*Published online 5 October 2011* *Nucleic Acids Research, 2012, Vol. 40, No. 3* ***1077–1090***
*doi:10.1093/nar/gkr838*

# Identification of essential and non-essential single-stranded DNA-binding proteins in a model archaeal organism

**Agnieszka Skowyra and Stuart A. MacNeill***

School of Biology, University of St Andrews, North Haugh, St Andrews, Fife KY16 9TF, UK

Received August 29, 2011; Revised and Accepted September 20, 2011

## ABSTRACT

**Single-stranded DNA-binding proteins (SSBs) play vital roles in all aspects of DNA metabolism in all three domains of life and are characterized by the presence of one or more OB fold ssDNA-binding domains. Here, using the genetically tractable euryarchaeon *Haloferax volcanii* as a model, we present the first genetic analysis of SSB function in the archaea. We show that genes encoding the OB fold and zinc finger-containing RpaA1 and RpaB1 proteins are individually non-essential for cell viability but share an essential function, whereas the gene encoding the triple OB fold RpaC protein is essential. Loss of RpaC function can however be rescued by elevated expression of RpaB, indicative of functional overlap between the two classes of haloarchaeal SSB. Deletion analysis is used to demonstrate important roles for individual OB folds in RpaC and to show that conserved N- and C-terminal domains are required for efficient repair of DNA damage. Consistent with a role for RpaC in DNA repair, elevated expression of this protein leads to enhanced resistance to DNA damage. Taken together, our results offer important insights into archaeal SSB function and establish the haloarchaea as a valuable model for further studies.**

## INTRODUCTION

Single-stranded DNA-binding proteins (SSBs) are indispensable for many aspects of DNA metabolism including replication, repair and recombination, and play a vital role in the maintenance of genomic stability in all three domains of life (1,2). SSBs are characterized by the presence of one or more OB (oligosaccharide–oligonucleotide binding) fold domains. OB folds consist of a five-stranded β-sheet that is coiled to form a closed β-barrel, often capped by an α-helix (2,3). They range in length from 75–150 residues with much of the length variation being due to the presence of variable loop regions located between structurally conserved β-strands, and display only a low level of primary sequence similarity (2). In addition to SSBs, where ssDNA binding is not sequence-specific, OB folds are also found in other contexts, for example in proteins that bind ssDNA in a sequence-specific manner such as the bacterial transcriptional terminator Rho. A number of structures of OB fold–DNA complexes have been solved, revealing the molecular details of the protein–ssDNA interaction. Particularly important are stacking interactions between the side chains of conserved aromatic amino acids and unpaired nucleotides (4).

SSBs display a wide variety of domain and subunit organizations across evolution (2). In bacteria, almost all SSB proteins comprise a single N-terminal OB fold and a relatively short, flexible C-terminal tail that mediates protein–protein interactions. Individual bacterial SSB proteins assemble into homotetrameric complexes. Eukaryotic mitochondrial SSBs closely resemble bacterial SSBs, indicative of their likely evolutionary origin. In some bacterial lineages, SSB proteins with two OB folds are found: these proteins form homodimers and thus have four OB folds per complex, similar to the homotetrameric SSB complexes described above (2).

In eukaryotes, the major cellular SSB is replication protein A (RPA), a conserved heterotrimeric complex comprising the RPA70, RPA32 and RPA14 proteins (5). RPA70 contains four OB-fold motifs, and RPA32 and RPA14 one each, although only four of these six OB folds actually bind ssDNA, the remaining two being involved in mediating protein–protein interactions. In addition to the OB folds, RPA32 also possesses an extended N-terminal domain and a C-terminal winged helix–turn–helix (wHTH) domain, both of which are involved in protein–protein interactions. The N-terminal domain is also a target for regulatory phosphorylation (1). In addition to the canonical RPA, alternative RPA

*To whom correspondence should be addressed. Tel: +44 1334 467268; Fax: +44 1334 462595; Email: stuart.macneill@st-andrews.ac.uk

(aRPA) complexes are also found in various eukaryotic lineages, including mammals (6,7) and plants (8) and the structurally similar CST (Cdc13–Stn1–Ten1) complex plays a key role in telomere maintenance (9). Two additional metazoan SSBs that function in the cellular response to DNA damage have also been identified; these proteins (hSSB1 and hSSB2) contain a single OB fold only (10–13).

In the third domain of life, the archaea, SSBs displaying a wide variety of architectures have been identified and characterized from species representative of several major lineages. The methanogenic euryarchaeal species *Methanosarcina acetivorans* encodes three RPA proteins termed MacRPA1–MacRPA3, each of which appears capable of self-assembling into a homomultimeric complex (14–18). MacRPA1 contains four tandem OB folds while MacRPA2 and MacRPA3 consist of two OB folds and a $CX_2CX_8CX_2H$ zinc finger motif (Figure 1). Mutation of individual conserved cysteine or histidine residues in the MacRPA3 zinc finger reduces ssDNA binding (16). The combination of OB folds and zinc fingers is seen in other methanogenic RPA proteins including *Methanopyrus kandleri* RPA (which resembles MacRPA2 and MacRPA3, with two OB folds and a zinc finger) (17), *Methanocaldococcus jannaschii* RPA (four OB folds and a zinc finger) (19) and *Methanothermobacter thermautotrophicus* RPA (five OB folds and a zinc finger), in RPA proteins from non-methanogenic species such as *Ferroplasma acidarmanus* and *Thermoplasma volcanium* (two OB folds and a zinc finger) (20) and *Pyrococcus furiosus* (one OB and a zinc finger), as well as in the eukaryotic RPA70 protein (four OB folds and a zinc finger) (5). In addition to MacRPA1, RPA proteins lacking zinc finger motifs are also seen in a broad range of archaeal organisms. In contrast to the complexity seen with the euryarchaeal SSB proteins, the best characterized crenarchaeal SSB, from the thermophile *Sulfolobus solfataricus*, contains only a single OB fold followed by a flexible C-terminal tail that is not required for ssDNA binding (21). Similar structures are predicted for other crenarchaeal SSB proteins.

Acquiring a detailed understanding of the function of the archaeal RPA proteins will require a combination of *in vivo* genetic and *in vitro* biochemical and biophysical approaches. To date, however, only a handful of archaeal species have proved tractable to molecular genetic analysis (22). Among these, the haloarchaeon *Haloferax volcanii* has proved particularly useful. This organism is easy to culture in the lab, growing aerobically with an optimal growth temperature of 42–45°C in medium containing 18% NaCl and a generation time of 2–3 h. A range of molecular tools and methods have been developed to facilitate genetic analysis in *Hfx. volcanii*, including selectable markers, shuttle vectors, transformation methods, a regulatable promoter, etc. (22). The complete sequence of the *Hfx. volcanii* genome has been determined and is publically available (23). *Hfx. volcanii* also has a natural system of genetic exchange (24,25) that can offer an additional route for strain construction (26).

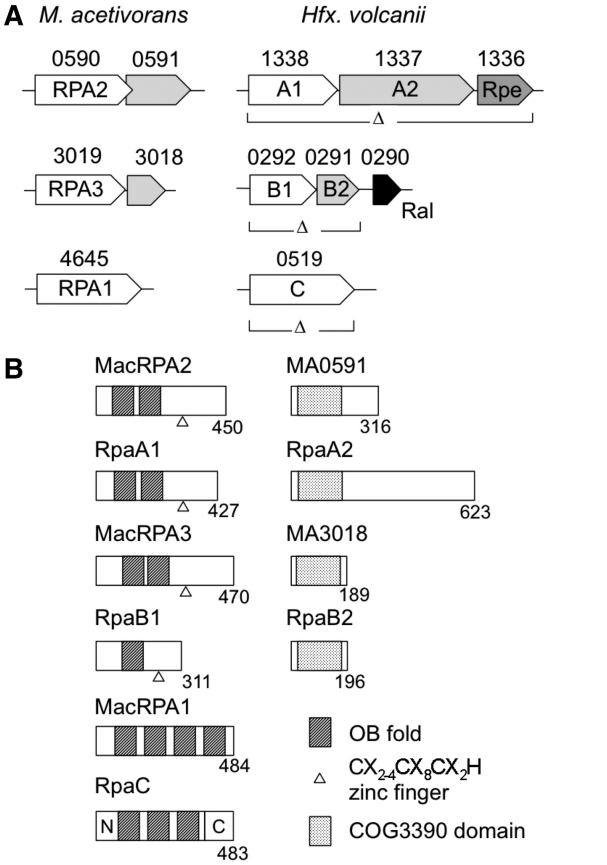


**Figure 1.** SSB proteins. (**A**) Gene organization of RPA proteins in *M. acetivorans* (left) and *Hfx. volcanii* (right). The arrows represent individual open reading frames with systematic gene designations shown above (for clarity the MA and HVO prefixes are omitted). The shading groups ORFs encoding OB fold-containing proteins (white), putative OB fold-containing COG3390 proteins (light grey), the phosphoesterase Rpe (dark grey) and the haloarchaeal-specific Ral protein (black). Regions deleted from the chromosome in this study are shown by the Δ symbol. (**B**) Schematic of SSB protein structures showing OB folds (hatched boxes), COG3390 domains (speckled boxes) and the zinc finger motifs (triangle). Protein lengths are indicated.

In this paper, we present the first genetic analysis of RPA function in any archaeal organism using *Hfx. volcanii* as a model system. We identify *Hfx. volcanii* genes encoding homologues of the biochemically characterized *M. acetivorans* RPA proteins MacRPA1–MacRPA3 and show that while the related *Hfx. volcanii* RpaA1 and RpaB1 proteins (homologues of *M. acetivorans* MacRPA2 and MacRPA3) are individually non-essential for cell viability, they appear to share an essential function. In contrast, *Hfx. volcanii* RpaC (homologue of *M. acetivorans* MacRPA1) is essential for cell viability on its own; however, the effects of reduced RpaC expression can be overcome by increased expression of RpaB. In addition, we demonstrate the importance for the OB folds of RpaC for its function and roles for the non-essential N- and C-terminal domains of RpaC in surviving DNA damage. Consistent with a role for RpaC in DNA repair, we show that elevated expression of RpaC leads to enhanced resistance to DNA damage.

Our results represent an important first step towards obtaining a full understanding of the diverse roles performed by the multiple SSB proteins found in haloarchaeal cells.

## MATERIALS AND METHODS

### Strains and media

*Hfx. volcanii* DS70 derivatives H26 (Δ*pyrE2*), H53 (Δ*pyrE2* Δ*trpA*), H98 (Δ*pyrE2* Δ*hdrB*) and H99 (Δ*pyrE2* Δ*trpA* Δ*hdrB*) have been described (27) and were a generous gift of Dr Thorsten Allers (University of Nottingham, Nottingham, UK). All the strains used in this study are listed in Table 1. *Hgm. borinquense* DSM11551 and *Hqr. walsbyi* DSM16790 were obtained from the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany). For routine cloning purposes, *Escherichia coli* DH5α (*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*) was used (Invitrogen). To prepare unmethylated plasmid DNA for *Hfx. volcanii* transformation, *E. coli* SCS110 (*rpsL* ($Str^r$) *thr leu endA thi-1 lacY galK galT ara tonA tsx dam dcm supE44* Δ(*lac-proAB*) [F′ *traD36 proAB lacI*$^q$*ZΔM15*]) was used (Stratagene).

