BMC Microbiology

Research article

Open Access

The single-stranded DNA-binding protein of *Deinococcus radiodurans*

Julie Malia Eggington, Nami Haruta, Elizabeth Anne Wood and Michael Matthew Cox*

Address: Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA

Email: Julie Malia Eggington - jeggington@biochem.wisc.edu; Nami Haruta - nharuta@biochem.wisc.edu; Elizabeth Anne Wood - ewood@biochem.wisc.edu; Michael Matthew Cox* - cox@biochem.wisc.edu

* Corresponding author

Published: 12 January 2004

BMC Microbiology 2004, **4**:2

Received: 17 October 2003

Accepted: 12 January 2004

This article is available from: http://www.biomedcentral.com/1471-2180/4/2

Abstract

Background: *Deinococcus radiodurans* R1 is one of the most radiation-resistant organisms known and is able to repair an unusually large amount of DNA damage without induced mutation. Single-stranded DNA-binding (SSB) protein is an essential protein in all organisms and is involved in DNA replication, recombination and repair. The published genomic sequence from *Deinococcus radiodurans* includes a putative single-stranded DNA-binding protein gene (*ssb;* DR0100) requiring a translational frameshift for synthesis of a complete SSB protein. The apparently tripartite gene has inspired considerable speculation in the literature about potentially novel frameshifting or RNA editing mechanisms. Immediately upstream of the *ssb* gene is another gene (DR0099) given an *ssb*-like annotation, but left unexplored.

Results: A segment of the *Deinococcus radiodurans* strain R1 genome encompassing the *ssb* gene has been re-sequenced, and two errors involving omitted guanine nucleotides have been documented. The corrected sequence incorporates both of the open reading frames designated DR0099 and DR0100 into one contiguous *ssb* open reading frame (ORF). The corrected gene requires no translational frameshifts and contains two predicted oligonucleotide/oligosaccharide-binding (OB) folds. The protein has been purified and its sequence is closely related to the *Thermus thermophilus* and *Thermus aquaticus* SSB proteins. Like the *Thermus* SSB proteins, the $\mathrm{SSB_{Dr}}$ functions as a homodimer. The *Deinococcus radiodurans* SSB homodimer stimulates *Deinococcus radiodurans* RecA protein and *Escherichia coli* RecA protein-promoted DNA three-strand exchange reactions with at least the same efficiency as the *Escherichia coli* SSB homotetramer.

Conclusions: The correct *Deinococcus radiodurans ssb* gene is a contiguous open reading frame that codes for the largest bacterial SSB monomer identified to date. The *Deinococcus radiodurans* SSB protein includes two OB folds per monomer and functions as a homodimer. The *Deinococcus radiodurans* SSB protein efficiently stimulates *Deinococcus radiodurans* RecA and also *Escherichia coli* RecA protein-promoted DNA strand exchange reactions. The identification and purification of *Deinococcus radiodurans* SSB protein not only allows for greater understanding of the SSB protein family but provides an essential yet previously missing player in the current efforts to understand the extraordinary DNA repair capacity of *Deinococcus radiodurans*.

Background

Deinococcus radiodurans R1, a tetrad-forming gram positive soil bacterium, is among the most radiation resistant organisms known [1]. The D_{37} γ irradiation dose, (the dose at which an irradiated population of cells is reduced to 37%) for *D. radiodurans* is approximately 6,000 Gy. This dose is 200 times that required to reduce *Escherichia coli* survival to the same extent [1]. A 6,000 Gy dose of radiation introduces approximately 300 DNA double strand breaks, greater than 3,000 single strand breaks, and more than 1,000 sites of base damage per *D. radiodurans* haploid genome ([1] and references therein). *D. radiodurans* is also able to sustain growth without induced mutation in the presence of considerable levels of ambient radiation (6 krads/h) [2]. Presumably, this organism possesses a robust DNA damage repair system. Single-stranded DNA-binding protein (SSB) is an essential protein in all known organisms and is required for DNA replication, recombination and repair [3].

