2006;7:44.

31. **Gatzeva-Topalova PZ, May AP, Sousa MC.** Crystal structure and mechanism of the Escherichia coli ArnA (PmrI) transformylase domain. An enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance. *Biochemistry.* 2005;44(14):5328-38.

32. **Breazeale SD, Ribeiro AA, McClerren AL, Raetz CR.** A formyltransferase required for polymyxin resistance in Escherichia coli and the modification of lipid A with 4-Amino-4-deoxy-L-arabinose. Identification and function oF UDP-4-deoxy-4-formamido-L-arabinose. *J Biol Chem.* 2005;280(14):14154-67.

33. **Yang S, Yu X, Seitz EM, Kowalczykowski SC, Egelman EH.** Archaeal RadA protein binds DNA as both helical filaments and octameric rings. *J Mol Biol.* 2001;314(5):1077-85.

34. D**ym O, Mevarech M, Sussman JL.** Structural features that stabilize halophilic malate dehydrogenase from an archaebacterium. *Science.* 1995;267(5202):1344-6.

35. **Madern D, Ebel C, Zaccai G.** Halophilic adaptation of enzymes. *Extremophiles.* 2000;4(2):91-8.

36. **Schlick T, Li B, Olson WK**. The influence of salt on the structure and energetics of supercoiled DNA. *Biophys J.* 1994;67(6):2146-66.

37. **Lee CD, Wang TF.** The N-terminal domain of Escherichia coli RecA have multiple functions in promoting homologous recombination. *J Biomed Sci.* 2009;16:37.

38. **Bergqvist S, Williams MA, O'Brien R, Ladbury JE.** Halophilic adaptation of protein-DNA interactions. *Biochem Soc Trans*. 2003;31(Pt 3):677-80.

39. **Hartman AL, Norais C, Badger JH, Delmas S, Haldenby S, et al.** The complete genome sequence of Haloferax volcanii DS2, a model archaeon. *PLoS One*. 2010;5(3):e9605.

Table

Table 1. List of strains and plasmids.

Strains and plasmids	Genotype/Description	Source
DH5α	*F-ϕ80lacZ ΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK-, mK+) phoA supE44 λ– thi-1 gyrA96 relA1*	Novagen
Rosetta 2 (DE3)	*F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE23 (CamR)*	Novagen
Rosetta 2 (DE3)Δ*arnA*	Rosetta 2 (DE3) with *arnA* inactivating mutation	This study
DS2	*H. volcanii* wild-type strain	(39) & (23)
pZero Blunt	Cloning vector	Invitrogen
pCPG42	Hvo *radB* encoding plasmid	(1515)
pET11-*radA*	*radA* cloned into NdeI/BamHI sites for overexpression of His-	This study
pETDUET-*radA*	*radA* cloned into NdeI/EcoRV sites of MCS2 for overexpression of	This study
pBPRAD2	*radB* cloned into NcoI/HindIII sites of pETDUET-*radA* MCS1 for co-expression of His-tagged RadA and untagged RadB	This study