*Hfx. volcanii* cells were routinely grown on Hv-YPC, Hv-Ca or Hv-Min media prepared as described in the Halohandbook v7.1 (www.haloarchaea.com/resources/halohandbook). For selection procedures, tryptophan and/or thymidine and hypoxanthine were added to Hv-Ca and Hv-Min medium at a final concentration of 50 μg/ml for tryptophan and thymidine and 40 μg/ml for hypoxanthine. For counter-selection using 5-fluoroorotic acid (5-FOA), Hv-Ca or Hv-Min was supplemented with uracil and 5-FOA at final concentrations of 10 μg/ml and 50 μg/ml, respectively. *Hgm. borinquense* DSM11551 and *Hqr. walsbyi* DSM16790 were grown in medium recommended by the DSMZ.

*E. coli* were grown in LB medium (Formedium) supplemented when necessary with ampicillin at 100 μg/ml.

### Plasmids for gene deletions

Plasmids for gene deletion were constructed by polymerase chain reaction (PCR) amplifying regions of ~500 bp from 5′ and 3′ to the region to be deleted using *Hfx. volcanii* DS70 genomic DNA as template. Details of the oligonucleotides used (P1–P12) can be found in Supplementary Table S1. The PCR products were then restricted with EcoRI and BamHI (5′ flanking region for *rpaA* and *rpaB* operons, 3′ flanking region for *rpaC* ORF) or BamHI and SpeI (5′ flanking region for *rpaC*, 3′ flanking region for *rpaA* and *rpaB*) and ligated together into EcoRI- and SpeI-digested plasmid pTA131 (27). The resulting plasmids (PL5, PL7 and PL9, Supplementary Table S2) were sequenced to ensure the absence of unwanted sequence changes. Next, BamHI or BamHI–BglI restriction fragments carrying the $trpA^+$ or $hdrB^+$ selectable marker genes from plasmids pTA230 (27) and pBBHrdB, respectively, were then cloned into the unique BamHI site located in the centre of the 5′ and 3′ flanking regions. Plasmid pBBHrdB is a derivative of pTA187 (27) in which a BglI site has been introduced downstream of the $hdrB^+$ gene to ease subsequent sub-cloning (unpublished data). The resulting plasmids (PL6, PL8 and PL10, Supplementary Table S2) were passaged through *E. coli* SCS110 to demethylate the DNA prior to transformation into *Hfx. volcanii*.

### Plasmids for regulated expression

Plasmid pNPM-tna was constructed by amplifying the 124 bp *tnaA* promoter from *Hfx. volcanii* DS70 genomic DNA using oligonucleotides P13 and P14, digesting the PCR product with BclI and NotI, and ligating into BclI- and NotI-digested pTA409 (28). The resulting plasmid (pNPM-tna, PL11, Supplementary Table S2) contains a multiple cloning site downstream of the *tna* promoter with unique NdeI, PacI and NotI sites along with additional unique sites that were present in pTA409; the plasmid was sequenced to ensure the absence of

**Table 1.** Strains used in this study

| Strain no. | Genotype | Construction method or reference |
|---|---|---|
| DS70 | | (37) |
| H98 | *ΔhdrB ΔpyrE2* | (27) |
| H99 | *ΔtrpA ΔhdrB ΔpyrE2* | (27) |
| SMH787 | *rpaA::trpA*$^+$ *ΔtrpA ΔhdrB ΔpyrE2* | H99, PL6 pop-in and pop-out |
| SMH788 | *rpaB::hdrB*$^+$ *ΔtrpA ΔhdrB ΔpyrE2* | H99, PL8 pop-in and pop-out |
| SMH728 | *rpaA*::[pNPM-tna-HfxRpaA] *ΔhdrB ΔpyrE2* | H98, PL12 pop-in |
| SMH729 | *rpaB*::[pNPM-tna-HfxRpaB] *ΔhdrB ΔpyrE2* | H98, PL13 pop-in |
| SMH730 | *rpaC*::[pNPM-tna-HfxRpaC] *ΔhdrB ΔpyrE2* | H98, PL14 pop-in |
| SMH738 | p*tna–rpaC ΔhdrB ΔpyrE2* | H98, PL15 pop-in and pop-out |
| SMH763 | *rpaA*::[pNPM-fdx-HfxRpaA] *ΔhdrB ΔpyrE2* | H98, PL16 pop-in |
| SMH764 | *rpaB*::[pNPM-fdx-HfxRpaB] *ΔhdrB ΔpyrE2* | H98, PL17 pop-in |
| SMH765 | *rpaC*::[pNPM-fdx-HfxRpaC] *ΔhdrB ΔpyrE2* | H98, PL18 pop-in |
| SMH789 | *rpaA*::[pNPM-fdx-HfxRpaA] *ΔhdrB ΔpyrE2* p*tna–rpaC* | SMH738, PL16 pop-in |
| SMH790 | *rpaB*::[pNPM-fdx-HfxRpaB] *ΔhdrB ΔpyrE2* p*tna–rpaC* | SMH738, PL17 pop-in |
| SMH791 | *rpaB*::[pNPM-tna-HfxRpaB] *rpaA::trpA*$^+$ *ΔtrpA ΔhdrB ΔpyrE2* | SMH787, PL13 pop-in |

See Supplementary Table S2 for plasmid details.
Strains carrying non-integrated plasmids are not shown.

sequence changes in the BclI–NotI region. Next, oligonucleotides P15–P20 were used to amplify the first 300 bp of the *rpaA1*, *rpaB1* and *rpaC* ORFs from *Hfx. volcanii* DS70 genomic DNA. The PCR products were restricted with NdeI and EcoRV, cloned separately into NdeI- and EcoRV-digested pNPM-tna and sequenced. The resulting plasmids (pNPM-tna-HfxRpaA, etc., PL12-PL14, Supplementary Table S2) were subsequently passaged through *E. coli* SCS110 prior to *Hfx. volcanii* transformation.

Plasmid pNPM-tna-HfxRpaC-int (PL15, Supplementary Table S2) was constructed from pNPM-tna–rpaC by insertion of a 500-bp BclI fragment carrying the upstream flanking region of *rpaC* amplified by PCR using oligonucleotides P21 and P22 and DS70 DNA as template into the unique BclI site upstream of the *tna* promoter. The plasmid was sequenced and passaged through *E. coli* SCS110 prior to transformation into *Hfx. volcanii*.

### Plasmids for constitutive high-level expression

To facilitate constitutive high-level RPA expression, the BclI–NdeI fragment carrying the *tna* promoter in the pNPM-tna-HfxRpa plasmids PL12–PL14 (Supplementary Table S2) was replaced with a 185-bp fragment carrying the *Hbt. salinarum fdx* promoter, amplified from genomic DNA using oligonucleotide primers P23 and P24 (Supplementary Table S1) and digested with BamHI and NdeI. The resulting plasmids (pNPM-fdx-HfxRpaA, etc., PL16–18, Supplementary Table S2) were sequenced and passaged through SCS110 prior to transformation into *Hfx. volcanii*.

### Plasmids for expression of full-length RpaC proteins

Plasmids containing the full-length *Hfx. volcanii*, *Hgm. borinquense* and *Hqr. walsbyi rpaC* genes under the control of the *fdx* promoter were constructed by amplifying the corresponding ORFs by PCR using the appropriate genomic DNA as template and oligonucleotides P25–P30 (Supplementary Table S1). The PCR products were then digested with NdeI and HindIII and ligated together with the BamHI–NdeI 185-bp fragment carrying the *Hbt. salinarum fdx* promoter, described above, into plasmid pTA230 (27) that had been digested with BamHI and HindIII. The resulting plasmids (pTA230-pfdx-HfxRpaC, etc., PL19–PL21, Supplementary Table S2) were then sequenced to ensure the absence of unwanted sequence changes in the BamHI–HindIII region and passaged through SCS110 prior to transformation into *Hfx. volcanii*.

### Plasmids for expression of mutated RpaC proteins

Plasmids expressing eight of nine mutant forms of the *Hgm. borinquense* RpaC protein were constructed by overlap extension PCR mutagenesis using plasmid pTA230-pfdx-HgmRpaC as template and oligonucleotides P27 and P28 in combination with oligonucleotides P31–P42 (Supplementary Table S1). The mutated ORFs were re-cloned (as NdeI–HindIII fragments) into pTA230-pfdx-HfxRpaC, from which the NdeI–HindIII fragment carrying the full-length *Hfx. volcanii rpaC* gene had been excised, and the resulting plasmids (PL22–PL30, Supplementary Table S2) sequenced to ensure the absence of unwanted sequence changes. The mutated ORFs expressed the following proteins: Hgm RpaC-ΔNTD (deletion of amino acids 1–62 inclusive), $\Delta OB^{A}$ (deletion of amino acids 58–167), $\Delta OB^{B}$ (amino acids 168–267), $\Delta OB^{C}$ (amino acids 268–367), $\Delta OB^{AB}$ (amino acids 58–267), $\Delta OB^{AC}$ (amino acids 58–167 and 268–367), $\Delta OB^{BC}$ (amino acids 168–367) and $\Delta OB^{ABC}$ (amino acids 58–367).

The ORF to express Hgm RpaC-ΔNTD was amplified using plasmid pTA230-pfdx-Hgm RpaC as template and oligonucleotides P43 and P28. The PCR product was then digested with NdeI and HindIII and cloned as above (PL22, Supplementary Table S2). The truncated ORF is deleted for sequences encoding amino acids 1–62 (incl.) of the *Hgm. borinquense* RpaC protein.

The ORF to express Hgm RpaC-ΔCTD was constructed by digesting plasmid pTA230-fdx-HgmRpaC with EcoRI and ligating in a short dsDNA with EcoRI cohesive ends assembled from oligonucleotides P44 and P45 (Supplementary Table S1) containing a stop codon. The EcoRI site is located in the gene at a position that corresponds to the poorly conserved linker between the third OB fold and the CTD. Correct insertion of stop codon-containing DNA was confirmed by sequencing (PL30, Supplementary Table S2). The truncated ORF encodes amino acids 1–404 of Hgm RpaC only.

### Plasmids for expression of Flag epitope-tagged RpaC proteins

Plasmids expressing Flag-tagged full-length and mutated forms of Hgm RpaC (PL31–39, Supplementary Table S2) were constructed by digesting the appropriate pTA230-fdx-HgmRpaC plasmids with NdeI and ligating in a short dsDNA with NdeI cohesive ends assembled from oligonucleotides P46 and P47 (see Supplementary Table S1). The resulting proteins are expressed with the N-terminal sequence $M_1$<u>DYKDDDDK</u>H$M_2$ where $M_2$ corresponds to the native N-terminal methionine (Met63 in the case of the RpaC–ΔNTD protein) and the Flag tag is underlined.