The genome of *D. radiodurans* consists of 2 chromosomes, a megaplasmid, and a plasmid. Each of these genomic elements is present at 4–10 copies per cell [4]. The genome has been sequenced [4]. In this sequence, the SSB protein appeared to be encoded by a tripartite *ssb* gene, requiring two translational frameshifts or some other mechanism to allow expression of a full length functional SSB protein [2,4-6]. The apparent need for frameshifting in this gene has led to considerable speculation about both the presence of novel frameshifting or RNA editing mechanisms in *D. radiodurans* and their possible roles in normal metabolism [2,4-6]. There has even been some uncertainty as to whether *D. radiodurans* possesses a functional *ssb* gene. We now report that the *ssb* sequences as originally published (genes DR0099 and DR0100) include two errors that render the gene interpretation opaque. The correct sequence reveals a longer and completely contiguous ORF for *ssb*, encoding an SSB protein closely related to the recently characterized *Thermus thermophilus* and *Thermus aquaticus* SSB proteins [7]. The *D. radiodurans* SSB protein has been over-expressed and purified, and like the *Thermus* SSB proteins, has been shown to form a homodimer in solution. The *D. radiodurans* SSB homodimer stimulates both the *D. radiodurans* RecA protein and *E. coli* RecA protein-promoted DNA three-strand exchange reactions at a level comparable to or greater than the stimulation of these reactions by the *E. coli* SSB homotetramer.

Results and Discussion

The D. radiodurans *ssb gene encodes a contiguous ORF*

The region encompassing the *ssb* gene, which is flanked by DR0098, the ribosomal protein S6 gene and DR0101, the ribosomal protein S18 gene, was PCR amplified from genomic DNA and sequenced (Fig. 1). The genomic DNA was derived from the *Deinococcus radiodurans* R1 type strain obtained directly from the ATCC (ATCC 13939). The *ssb* gene was also independently cloned into two different cloning vectors and re-sequenced. In each case, the sequence revealed two additional guanine nucleotides that are not present in the sequence deposited in the database of the National Center for Biotechnology Information (NCBI) (Fig. 1) [4]. (The original submitted sequence has accession #AE000513. The current sequence has accession #NC_001263 and is identical to the original). With the corrections reported here, a contiguous ORF that encodes a complete *ssb* protein is revealed. A recognizable *D. radiodurans* ribosomal binding site (AAGGAG) exists at an optimal 8 nucleotides upstream of the initiation codon (Fig. 1) [8-10]. The corrected *ssb* sequence has been deposited with GenBank, accession number AY293617. The predicted *D. radiodurans* SSB protein is the largest bacterial SSB polypeptide identified to date (301 amino acids, including the initiating methionine), exceeding in length the *T. thermophilus* VK1 SSB by 35 amino acid residues.

The presence of a contiguous open reading frame encompassing the regions designated DR0099 and DR0100 was confirmed by PCR amplification of this region from genomic DNA, and cloning it into an *E. coli* expression vector. Excellent inducible expression of a protein of the predicted size was obtained from the pEAW328 vector (see Methods) (Fig. 2). This protein was purified and confirmed to be *D. radiodurans* SSB protein by N-terminal sequencing of the first nine amino acids (ARGMNHVYL). As expected, the N-terminal methionine residue was absent in the *D. radiodurans* SSB protein purified from *E. coli*. The calculated molecular weight of the purified *D. radiodurans* SSB protein monomer (300 amino acids without the initiating methionine) is 32,591. We estimate that the *D. radiodurans* SSB protein purified as described here is more than 99% homogeneous. Additionally, the purified *D. radiodurans* SSB protein was free of detectable DNA endo- and exonucleases.

D. radiodurans *SSB protein is closely related to the* T. thermophilus *and* T. aquaticus *SSB proteins*

The protein sequence predicted for *D. radiodurans* SSB protein shares 43% identity and 58% similarity with the *T. thermophilus* HB-8 SSB protein, 43% identity and 58% similarity with the *T. thermophilus* VK-1 SSB protein and 44% identity and 61% similarity with the *T. aquaticus* YT-1 SSB protein (Fig 3A and 3B). The N-terminal segment of the *D. radiodurans* SSB protein shares 38% identity and 49% similarity with the *E. coli* SSB protein. The C-terminal segment of the *D. radiodurans* SSB protein shares 39% identity and 64% similarity with the *E. coli* SSB protein.