Figure legends

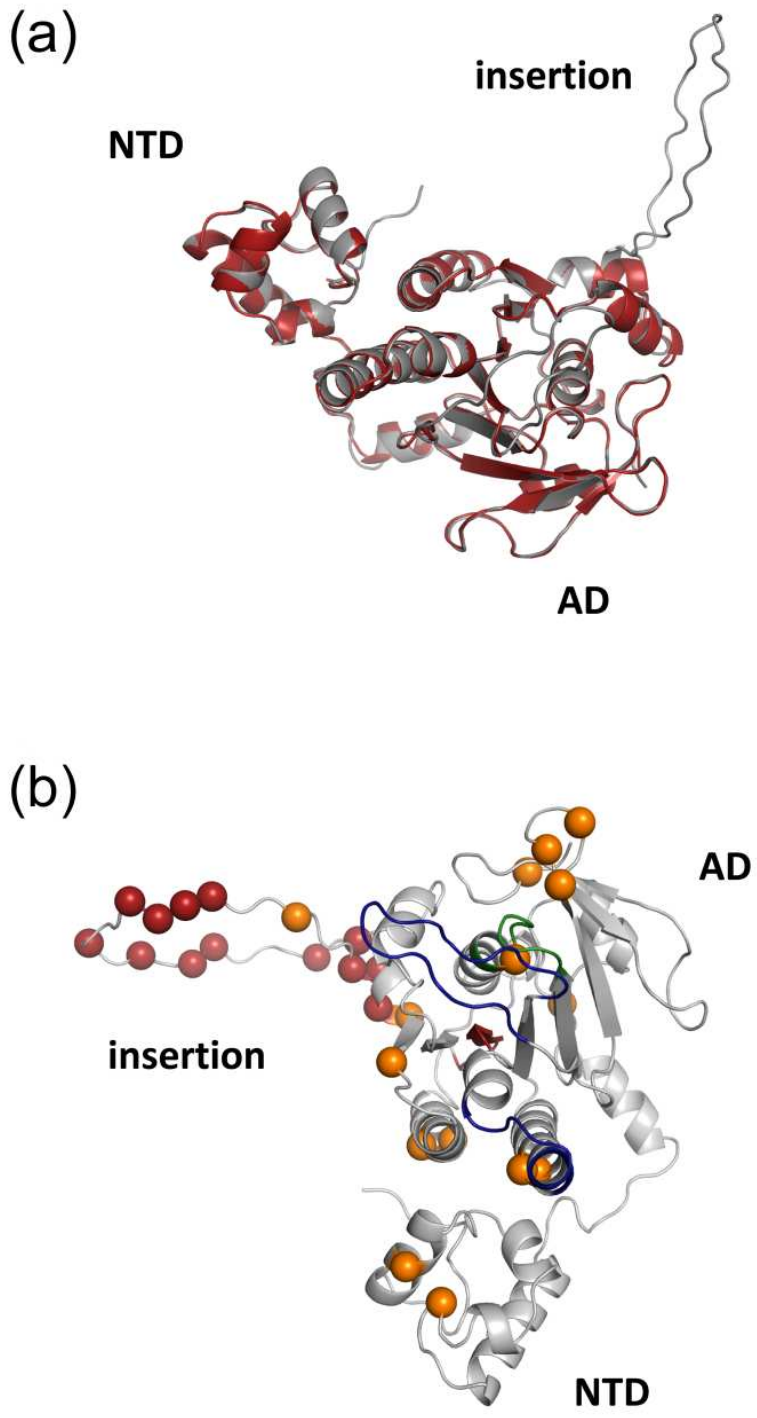
Fig. 1(a). Superposition of the Hvo RadA homology model with the Pfu RadA crystal structure. The Hvo RadA homology model is shown in grey and the Pfu RadA structure (1PZN:Chain A) in red (cartoon representation). Good correspondence is apparent between major structural elements, although Hvo RadA can be distinguished by an additional large loop insertion. (**b). Cartoon representation of the backbone of the modelled Hvo RadA structure.** The Walker A motif (green), Walker B motif (red) and DNA binding loops (blue) are indicated. The Cα position of the lysine residues in Hvo RadA are indicated with orange spheres. The Cα position of acidic residues located in an insertion loop in Hvo RadA relative to related sequences (see supplementary Fig. S1) are indicated with red spheres. NTD, AD and the insertion loop are labelled.

Fig. 2(a). Coomassie Brilliant-Blue stained SDS-PAGE gel demonstrating co-over-expression of Hvo RadA and RadB in *E. coli.* Broadly equivalent levels of recombinant RadA and RadB were observed in whole cell lysates when expressed independently (lanes 3 & 4, respectively) compared with reduced levels of RadB observed when co-expressed with RadA (lane 2, arrow). Lane 1 shows the pre-induction sample. RadA and RadB positions are indicated in (b). Molecular mass markers are indicated (kDa). **(b) Purification of co-expressed RadA-RadB proteins by cobalt-based immobilized metal affinity chromatography.** Untagged RadB is observed to co-purify with His-tagged RadA.

Fig. 3(a). Purification of RadA-RadB complex by Size Exclusion Chromatography. Chromatogram monitoring UV absorbance at 280 nm indicating the elution of the RadA-RadB complex at 130-170mL volume (arrow). Fractions as collected are indicated on the *x*-axis. The relative elution peaks for a series of molecular mass standards are shown with arrows. **(b). Coomassie Brilliant-Blue stained SDS-PAGE gel representing the purified proteins after SEC.** Lane 1 showing 10 µl Talon purified load prior to purification on a S200 column. All S200 purified fractions corresponding to the UV peak at 100-180 mL (lanes 2-9) show RadA and RadB co-purification at their estimated MW markers. An ~64kDa contaminating host protein (ArnA) consistently co-eluted with the RadA-B complex as indicated in lanes 5-8.