### Gene deletions

Unmethylated plasmids for gene deletion (PL6, PL8 and PL10, Supplementary Table S2) were transformed into *Hfx. volcanii* strain H99 (*ΔpyrE2 ΔtrpA ΔhdrB*) and transformants obtained on Hv-Ca medium with appropriate supplements at 45°C. For each gene deletion, three integrant colonies were picked, resuspended in Hv-Ca medium and plated on Hv-Ca or Hv-Min plates (see 'Results' section), supplemented with both uracil and 5-FOA, with or without tryptophan (for selection of deletions marked with $trpA^{+}$) or with or without thymidine and hypoxanthine (for selection of deletions marked with $hdrB^{+}$). Candidate gene deletion colonies (typically 16 colonies for each) were screened by PCR using oligonucleotides P48–P57 indicated in Supplementary Table S1. Deletion strains were then purified by streaking to

single colonies on Hv-YPC medium, re-tested by PCR and stored at −80°C in 20% glycerol.

### Mating assay

Mating of *Hfx. volcanii* strains was performed essentially as described previously (26). Briefly, strains to be mated (see below) were grown to mid-exponential phase ($OD_{650\ nm}$ of 0.4) in Hv-YPC medium before being combined and filtered onto a 0.45-μm filter. The filter was then placed face-up on an Hv-YPC plate. After overnight incubation at 45°C, the cells were washed from the filter using 1 ml of sterile 18% SW and 100-μl aliquots plated onto Hv-Ca plates with necessary supplements. Colonies formed after 5 days incubation at 45°C were then analysed by PCR.

### Plasmid integration

For expression from the *tna* promoter, plasmids PL12–14 (Supplementary Table S2) were transformed into *Hfx. volcanii* H98 and integrative transformants obtained on Hv-Ca medium without additional supplements. Note that Hv-Ca contains a low level of tryptophan sufficient to activate the *tna* promoter (29). Plasmid integration at the correct chromosomal loci by homologous recombination was confirmed by PCR using oligonucleotide primers P48–P53 (Supplementary Table S1). Integrant strains (together with strain DS70 as a control) were then grown in liquid Hv-Min medium containing 0.075 mM tryptophan to mid-exponential phase ($OD_{650}$ 0.5). After pelleting at 2000 rpm for 8 min and washing with Hv-Min medium, the cultures were adjusted to an $OD_{650}$ of 0.5, serially diluted in Hv-Min from $10^{-1}$ to $10^{-5}$ and spotted onto Hv-Min plates with or without 0.075 mM tryptophan. The plates were incubated at 45°C until colonies were formed.

For expression from the *fdx* promoter, plasmids PL16–PL18 (Supplementary Table S2) were transformed into *Hfx. volcanii* H98 and integrative transformants obtained on Hv-Ca medium without additional supplements. Correct integration was confirmed as above.

### Creation of a stable p*tna–rpaC* strain

Plasmid PL15 (Supplementary Table S2) was transformed into *Hfx. volcanii* H98 and integrative transformants obtained on Hv-Min medium supplemented with 0.075 mM tryptophan at 45°C. Colonies were picked, re-suspended in 18% SW and plated on Hv-Min plates containing 0.075 mM tryptophan, uracil, 5-FOA, thymidine and hypoxanthine, and incubated at 45°C until colonies formed. Colonies were then screened by PCR using oligonucleotide primers P58 and P59 (Supplementary Table S1) to identify p*tna–rpaC* strains. Subsequently, p*tna–rpaC* strains were purified by streaking to single colonies on Hv-Min medium containing 0.075 mM tryptophan, thymidine and hypoxanthine, re-tested by PCR and stored at −80°C in 20% glycerol.

### Full-length and truncated RpaC proteins

Plasmids P19–P30 (Supplementary Table S2) were transformed into *Hfx. volcanii* p*tna–rpaC* and transformants obtained on Hv-Min medium supplemented with 0.075 mM tryptophan, thymidine and hypoxanthine at 45°C. Transformant strains were then cultured on HvMin with tryptophan to mid-exponential phase, washed, serially diluted, spotted on HvMin plates with or without tryptophan and incubated at 45°C for 3 days.

### Detection of Flag-tagged RpaC proteins

Plasmids PL31–PL40 (Supplementary Table S2) were transformed into *Hfx. volcanii* H98 and transformants obtained on Hv-Ca medium (supplemented with thymidine and hypoxanthine) at 45°C. Proteins extracts were prepared from cell cultures grown to mid-log phase in Hv-Min medium (supplemented with thymidine and hypoxanthine), subjected to sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE), transferred to PDVF membrane and immunoblotted using anti-Flag monoclonal antibodies (Sigma). Following washing and incubation with HRP-linked sheep anti-mouse secondary antibodies (GE Healthcase), signals were visualized using an LAS 3000 imaging system (Fujifilm).

### DNA damage sensitivity assays

Cells were grown in Hv-Min medium supplemented with 0.075 mM tryptophan to an $OD_{650nm}$ of 0.3. For UV sensitivity assay, cells were 10-fold serial diluted, plated onto Hv-Min with tryptophan, irradiated with doses from 50 to 200 $J/m^2$ with UV-C (254 nm) using a Stratalinker UV1800 (Agilent) and incubated in dark at 45°C. For 4NQO, MMS and phleomycin sensitivity assays, cells were divided into several aliquots of 1 ml and incubated with 20 μl of appropriated concentration of drug for 1 h with shaking at 45°C. Ten-fold serial dilutions were plated on Hv-Min with tryptophan. Plates were incubated at 45°C until colonies formed. Final concentrations for 4NQO were 0.2, 0.4, 0.6 and 0.8 μg/ml; for MMS 0.02, 0.04, 0.06 and 0.08% and for phleomycin 5 and 10 mg/ml.

## RESULTS

### Identification of putative haloarchaeal SSB proteins

As described in the 'Introduction' section, combined bioinformatic and biochemical studies have previously identified three OB fold-containing single-stranded DNA binding proteins MacRPA1, MacRPA2 and MacRPA3 in the methanogenic archaeal organism *M. acetivorans* C2A (15,17,18). MacRPA2 and MacRPA3 each contains two OB fold domains and a $CX_2CX_8CX_2H$ zinc finger motif while MacRPA1 encodes a protein with four OB fold domains and no zinc finger (Figure 1). In order to identify genes encoding putative SSB proteins in the haloarchaea (the *Halobacteriales* are a sister clade of the *Methanosarcinales*), BLAST searching (30) was performed using the three *M. acetivorans* RPA proteins as query sequences. Genes encoding orthologues of MacRPA1, MacRPA2 and MacRPA3 can be readily identified in all 14 completely sequenced haloarchaeal genomes in current databases (see Supplementary Tables S3 and S4 for

accession numbers and identity/similarity values with *M. acetivorans* RPA proteins).

In *Hfx. volcanii*, the MacRPA1, MacRPA2 and MacRPA3 orthologues are encoded by the *HVO_0519*, *HVO_1338* and *HVO_0292* genes, respectively (Figure 1A). Like MacRPA2, the product of *HVO_1338* possesses two putative OB folds, whereas the product of *HVO_0292* comprises only a single OB fold (MacRPA3 has two). Both proteins contain a zinc finger motif (although in the HVO_0292 protein and its haloarchaeal orthologues this has the sequence $CX_4CX_{8\text{-}12}CX_2H$, in contrast to $CX_2CX_8CX_2H$ in MacRPA2). It is highly likely that both *HVO_1338* and *HVO_0292* are co-transcribed with the downstream ORFs *HVO_1337* and *HVO_0291*, respectively (Figure 1A). The products of these downstream ORFs are related to one another and may also contain OB-fold motifs (see 'Discussion' section).

For simplicity, we designated the proteins encoded by the *Hfx. volcanii* genes as RpaA1 (encoded by *HVO_1338*, homologue of MacRPA2) and RpaB1 (*HVO_0292*, homologue of MacRPA3) and the proteins encoded by the downstream ORFs, RpaA2 (*HVO_1337*) and RpaB2 (*HVO_0291*). In *Hfx. volcanii* the ORFs encoding RpaA1 and RpaA2 overlap by three nucleotides, while those encoding RpaB1 and RpaB2 are separated by one nucleotide only. Similar gene organization is seen in all other haloarchaeal species with sequenced genomes (see Supplementary Tables S3 and S4) and in *M. acetivorans*, i.e. *RPA2* and *RPA3* are located adjacent to genes encoding homologues of RpaA2 and RpaB2, respectively (Figure 1A), but the biochemical properties of these proteins have not been analysed.

Downstream of the RpaA2 ORF in *Hfx. volcanii* is an ORF encoding a putative metallophosphoesterase protein of unknown function, Rpe (*HVO_1336*), that may also be part of the same operon; the RpaA2 and Rpe ORFs are separated by five nucleotides only. Although an Rpe-related ORF is found in the *M. acetivorans* genome, this is not linked to MacRPA2. Downstream of RpaB2 is the ORF encoding the haloarchaeal-specific Ral protein. In this case, the gap between the RpaB2 and Ral ORFs is 165 nucleotides, suggesting that Ral is not part of the same transcription unit, at least in *Hfx. volcanii* (see 'Discussion' section).

The *HVO_0519* gene encoding the closest MacRPA1 homologue in *Hfx. volcanii* is separated from its 5′ and 3′ flanking ORFs by almost 200 nt in each case and is presumably transcribed as a monocistronic mRNA. We designate the protein encoded by *HVO_0519* as RpaC. RpaC possesses three clear OB fold domains (in contrast to MacRPA1, which has four, see 'Discussion' section) plus conserved N- and C-terminal domains of unknown function (Figure 1B).

## Essential and non-essential RPA proteins

To investigate which, if any, of the *Hfx. volcanii* RPA proteins is essential for cell viability, we attempted to delete the corresponding genes from their native chromosomal loci. In the case of RpaA1 and RpaB1, we initially decided to delete the entire *rpaA1*$^+$–*rpaA2*$^+$–*rpe*$^+$ and *rpaB1*$^+$–*rpaB2*$^+$ loci as a precursor to more detailed analysis. To delete the genes, the pop-in/pop-out method was used with *Hfx. volcanii* H99 (Δ*trpA* Δ*hdrB* Δ*pyrE2*) (27,31). Using this method—full details of which can be found in the 'Materials and Methods' section—the regions of interest are replaced by an auxotrophic marker: *trpA*$^+$ for deletion of the *rpaA1*$^+$–*rpaA2*$^+$–*rpe*$^+$ operon and for deletion of *rpaC*$^+$, and *hdrB*$^+$ for deletion of *rpaB1*$^+$–*rpaB2*$^+$.

**Table 2.** Gene deletion experiments

| | Strain background | Attempted deletion | Non-selective medium | Selective medium |
|---|---|---|---|---|
| 1 | H99 | *rpaA::trpA*$^+$ | $9.9 \times 10^5$ | $8.8. \times 10^5$ |
| 2 | H99 | *rpaB::hdrB*$^+$ | $3.0 \times 10^4$ | $1.2 \times 10^2$ |
| 3 | H99 | *rpaC::trpA*$^+$ | $1.8 \times 10^5$ | 20 |

Cells were grown as described in 'Materials and Methods' section and plated onto non-selective or selective medium.
Colony-forming units per ml of cell culture are shown, determined by plating serial dilutions.