Virtually all bacterial SSB proteins identified to date contain only one oligonucleotide/oligosaccharide-binding

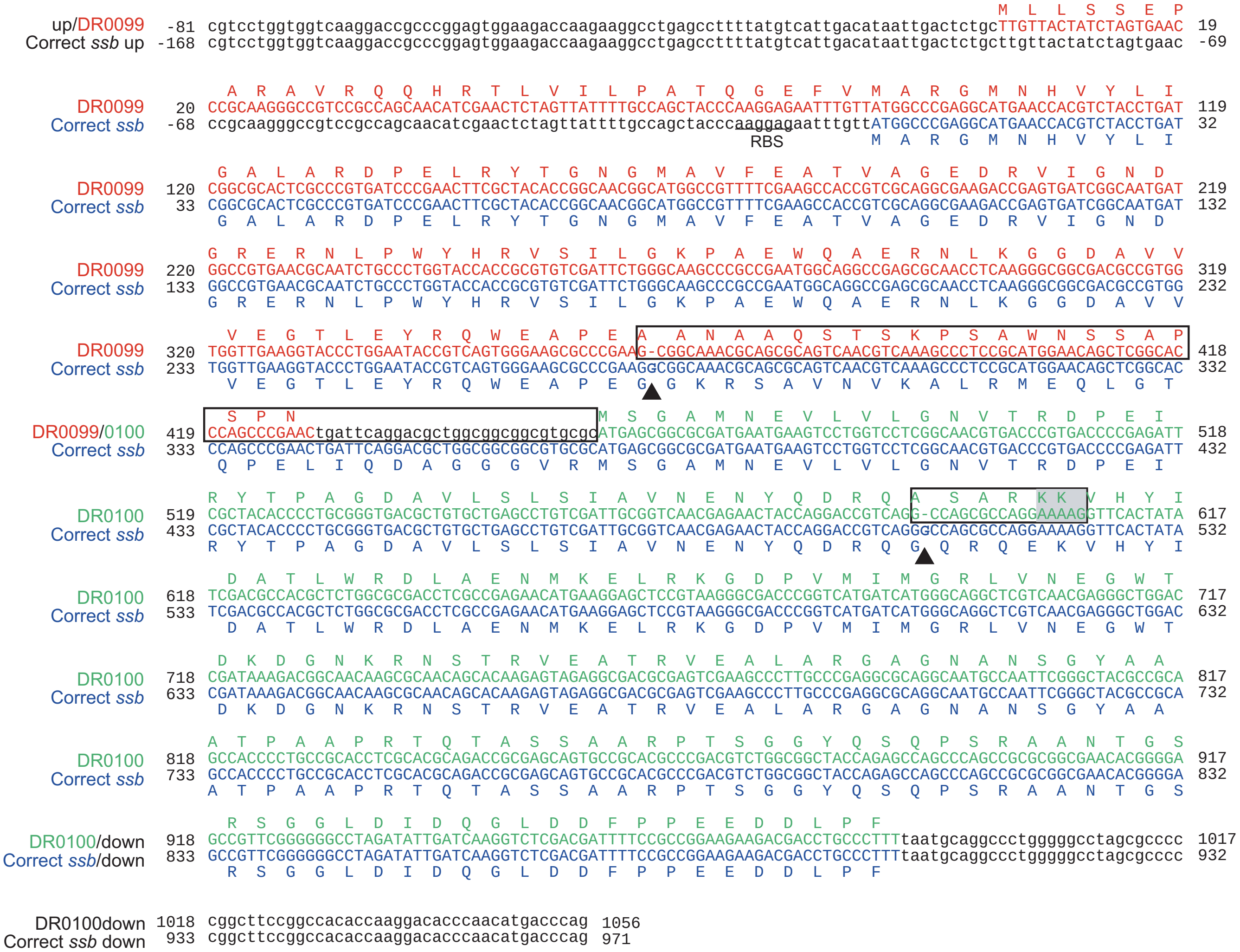

Figure 1

Nucleotide sequence of the *D. radiodurans* R1 *ssb* gene. The predicted amino acid sequence is shown above the nucleotide sequence for the DR0099 (red) and DR0100 (green)[4] and below for the corrected *ssb* (blue) sequences. Predicted non-coding regions are shown in lower case. The nucleotides omitted in the published sequence are shown in bold in the corrected *ssb* sequence and are indicated by triangles. The reading frames affected by the previous errors are shown in boxes. The translational frameshift region predicted by the earlier sequence is highlighted by a gray box. The putative Ribosomal Binding Site in the corrected *ssb* sequence is underlined. The sequences are numbered according to the first predicted initiation codon in the sequences (TTG in DR0099 and ATG in the corrected *ssb*). The DR0099 and DR0100 genes have accession number NC_001263. The corrected *ssb* sequence, reported here, has accession number AY293617.

(OB) fold per monomer and function as homotetramers [11,12]. The very recently characterized *T. thermophilus* and *T. aquaticus* SSB proteins have broken that pattern, and contain 2 OB folds per monomer [7,11] and function