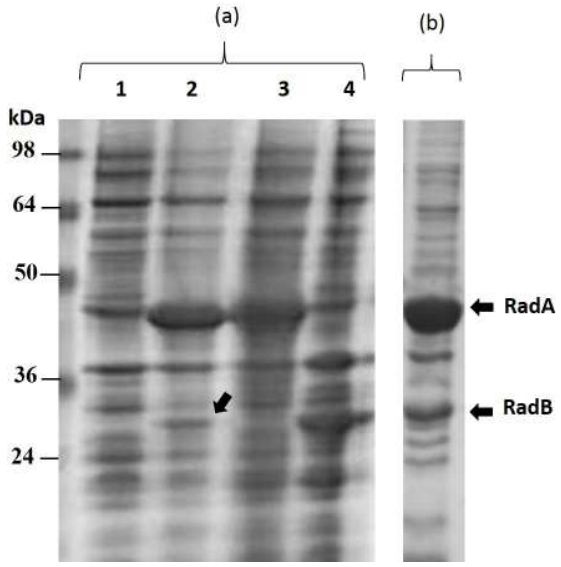
Fig. 4. Transmission Electron Microscopy (TEM) images of co-expressed RadA-B complex. TEM images of co-expressed RadA-B complex at 100,000X magnification. All of the protein samples were negatively stained with either (a) 1% or (b) 0.5% phosphotungstic acid (PTA). Consistently sized (~13nm diameter) ring-like RadA-RadB complexes are indicated with arrows. Boxed ring-like structures are shown enlarged in inset images.

630 **Figure 1**



631

632 **Figure 2**



633

634

635

636

637

638

639

640

641

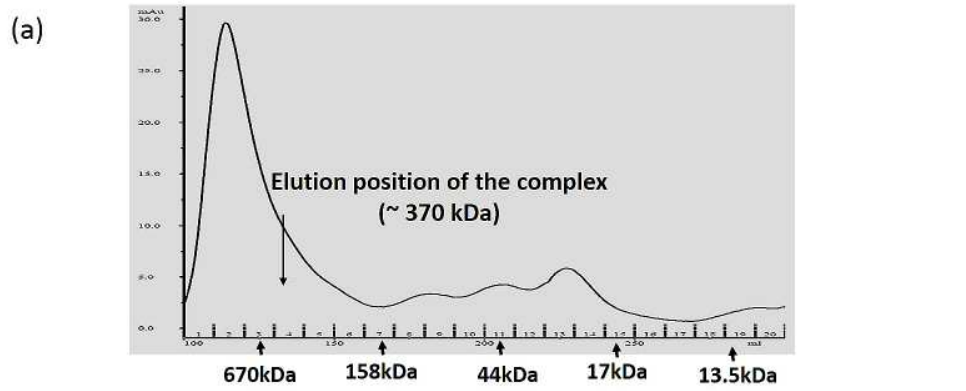
642

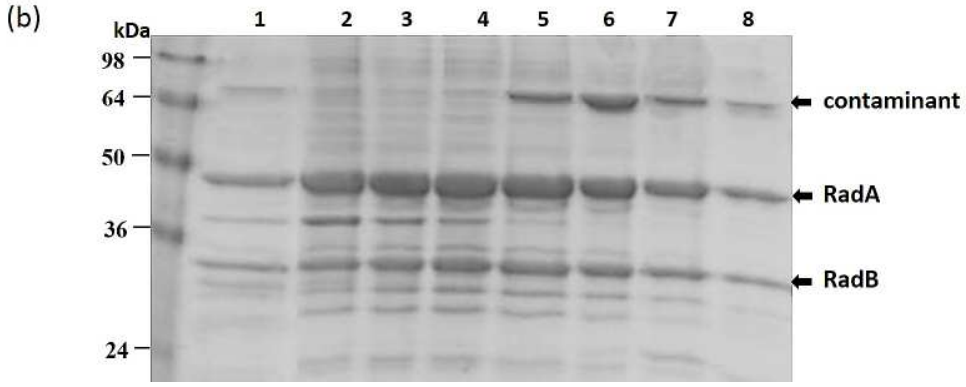
643

644

645

646 **Figure 3**





647

648

649

650

651

652

653

654

655

656

Figure 4