The results of these experiments were as follows. For deletion of the *rpaA1*$^+$–*rpaA2*$^+$–*rpe*$^+$ operon, we found comparable numbers of colonies on plates on selective and non-selective media (i.e. with and without tryptophan) following the pop-out step, implying that the products of the *rpaA1*$^+$, *rpaA2*$^+$ and *rpe*$^+$ genes are non-essential for cell viability in *Hfx. volcanii*. Loss of the *rpaA1*$^+$–*rpaA2*$^+$–*rpe*$^+$ *operon* was confirmed by PCR in all 16 putative *ΔrpaA1 ΔrpaA2 Δrpe* colonies analysed (Table 2, Figure 2A).

For *rpaB1*$^+$–*rpaB2*$^+$, the number of colonies on selective medium (i.e. Hv-Ca medium lacking thymidine and hypoxanthine to select for *hdrB*$^+$, see 'Materials and Methods' section) following pop-out was substantially lower than on non-selective medium (Table 2). However, 16 colonies from the selective plates screened by PCR all showed loss of the *rpaB1*$^+$–*rpaB2*$^+$ locus, indicating the products of the *rpaB1*$^+$ and *rpaB2*$^+$ genes are also non-essential (Figure 2B). We have observed a similar disparity between actual and expected colony number when creating gene deletions of other non-essential genes using *hdrB*$^+$ (but not *trpA*$^+$) as the selectable marker on Hv-Ca medium (unpublished results). When we repeated the *rpaB1*$^+$–*rpaB2*$^+$ deletion using *Hfx. volcanii* minimal medium (Hv-Min) instead of Hv-Ca, the number of colonies found on selective and non-selective plates was comparable and all the pop-outs screened from selective plates were *rpaB1*$^+$–*rpaB2*$^+$ deletions (data not shown). That indicates that Hv-Min is a more appropriate medium for gene deletion using the *hdrB*$^+$ marker than Hv-Ca.

For deletion of *rpaC*$^+$, we found $\sim 10^4$ more colonies on non-selective medium than on selective medium (i.e. medium lacking tryptophan) following pop-out. When screened by PCR for the presence or absence of *rpaC*$^+$, all 32 colonies analysed from plates lacking tryptophan still contained the *rpaC*$^+$ gene (Table 2, Figure 2C). These observations suggest strongly that *rpaC*$^+$ is an

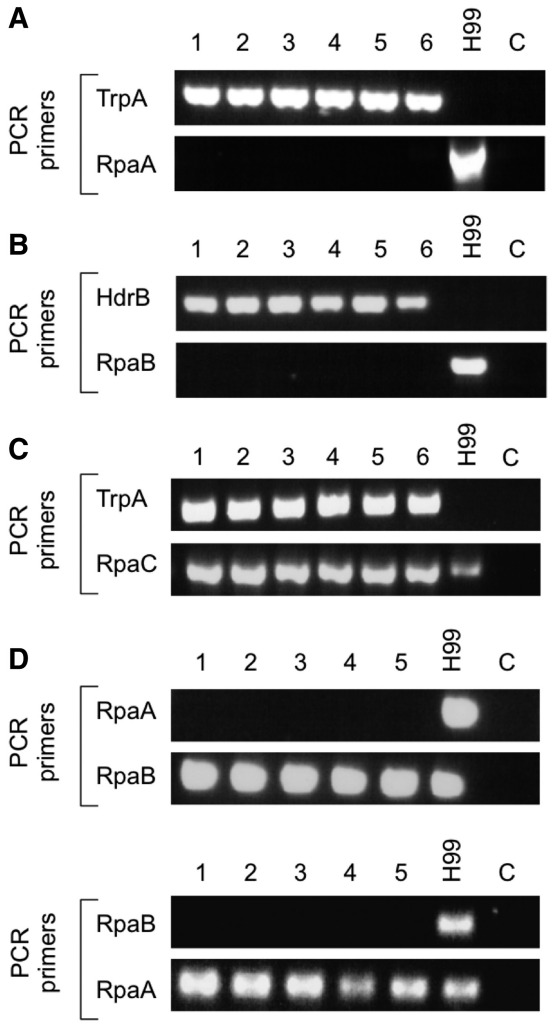


**Figure 2.** Gene deletions. (**A**) PCR analysis of chromosomal DNA prepared from six *rpaA::trpA*$^+$ colonies isolated on selective medium (i.e. Hv-Ca or Hv-Min medium containing 5FOA and uracil for counterselection but lacking either tryptophan or thymine/hypoxanthine for selection—see 'Materials and Methods' section for details) demonstrating successful deletion of the *rpaA1-rpaA2-rpe* locus. (**B**) Similar analysis of six *rpaB::hdrB*$^+$ colonies demonstrating successful deletion of the *rpaB1-rpaB2* locus. (**C**) PCR analysis of six representatives of the rare colonies isolated on selective medium following attempted replacement of *rpaC*$^+$ with *trpA*$^+$. All six still contain an intact *rpaC*$^+$ locus in addition to the *trpA*$^+$ marker (see 'Results' section). (**D**) Similar analysis of five of the rare colonies isolated on selective medium following attempted replacement of *rpaB*$^+$ with *hdrB*$^+$ in an *ΔrpaA1-rpaA2-rpe* genetic background (upper panel) or of *rpaA*$^+$ with *trpA*$^+$ in an *ΔrpaB1-rpaB2* background (lower case). In both cases, all the transformants analysed still contained the gene targeted for deletion as well as the marker gene, either *hdrB*$^+$ or *trpA*$^+$ (data not shown). Note that in each case (A–D) the primers used for PCR were located within the ORFs targeted for deletion (see Supplementary Table S1 for primer sequences). See text for further details.

essential gene in *Hfx. volcanii*—the colonies obtained on the selective medium are presumably either merozygotic (that is, they possess both *rpaC*$^+$ and *rpaC::trpA*$^+$ loci on different chromosomes after pop-out) or carry mutated *pyrE2* genes that confer resistance to 5-FOA (with the result that both *rpaC*$^+$ and *rpaC::trpA*$^+$ loci are present on the same chromosome).

The relatedness of the products of *rpaA1*$^+$ and *rpaB1*$^+$, and to lesser extent *rpaA2*$^+$ and *rpaB2*$^+$, raised the possibility that these genes might share an essential function. To test this, we attempted to create an Δ*rpaA* Δ*rpaB* double deletion strain. This was done in three ways. First, we attempted to replace the *rpaA1*$^+$*–rpaA2*$^+$*–rpe*$^+$ operon with *trpA*$^+$ in an *rpaB*::*hdrB*$^+$ background. For this purpose, plasmid PL6 (see 'Materials and Methods' section) was transformed into the *rpaB*::*hdrB*$^+$ strain and integrant (pop-in) colonies were plated on medium containing uracil and 5-FOA with and without tryptophan to identify (pop-out) colonies of cells that had lost the plasmid. Colonies from plates lacking tryptophan were screened by PCR to detect the presence or absence of the *rpaA1*$^+$*–rpaA2*$^+$*–rpe*$^+$ operon. However, all the colonies examined contained the *rpaA1*$^+$*–rpaA2*$^+$*–rpe*$^+$ operon, suggesting that a Δ*rpaA* Δ*rpaB* double mutant is not viable (Figure 2D, upper panel). Consistent with this, an identical result was obtained when we attempted to replace the *rpaB1*$^+$*–rpaB2*$^+$ operon with *hdrB*$^+$ in *rpaA*::*trpA*$^+$ background using plasmid PL8. While deletion of *rpaB1*$^+$*–rpaB2*$^+$ in the wild-type was straightforward (see above), we were unable to delete the gene in a *rpaA*::*trpA*$^+$ strain (Figure 2D, lower panel).

Finally, we attempted to construct an Δ*rpaA* Δ*rpaB* double deletion strain by mating Δ*rpaA* and Δ*rpaB* strains, as described previously (26). Briefly, exponentially growing cultures of *rpaA*::*trpA*$^+$ *ΔhdrB* and *rpaB*::*hdrB*$^+$ Δ*trpA* strains were combined, harvested by filtration and incubated overnight on rich medium. The cells were then plated on medium lacking both tryptophan and thymidine to select for double mutants. Control matings were performed with *Hfx. volcanii* H53 (Δ*trpA hdrB*$^+$) and H98 (*trpA*$^+$ Δ*hdrB*) strains (see 'Materials and Methods' section). In two independent mating experiments we were unable to obtain *rpaA*::*trpA*$^+$ *rpaB*::*hdrB*$^+$ double mutants from the *rpaA*::*trpA*$^+$ *ΔhdrB* × *rpaB*::*hdrB*$^+$ Δ*trpA* mating, despite *rpaA*::*trpA*$^+$ *hdrB*$^+$ and *trpA*$^+$ *rpaB*::*hdrB*$^+$ strains being abundant following the control matings (*rpaA*::*trpA*$^+$ *ΔhdrB* × Δ*trpA hdrB*$^+$ and *trpA*$^+$ Δ*hdrB* × *rpaB*::*hdrB*$^+$ Δ*trpA*, respectively). Taken together, the results of these experiments imply that the products of the *rpaA1*$^+$*–rpaA2*$^+$*–rpe*$^+$ and *rpaB1*$^+$*–rpaB2*$^+$ operons share at least one essential function (see 'Discussion' section).

### Conditional down-regulation of gene expression

In order to analyse the *in vivo* function of RpaC in greater detail, we used a tryptophan-repressible promoter system that allows conditional down-regulation of gene expression. The promoter of the *Hfx. volcanii* tryptophanase gene *tnaA* is strongly induced in the presence of tryptophan and repressed in its absence (29). Initially, we engineered a strain in which the *tnaA* promoter was positioned upstream of *rpaC*$^+$, displacing the native promoter. This was achieved by constructing an integrating plasmid pNPM-tna–rpaC carrying the *tna* promoter 5′ to 300 bp from the 5′ end of the *rpaC*$^+$ ORF (PL14, see 'Materials and Methods' section). This plasmid was transformed into *Hfx. volcanii* H98 and transformant colonies obtained on

Hv-Ca medium [which contains low levels of tryptophan, enough to activate p*tna*, see (29)] were analysed by PCR to confirm correct integration of the plasmid (data not shown).

To test the ability of the integrated p*tna–rpaC*$^+$ strain to grow in the absence of tryptophan, serial dilutions of exponentially-growing cultures were spotted onto Hv-Min medium plates lacking tryptophan and, as a control, onto Hv-Min plates supplemented with 0.075 mM tryptophan (see 'Materials and Methods' section). Little or no growth was seen on medium lacking tryptophan, consistent with *rpaC*$^+$ being an essential gene in *Hfx. volcanii* (Figure 3A, upper panel). Analysis of expression levels by quantitative reverse transcription real-time PCR (qRT-PCR) confirmed that the *rpaC*$^+$ mRNA level in cells grown in Hv-Min medium lacking tryptophan was reduced to below the level seen in wild-type cells. At the same time, we observed that the *rpaC*$^+$ expression level in Hv-Min medium with 0.075 mM tryptophan was elevated above the normal wild-type level (data not shown) but since there is no detectable difference in growth rate between the wild-type and the integrated p*tna–rpaC*$^+$ strain overexpression does not appear to have a detrimental effect on cells.

Taken together, these results show that the *tnaA* promoter can be successfully used for down-regulation of *rpaC*$^+$ expression. However, the plasmid integration strategy described above creates a strain in which homologous recombination (between truncated and full-length copies of *rpaC*$^+$) will lead to reversion to wild-type. In order to prevent this, a genetically stable p*tna–rpaC*$^+$ strain was constructed using the pop-in/pop-out method, in which the *rpaC*$^+$ promoter was replaced by p*tna* (see 'Material and Methods' section for details of strain construction). As expected, the stable p*tna–rpaC*$^+$ strain also showed reduced growth rate on medium lacking tryptophan (Figure 3A, lower panel).

As shown above, the products of the *rpaA1*$^+$*–rpaA2*$^+$*–rpe*$^+$ and *rpaB1*$^+$*-rpaB2*$^+$ operons appear to share at least one essential function—simultaneous deletion of the two loci is not possible. To extend this observation, we examined the effects of placing the *rpaB1*$^+$*-rpaB2*$^+$ locus under *tna* control in a *ΔrpaA* background. As expected, growth of the p*tna-rpaB ΔrpaA* strain was significantly retarded on medium lacking tryptophan (Figure 3B).

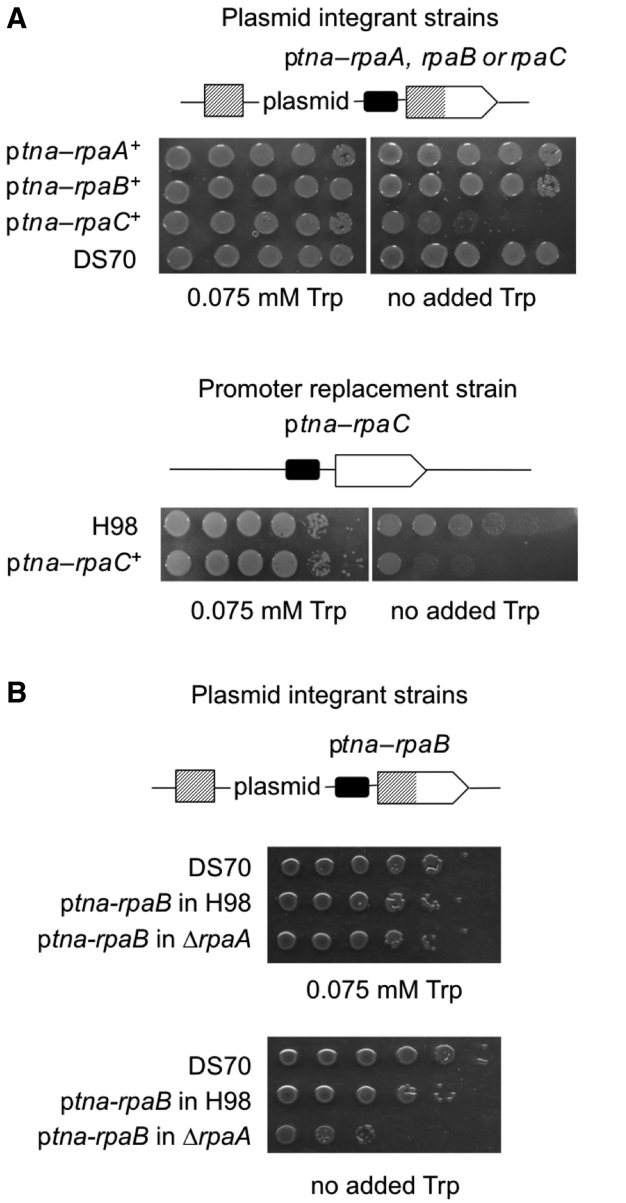


**Figure 3.** Construction of conditional lethal strains. (**A**) Upper panel: *Hfx. volcanii* H98 strains carrying integrated pNPM-tna-Rpa plasmids (see Table 1) were grown up to mid-exponential phase at 45°C in Hv-Min medium containing 0.075 mM tryptophan. The cells were then washed and 10-fold serial dilutions spotted onto Hv-Min plates with or without tryptophan, and incubated at 45°C for 3 days. Strain DS70 is a wild-type control (Table 1). Lower panel: the *Hfx. volcanii ptna–rpaC* promoter replacement strain (Table 1) was processed and incubated as above. Strain H98 is a wild-type control. (**B**) *Hfx. volcanii ΔrpaA* strains carrying integrated pNPM-tna-RpaB (see Table 1) were processed and incubated as above.

## Interplay between SSB proteins

The presence of multiple single-stranded DNA binding factors in *Hfx. volcanii* raises the question of whether there is interplay between them. To examine this, we first tested whether high level expression of either *rpaA1*$^+$*–rpaA2*$^+$*–rpe*$^+$ or *rpaB1*$^+$*–rpaB2*$^+$ would rescue loss of *rpaC*$^+$ function in p*tna–rpaC*$^+$ cells growing on medium lacking tryptophan. To this end, the *fdx* promoter was placed 5′ to the *rpaA1*$^+$*–rpaA2*$^+$*–rpe*$^+$ and *rpaB1*$^+$*–rpaB2*$^+$ loci in a p*tna–rpaC*$^+$ background by plasmid integration (see 'Material and Methods' section for details of strain construction). The *fdx* promoter, derived from the ferredoxin gene from *Hbt. salinarum*, promotes strong constitutive expression in *Hfx. volcanii* (32). Serial dilutions of exponentially growing p*fdx–rpaA1*$^+$*–rpaA2*$^+$*–rpe*$^+$ p*tna–rpaC*$^+$ and p*fdx–rpaB1*$^+$*–rpaB2*$^+$ p*tna–rpaC*$^+$ integrant strains were then spotted onto Hv-Min medium plates lacking tryptophan with appropriate controls. We found that elevated expression of *rpaB1*$^+$*–rpaB2*$^+$ but not *rpaA1*$^+$*–rpaA2*$^+$*–rpe*$^+$ could complement for lack of *rpaC*$^+$ (Figure 4A): growth of

p*fdx–rpaB1*$^{+}$*–rpaB2*$^{+}$ p*tna–rpaC* cells was as strong as wild-type cells on Hv-Min medium lacking tryptophan, whereas growth of the p*fdx–rpaA1*$^{+}$*–rpaA2*$^{+}$*–rpe*$^{+}$ p*tna–rpaC* strain was indistinguishable from that of the parental p*tna–rpaC* strain. Thus, at least one of the products of the *rpaB1*$^{+}$*–rpaB2*$^{+}$ operon is able to substitute, directly or indirectly, for loss of RpaC.

### Cross-species complementation

In many model organisms, construction of a conditional-lethal mutant strain would ordinarily facilitate rapid structure–function analysis of the corresponding protein. However, in our experience the highly efficient homologous recombination seen with *Hfx. volcanii* greatly hampers these studies, with gene conversion between the wild-type chromosomal copy of the gene and incoming mutated copies often resulting in loss of the mutant DNA sequences before any functional test can be applied (unpublished data). To bypass these difficulties, we tested whether RpaC orthologues from related haloarchaeal species, displaying only limited sequence similarity at the nucleotide sequence level to the *Hfx. volcanii rpaC*$^{+}$ gene but encoding closely related proteins, would function in *Hfx. volcanii* and rescue growth of p*tna–rpaC* cells grown in the absence of tryptophan. The RpaC homologues from *Halogeometricum borinquense* (Hbor_26830; GI:313127408) and *Haloquadratum walsbyi* (HQ_1435A; GI:110667397) are 75% and 61% identical to the *Hfx. volcanii* protein, respectively. We constructed replicating plasmids carrying these genes under the control of the *fdx* promoter (see 'Material and Methods' section) and transformed these into the p*tna–rpaC* strain. As shown in Figure 4B, expression of either the *Hgm. borinquense* or *Hqr. walsbyi rpaC*$^{+}$ genes from the *fdx* promoter restores growth of the p*tna–rpaC* strain in medium lacking tryptophan. No rescue is seen when the cells are transformed with the empty plasmid (Figure 4B, lower panel, top row). Subsequent experiments showed that the level of rescue observed with *Hgm. borinquense* RpaC was indistinguishable from that seen when *Hfx. volcanii* RpaC was expressed under equivalent conditions (data not shown).

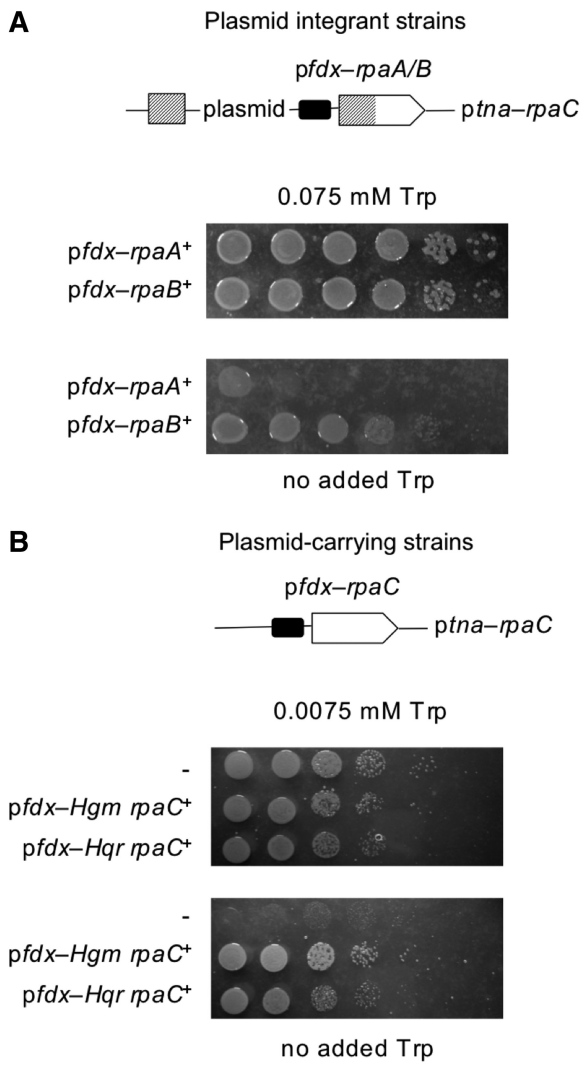


**Figure 4.** Rescue of p*tna–rpaC* strain by expression of RpaB and Hgm RpaC. (**A**) *Hfx. volcanii* p*tna–rpaC* cells carrying integrated pNPM-fdx-RpaA or pNPM-fdx-RpaB plasmids were grown mid-exponential phase at 45°C in Hv-Min medium containing 0.075 mM tryptophan. The cells were then washed and ten-fold serial dilutions spotted onto Hv-Min plates with or without tryptophan and incubated at 45°C for 3 days. (**B**) *Hfx. volcanii* p*tna–rpaC* cells carrying plasmids pTA230-fdx-HgmRpaC or pTA230-fdx-HqrRpaC were grown mid-exponential phase at 45°C in Hv-Min medium containing 0.075 mM tryptophan, washed, serially diluted as above and spotted onto Hv-Min plates with or without tryptophan and incubated at 45°C for 3 days.

### Structure–function analysis of RpaC

Haloarchaeal RpaC orthologues display high conservation in domain organization, with three OB folds and N- and C-terminal domains (Figure 1B). The N-terminal domain is present in all haloarchaeal RpaC proteins, and in some methanogens (notably the *Methanosarcinales*, including *M. acetivorans*), whereas the C-terminal domain appears to be unique to the haloarchaea (see 'Discussion' section).

Studies of the four OB fold-containing MacRPA1 protein have shown that single OB motifs can be deleted without significant loss of ssDNA-binding ability (14) but it is not clear how these changes to the structure of the protein impact on *in vivo* function. To examine this, and to probe the functions of the N- and C-terminal domains, we generated a collection of truncated and deleted derivatives of *Hgm. borinquense* RpaC and tested whether these were able to support cell growth when expressed in p*tna–rpaC* cells grown on medium lacking tryptophan. As shown schematically in Figure 5A, we expressed nine deleted or truncated versions of the *Hgm. borinquense* RpaC protein: N- and C-terminal deletions (ΔNTD and ΔCTD respectively), single OB fold deletions ($\Delta OB^{A}$, $\Delta OB^{B}$ and $\Delta OB^{C}$), double OB fold deletions ($\Delta OB^{AB}$, $\Delta OB^{BC}$ and $\Delta OB^{AC}$) and a triple OB deletion ($\Delta OB^{ABC}$). Note that, in contrast to previously reported biochemical analysis of the *M. acetivorans* MacRPA1 protein, we chose to delete individual OB fold domains in their entirety rather than create chimeric OB folds by fusing adjacent folds (15).

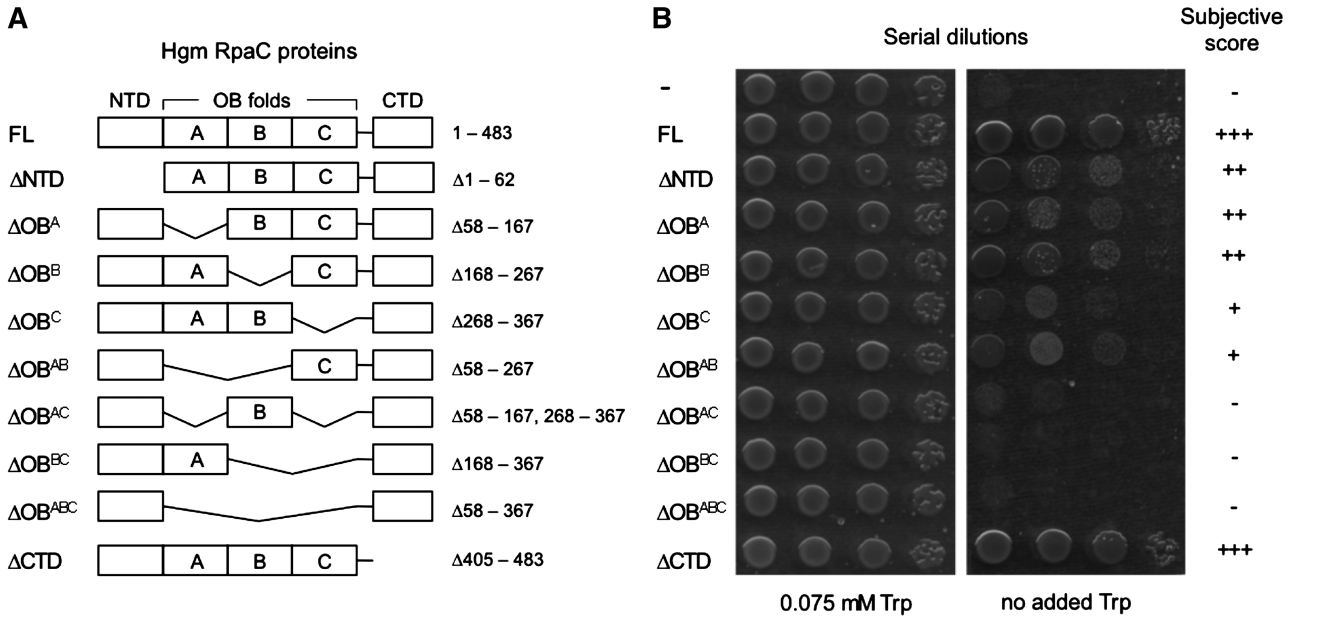


**Figure 5.** Structure–function analysis of RpaC. (**A**) Schematic representation of nine mutant *Hgm. borinquense* RpaC proteins indicating (boxed) the location of the N-terminal domain (NTD), OB folds A–C and the C-terminal domain (CTD) and the extent of the deleted regions. (**B**) *Hfx. volcanii* p*tna–rpaC* cells carrying various pTA230-fdx-HgmRpaC plasmids (see Supplementary Table S2) were grown up to mid-exponential phase at 45°C in Hv-Min medium containing 0.075 mM tryptophan. The cells were then washed and 10-fold serial dilutions spotted onto Hv-Min plates with or without tryptophan, and incubated at 45°C for 3 days.

Plasmids expressing each of the nine alleles from the *fdx* promoter were transformed into the p*tna–rpaC* strain and transformants tested for their ability to grow on medium lacking tryptophan by spotting of serial dilutions of exponentially growing cultures (Figure 5B). The results were as follows—none of the three OB folds, the NTD nor the CTD is individually essential for RpaC function: all five single domain deletions (ΔNTD, ΔCTD, $\Delta OB^A$, $\Delta OB^B$ and $\Delta OB^C$) were able to restore growth to the p*tna–rpaC* strain, although in each the degree of rescue was reduced when compared to the full-length RpaC protein (Figure 5B). OB fold C appears to have the most important role, as no growth was observed when $\Delta OB^{AC}$ and $\Delta OB^{BC}$ were expressed, while $\Delta OB^{AB}$ retained some rescue activity. Consistent with this, $\Delta OB^A$ and $\Delta OB^B$ rescued better than $\Delta OB^C$, although none of these rescued as well as the full-length protein, indicating that all three OB folds are necessary for optimal RpaC function. No rescue was observed in cells expressing the triple OB fold deletion mutant $\Delta OB^{ABC}$.

Cells expressing ΔNTD also grew more poorly than cells expressing the full-length protein while rescue by ΔCTD was comparable to wild type. Subsequent western blotting of N-terminally Flag-tagged versions of all nine mutant proteins plus full-length Hgm RpaC confirmed that all nine were expressed to similar levels in *Hfx. volcanii* (data not shown).

### Cells expressing ΔNTD and ΔCTD mutants are sensitive to DNA-damaging agents

Despite being non-essential for RpaC function, the N- and C-terminal domain sequences are reasonably well conserved across haloarchaeal species suggesting that they play important roles *in vivo*. As the RPA proteins are likely to be involved in DNA repair, we examined whether p*tna–rpaC* cells expressing truncated Hgm RpaC-ΔNTD and RpaC-ΔCTD proteins were more sensitive to the effects of various DNA damaging agents than cells expressing full length Hgm RpaC. As summarized in Figure 6, cells expressing RpaC-ΔNTD showed increased sensitivity to UV light (Figure 6A) and to MMS and phleomycin (Figure 6B), whereas cells expressing RpaC-ΔCTD mutant displayed increased sensitivity to UV irradiation and MMS but not to phleomycin. These results suggest that the N-terminal domain may have a general role in RpaC function in DNA repair whereas the role of the C-terminal domain may be confined to certain types of DNA repair only (see 'Discussion' section).

### Increased expression of RpaC leads to resistance to DNA damage

Previous work with *Hbt. salinarum* has established a link between increased expression of the orthologous $rpaB1^+$–$rpaB2^+$ operon and resistance to ionizing radiation (33). To investigate whether increased expression of the *Hfx. volcanii* $rpaC^+$ might also lead to resistance to DNA damage, p*fdx–rpaC* and p*tna–rpaC* strains (the latter grown on medium containing tryptophan) were exposed to a variety of DNA-damaging agents in liquid culture and cell viability determined by colony counting. We observed enhanced resistance of both p*fdx–rpaC* and p*tna–rpaC* strains to all DNA-damaging agents tested (Figure 7). Taken together with essentiality of RpaC and

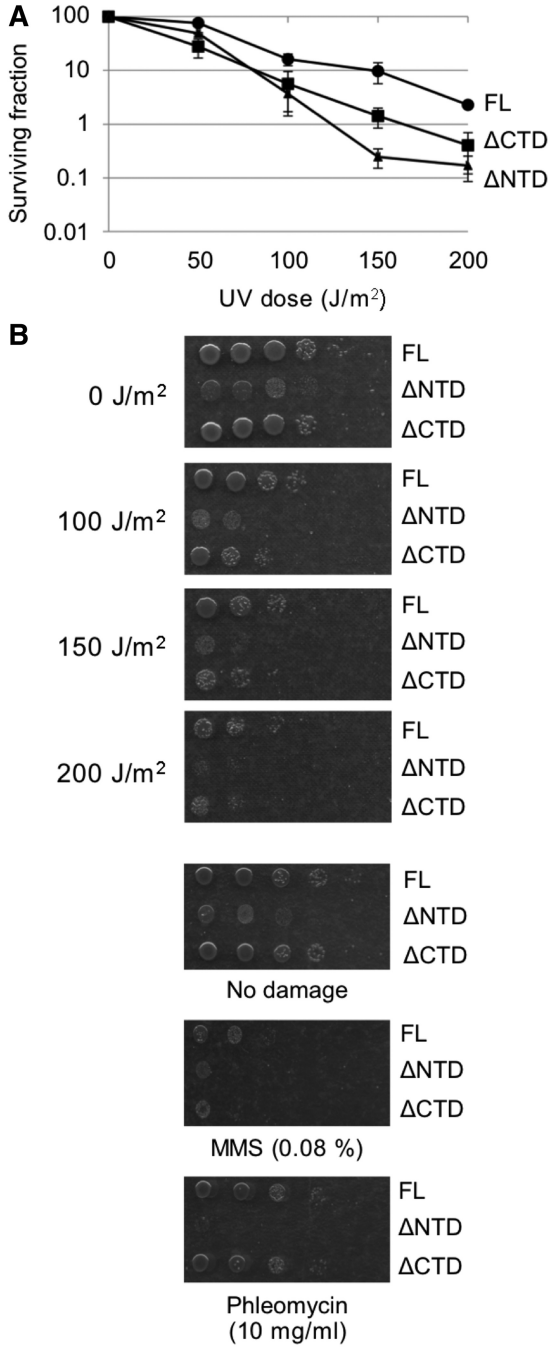


**Figure 6.** Sensitivity of cells expressing Hgm RpaC-ΔNTD and -ΔCTD to DNA damage. (**A**) Response to UV treatment. *Hfx. volcanii* p*tna–rpaC* cells carrying various pTA230-fdx-HgmRpaC plasmids were grown up to mid-exponential phase at 45°C in Hv-Min medium containing 0.075 mM tryptophan. The cells were then washed and ten-fold serial dilutions spotted onto Hv-Min plates, irradiated with the indicated dose of UV and incubated at 45°C for 3 days. Upper panel: survival curves. Lower panel: spotting assay. (**B**) Response to MMS and phleomycin treatment. The same strains were grown up and exposed to MMS (0.08%) or phleomycin (10 mg/ml) for 1 h, before being serially diluted, spotted and incubated at 45°C for 3 days. See 'Materials and Methods' section for details.

the DNA-damaging sensitivity seen in cells expressing N- and C-terminally truncated RpaC proteins, these results underline the contribution of RpaC to maintaining genome integrity in the haloarchaea.

## DISCUSSION

In all three domains of life, SSBs play vital roles in DNA replication, repair and recombination. In this report we present the first genetic analysis of RPA function in the third domain of life, the archaea, using the euryarchaeon *Hfx. volcanii* as a model to take advantage of the wide range of genetic tools available for this organism (22).

Previously, detailed biochemical analysis of the SSB proteins MacRPA1, MacRPA2 and MacRPA3 from the methanogenic archaeal species *M. acetivorans* has been described (14–18). The MacRPA2 and MacRPA3 proteins are related to one another, indicative of descent from a common ancestor: each possesses two OB folds plus a zinc finger motif (Figure 1B). Database searching readily identified homologues of these proteins in the haloarchaea, a sister clade to the methanogens. The haloarchaeal MacRPA2 and MacRPA3 homologues (RpaA1 and RpaB1 in *Hfx. volcanii*, respectively) have a similar domain organization to their methanogen counterparts, with one or two OB folds and a zinc finger motif. Intriguingly, in both cases (and in both organisms, *Hfx. volcanii* and *M. acetivorans*) the ORFs encoding these proteins are most likely co-transcribed with adjacent ORFs encoding proteins of the COG3390 family (RpaA2 and RpaB2 in *Hfx. volcanii*, MA0591 and MA3018 in *M. acetivorans*, see Figure 1A). Iterative database searching using PSI-BLAST (30) reveals that these proteins also contain at least one OB fold (data not shown), raising the possibility that RpaA1 and RpaA2, and RpaB1 and RpaB2, may form heteromeric SSB complexes *in vivo*. Also possible are mixed complexes comprising, for example, RpaA1 and RpaB2, or RpaB1 and RpaA2, although the non-essentiality of the individual *rpaA1*$^+$–*rpaA2*$^+$–*rpe*$^+$ and *rpaB1*$^+$–*rpaB2*$^+$ operons indicates that mixed complexes, should they exist, are not obligatory for cell growth.

The *rpaA* locus also contains a gene, *rpe*$^+$, downstream of *rpaA2*$^+$, encoding a putative metallophosphoesterase of unknown function. In *Hfx. volcanii*, the *rpaA2*$^+$ and *rpe*$^+$ ORFs are separated by five nucleotides only, suggesting that they are likely to be co-transcribed although this has not been tested. A related gene is found in methanogens, including *M. activorans*, but is unlinked to the *MacRPA2–MA0591* locus (Figure 1A, data not shown). Our deletion analysis shows that *rpe*$^+$ is a non-essential gene (Figure 2). An additional ORF, *ral*$^+$, is also found downstream of *rpaB1*$^+$ and *rpaB2*$^+$. In *Halobacterium* sp. NRC-1, this appears to be co-transcribed with the *rpaB1*$^+$ and *rpaB2*$^+$ orthologues but whether this is true in *Hfx. volcanii* (where the *rpaB2*$^+$ and *ral*$^+$ ORFs are separated by 165 bp, in contrast to only 45 bp in *Halobacterium* sp. NRC-1) is not known: in several haloarchaeal species, the *rpaB2*$^+$ and *ral*$^+$ ORFs are physically separated by ORFs encoding (or predicted to encode) unrelated proteins,

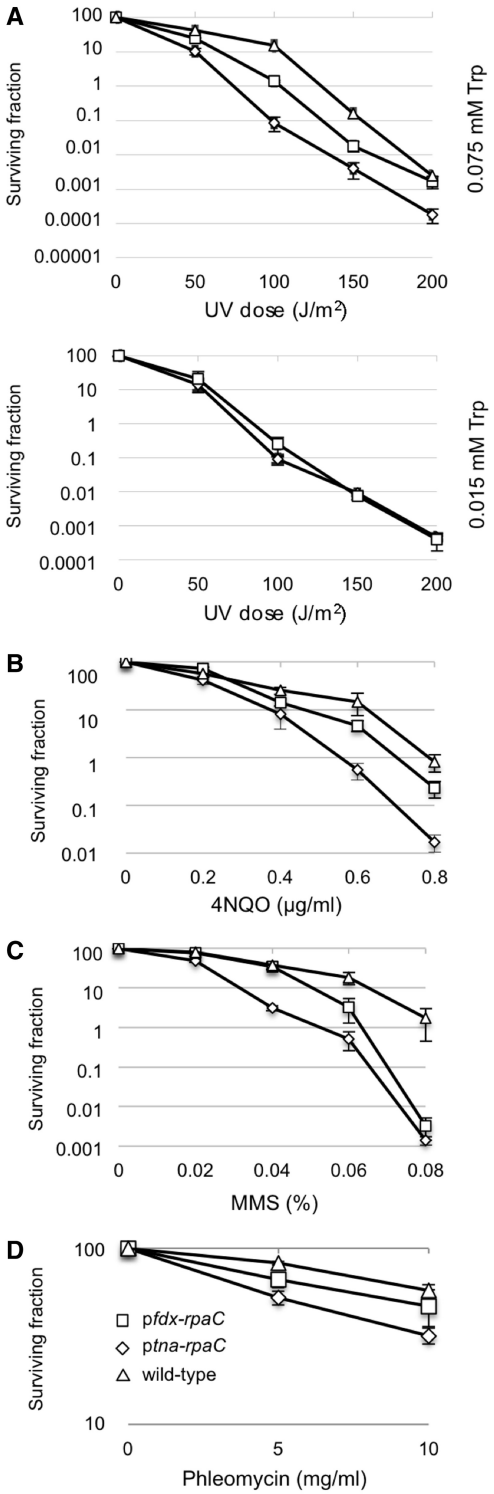


**Figure 7.** Increased expression of RpaC leads to enhanced resistance to DNA damage. (**A**) *Hfx. volcanii* H98 cells carrying integrated pNPM-tna–rpaC (squares) or pNPM-fdx-RpaC (triangles) plasmids, or *Hfx. volcanii* DS70 cells (diamonds), were plated on Hv-Min plates containing either 0.075 or 0.015 mM tryptophan and irradiated with UV at the doses shown. Survival was determined by colony counting after 4 days incubation at 45°C. (**B–D**) The same three strains were grown to mid-exponential phase in Hv-Min medium containing 0.075 mM tryptophan and treated with the indicated DNA-damaging agents for 1 h before being plated on the same medium. Survival was determined by colony counting after 4 days incubation at 45°C. See 'Materials and Methods' section for details.

indicating that co-expression of *rpaB1*$^+$–*rpaB2*$^+$ with *ral*$^+$ is not obligatory (data not shown). Importantly, levels of the *Halobacterium rpaB1*$^+$–*rpaB2*$^+$–*ral*$^+$ transcript were shown to be increased in two mutant strains displaying increased resistance to ionizing radiation (33), possibly implying a role for these proteins in the repair of DNA double-strand breaks, and following UV treatment (34). Consistent with latter result, deletion of the *Hfx. volcanii rpaB1*$^+$–*rpaB2*$^+$ operon results in increased UV sensitivity (data not shown). The Ral protein does not contain an OB fold and is unlikely to have any role in ssDNA binding.

While the individual *rpaA1*$^+$–*rpaA2*$^+$–*rpe*$^+$ and *rpaB1*$^+$–*rpaB2*$^+$ deletions were viable, we were unable to construct the double *ΔrpaA ΔrpaB* deletion, either by the pop-in/pop-out method or by mating, suggesting that the products of these loci share an essential function. The construction of a p*tna-rpaB ΔrpaA* strain (Figure 3B) offers a useful tool for future analysis of this function.

In contrast to the *rpaA1*$^+$–*rpaA2*$^+$–*rpe*$^+$ and *rpaB1*$^+$–*rpaB2*$^+$ loci, *rpaC*$^+$ is individually essential for cell growth. We were unable to delete the *rpaC* gene from the chromosome and downregulation of *rpaC*$^+$ expression using the *tna* promoter caused growth arrest of p*tna–rpaC* strains on medium lacking tryptophan (Figure 3A). Interestingly, elevated expression of the *rpaB1-rpaB2* operon, driven by the constitutive *fdx* promoter, is sufficient to partially suppress the growth defect of p*tna–rpaC*$^+$ cells (Figure 4A), implying that the RpaB1 protein, possibly acting in concert with RpaB2, is able to perform some of the cellular functions of RpaC.

Using the closely related *Hgm. borinquense* protein as a model, we examined the consequences of deleting the N- and C-terminal and OB fold domains of RpaC by testing the ability of the mutated proteins to rescue growth of the p*tna–rpaC* strain on medium lacking tryptophan (Figure 5). Individual deletion of the three OB folds reduced but did not abolish growth, indicating that RpaC function is not dependent on the presence of all three OB folds. In contrast, deletion of OB fold C together with either A or B prevented rescue of p*tna–rpaC*. Some growth was seen when OB folds A and B were deleted, implying that an RpaC protein containing OB fold C alone retains residual activity. It remains to be seen how these deletions affect DNA binding by the Hgm RpaC protein and whether there is a direct correlation between *in vivo* function and affinity for ssDNA. Previous work with the *M. acetivorans* MacRPA1 protein has shown that ssDNA binding by this protein is largely unaffected by deletion of single OB folds (achieved by the creation of chimeric OB folds by, for example, fusing the N-terminal part of OB fold A with the C-terminal part of OB fold B), although ssDNA-binding affinities are somewhat reduced and the ability to discriminate between ssDNA and dsDNA impaired (15). Interestingly, deletion of single OB folds in this way did not affect the size of the ssDNA-binding site.

While MacRPA1 comprises four OB folds A–D, the haloarchaeal RpaC protein contains only three, equivalent to folds A-C in the MacRPA1 protein. A low level of sequence similarity can be observed between the MacRPA1 OB fold D and the region of RpaC that we

have designated the C-terminal domain (data not shown) suggesting that the latter is a highly divergent form of the former that is no longer recognizable as an OB fold. The RpaC CTD is also separated from OB fold C by a linker sequence (indicated by the short black line in Figure 5A) that is not conserved even among the haloarchaea. Deletion of the CTD from the RpaC protein did not affect its ability to rescue the p*tna–rpaC*$^{+}$ strain, although the rescued cells were significantly more sensitive to UV- and MMS-induced DNA damage (Figure 6). No increase in sensitivity to phleomycin was seen, however, suggesting that the CTD does not play an important role in the repair of dsDNA breaks.

Another highly conserved feature of the haloarcheal RpaC proteins is the N-terminal domain (NTD). This 60-amino-acid domain, which centres on a 15-amino-acid sequence (residues 27–41 in the *Hfx. volcanii* RpaC protein) that is identical in all 14 haloarchaeal RpaC proteins in current databases, is also found in the methanogen RpaC-like proteins such as MacRPA1. There is no evidence to suggest that the NTD forms an OB fold; indeed, structure modelling using Phyre2 (35,36) suggests, albeit with low confidence, that the NTD may form a three-helix bundle structure, with the perfectly conserved 15-amino-acid sequence spanning the C-terminal end of the second helix, an unstructured loop and the N-terminal end of the third helix (data not shown). Whether this prediction is a valid one awaits determination of the structure of RpaC. Deletion of the NTD impairs RpaC function; cells expressing the ΔNTD protein are more sensitive to DNA damage (Figures 5 and 6). We are currently screening for proteins that interact directly with the NTD in an effort to elucidate the role of this part of the RpaC protein.

In contrast to the DNA damage sensitivities observed when domains of RpaC are deleted, overexpression of *rpaC* (either from the constitutive *H. salinarum fdx* promoter or from the *tna* promoter in cells growing on tryptophan) resulted in an increased resistance to several DNA-damaging agents without any measurable affect on growth rate (Figure 7). We used four DNA-damaging agents that cause different types of damage: UV irradiation and 4NQO (4-nitroquinoline 1-oxide) introduce bulky adducts such as cyclobutane-pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4PPs) into DNA, MMS (methyl methanesulfonate) methylates DNA on $N^7$-deoxyguanine and $N^3$-deoxyadenine which can lead to replication fork stalling, and phleomycin causes DNA double-strand breaks. The increased resistance observed to each type of damage when *rpaC*$^{+}$ is overexpressed implies that RpaC function (or more likely, SSB function in general) is limiting in wild-type cells. A similar conclusion can be drawn from the earlier observation that expression of the *rpaB1-rpaB2* operon in *Halobacterium* sp. NRC1 is induced in response to both ionizing radiation and UV light treatments (33,34). In neither of these cases was *rpaC*$^{+}$ expression reported to be increased, and we do not know if there is similar transcriptional regulation of SSB function in *Hfx. volcanii*.

In summary, our results offer the first insights into the functional and structural properties of the haloarchaeal RPA proteins. Future work will focus on determining the exact roles of these proteins in haloarchaeal DNA replication, repair and recombination, on understanding precisely how each interacts with ssDNA in the high-salt intracellular environment of the haloarchaeal cell, and on identifying the proteins with which each RPA interacts to perform its functions.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables S1–S4.

## ACKNOWLEDGEMENTS

We are grateful to our colleagues in St Andrews and elsewhere for their help during the course of this work, in particular Dr Thorsten Allers (University of Nottingham, UK) for plasmids and strains and Prof. Malcolm White (University of St Andrews) for comments on the manuscript.

## FUNDING

The Scottish Universities Life Sciences Alliance (SULSA) and by the USAF Office of Scientific Research under award number FA9550-10-1-0421. Funding for open access charge: University of St Andrews.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Broderick,S., Rehmet,K., Concannon,C. and Nasheuer,H.P. (2010) Eukaryotic single-stranded DNA binding proteins: central factors in genome stability. *Subcell Biochem.*, **50**, 143–163.
2. Pestryakov,P.E. and Lavrik,O.I. (2008) Mechanisms of single-stranded DNA-binding protein functioning in cellular DNA metabolism. *Biochemistry*, **73**, 1388–1404.
3. Flynn,R.L. and Zou,L. (2010) Oligonucleotide/oligosaccharide-binding fold proteins: a growing family of genome guardians. *Crit. Rev. Biochem. Mol.*, **45**, 266–275.
4. Theobald,D.L., Mitton-Fry,R.M. and Wuttke,D.S. (2003) Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.*, **32**, 115–133.
5. Iftode,C., Daniely,Y. and Borowiec,J.A. (1999) Replication protein A (RPA): The eukaryotic SSB. *Crit. Rev. Biochem. Mol.*, **34**, 141–180.
6. Haring,S.J., Humphreys,T.D. and Wold,M.S. (2010) A naturally occurring human RPA subunit homolog does not support DNA replication or cell-cycle progression. *Nucleic Acids Res.*, **38**, 846–858.
7. Kemp,M.G., Mason,A.C., Carreira,A., Reardon,J.T., Haring,S.J., Borgstahl,G.E., Kowalczykowski,S.C., Sancar,A. and Wold,M.S. (2010) An alternative form of replication protein A expressed in normal human tissues supports DNA repair. *J. Biol. Chem.*, **285**, 4788–4797.
8. Sakaguchi,K., Ishibashi,T., Uchiyama,Y. and Iwabata,K. (2009) The multi-replication protein A (RPA) system–a new perspective. *FEBS J.*, **276**, 943–963.
9. Price,C.M., Boltz,K.A., Chaiken,M.F., Stewart,J.A., Beilstein,M.A. and Shippen,D.E. (2010) Evolution of CST function in telomere maintenance. *Cell Cycle*, **9**, 3157–3165.
10. Huang,J., Gong,Z., Ghosal,G. and Chen,J. (2009) SOSS complexes participate in the maintenance of genomic stability. *Mol. Cell*, **35**, 384–393.

11. Li,Y., Bolderson,E., Kumar,R., Muniandy,P.A., Xue,Y., Richard,D.J., Seidman,M., Pandita,T.K., Khanna,K.K. and Wang,W. (2009) HSSB1 and hSSB2 form similar multiprotein complexes that participate in DNA damage response. *J. Biol. Chem.*, **284**, 23525–23531.
12. Richard,D.J., Bolderson,E., Cubeddu,L., Wadsworth,R.I., Savage,K., Sharma,G.G., Nicolette,M.L., Tsvetanov,S., McIlwraith,M.J., Pandita,R.K. *et al.* (2008) Single-stranded DNA-binding protein hSSB1 is critical for genomic stability. *Nature*, **453**, 677–681.
13. Richard,D.J., Bolderson,E. and Khanna,K.K. (2009) Multiple human single-stranded DNA binding proteins function in genome maintenance: structural, biochemical and functional analysis. *Crit. Rev. Biochem. Mol.*, **44**, 98–116.
14. Lin,Y., Guzman,C.E., McKinney,M.C., Nair,S.K., Ha,T. and Cann,I.K. (2006) *Methanosarcina acetivorans* flap endonuclease 1 activity is inhibited by a cognate single-stranded-DNA-binding protein. *J. Bacteriol.*, **188**, 6153–6167.
15. Lin,Y., Lin,L.J., Sriratana,P., Coleman,K., Ha,T., Spies,M. and Cann,I.K. (2008) Engineering of functional replication protein A homologs based on insights into the evolution of oligonucleotide/oligosaccharide-binding folds. *J. Bacteriol.*, **190**, 5766–5780.
16. Lin,Y., Robbins,J.B., Nyannor,E.K., Chen,Y.H. and Cann,I.K. (2005) A CCCH zinc finger conserved in a replication protein A homolog found in diverse Euryarchaeotes. *J. Bacteriol.*, **187**, 7881–7889.
17. Robbins,J.B., McKinney,M.C., Guzman,C.E., Sriratana,B., Fitz-Gibbon,S., Ha,T. and Cann,I.K.O. (2005) The euryarchaeota, Nature's medium for engineering of single-stranded DNA-binding proteins. *J. Biol. Chem.*, **280**, 15325–15339.
18. Robbins,J.B., Murphy,M.C., White,B.A., Mackie,R.I., Ha,T. and Cann,I.K.O. (2004) Functional analysis of multiple single-stranded DNA-binding proteins from *Methanosarcina acetivorans* and their effects on DNA synthesis by DNA polymerase BI. *J. Biol. Chem.*, **279**, 6315–6326.
19. Kelly,T.J., Simancek,P. and Brush,G.S. (1998) Identification and characterization of a single-stranded DNA-binding protein from the archaeon *Methanococcus jannaschii*. *Proc. Natl Acad. Sci. USA*, **95**, 14634–14639.
20. Pugh,R.A., Lin,Y., Eller,C., Leesley,H., Cann,I.K. and Spies,M. (2008) *Ferroplasma acidarmanus* RPA2 facilitates efficient unwinding of forked DNA substrates by monomers of FacXPD helicase. *J. Mol. Biol.*, **383**, 982–998.
21. Kerr,I.D., Wadsworth,R.I., Cubeddu,L., Blankenfeldt,W., Naismith,J.H. and White,M.F. (2003) Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of *Sulfolobus* SSB protein. *EMBO J.*, **22**, 2561–2570.
22. Leigh,J.A., Albers,S.V., Atomi,H. and Allers,T. (2011) Model organisms for genetics in the domain *Archaea*: methanogens, halophiles, *Thermococcales* and *Sulfolobales*. *FEMS Microbiol. Rev.*, **35**, 577–608.
23. Hartman,A.L., Norais,C., Badger,J.H., Delmas,S., Haldenby,S., Madupu,R., Robinson,J., Khouri,H., Ren,Q., Lowe,T.M. *et al.* (2010) The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. *PLoS One*, **5**, e9605.
24. Mevarech,M. and Werczberger,R. (1985) Genetic transfer in *Halobacterium volcanii*. *J. Bacteriol.*, **162**, 461–462.
25. Rosenshine,I., Tchelet,R. and Mevarech,M. (1989) The mechanism of DNA transfer in the mating system of an archaebacterium. *Science*, **245**, 1387–1389.
26. Zhao,A., Gray,F.C. and MacNeill,S.A. (2006) ATP- and $NAD^+$-dependent DNA ligases share an essential function in the halophilic archaeon *Haloferax volcanii*. *Mol. Microbiol.*, **59**, 743–752.
27. Allers,T., Ngo,H.P., Mevarech,M. and Lloyd,R.G. (2004) Development of additional selectable markers for the halophilic Archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. *Appl. Environ. Microb.*, **70**, 943–953.
28. Holzle,A., Fischer,S., Heyer,R., Schutz,S., Zacharias,M., Walther,P., Allers,T. and Marchfelder,A. (2008) Maturation of the 5S rRNA 5′ end is catalyzed *in vitro* by the endonuclease tRNase Z in the archaeon *H. volcanii*. *RNA*, **14**, 928–937.
29. Large,A., Stamme,C., Lange,C., Duan,Z.H., Allers,T., Soppa,J. and Lund,P.A. (2007) Characterization of a tightly controlled promoter of the halophilic archaeon *Haloferax volcanii* and its use in the analysis of the essential *cct1* gene. *Mol. Microbiol.*, **66**, 1092–1106.
30. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
31. Bitan-Banin,G., Ortenberg,R. and Mevarech,M. (2003) Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the *pyrE* gene. *J. Bacteriol.*, **185**, 772–778.
32. Gregor,D. and Pfeifer,F. (2005) *In vivo* analyses of constitutive and regulated promoters in halophilic archaea. *Microbiology*, **151**, 25–33.
33. DeVeaux,L.C., Muller,J.A., Smith,J., Petrisko,J., Wells,D.P. and DasSarma,S. (2007) Extremely radiation-resistant mutants of a halophilic archaeon with increased single-stranded DNA-binding protein (RPA) gene expression. *Radiat. Res.*, **168**, 507–514.
34. McCready,S., Jochen,A., Müller,J.A., Boubriak,I., Berquist,B.R., Loon,W. and DasSarma,S. (2005) UV irradiation induces homologous recombination genes in the model archaeon, *Halobacterium* sp. NRC-1. *Saline Syst.*, **1**, 3.
35. Bennett-Lovsey,R.M., Herbert,A.D., Sternberg,M.J. and Kelley,L.A. (2008) Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins*, **70**, 611–625.
36. Kelley,L.A. and Sternberg,M.J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.*, **4**, 363–371.
37. Wendoloski,D., Ferrer,C. and Dyall-Smith,M.L. (2001) A new simvastatin (mevinolin)-resistance marker from *Haloarcula hispanica* and a new *Haloferax volcanii* strain cured of plasmid pHV2. *Microbiology*, **147**, 959–964.