as homodimers [7]. In the formation of homodimers, *T. thermophilus* and *T. aquaticus* SSB proteins maintain the bacterial trend of 4 OB folds per SSB protein oligomer. The *D. radiodurans* SSB protein gene structure is quite similar to that of the *T. thermophilus* and *T. aquaticus ssb* gene structures and is predicted to contain two OB folds. The high similarity of the *D. radiodurans* SSB protein sequence to the *Thermus* SSB proteins (Fig 3) suggested that the *D. radiodurans* SSB protein would also form a homodimer. This prediction was confirmed (see below). The structure of the *D. radiodurans* SSB protein is consistent with and reinforces the close phylogenetic relationship between

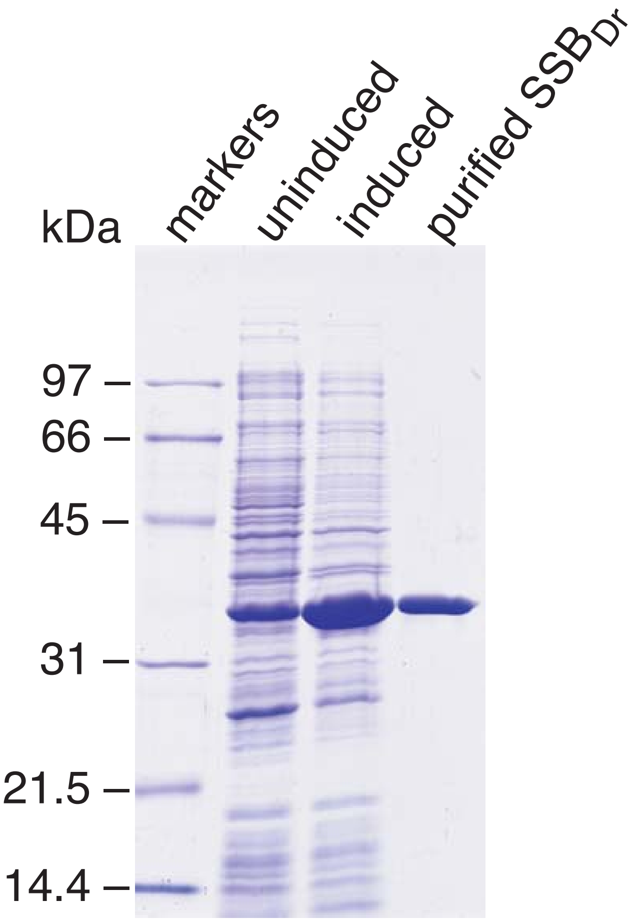

Figure 2
Expression and purification of the *D. radiodurans* R1 SSB protein. Protein expression was obtained from the pEAW328 vector in BL21 Codon Plus (DE3) (Stratagene) *E. coli* cells. Proteins were examined on a standard SDS-polyacrylamide gel. Lanes are (1) molecular weight markers (Bio-Rad, Low Range), as noted with the labels at left; (2) Whole cell extract of uninduced cells; (3) whole cell extract of cells after induction of the *D. radiodurans ssb* gene; (4) purified *D. radiodurans* R1 SSB protein.

Deinococcus and bacteria of the *Thermus* group of extremophiles [2,4-6].

D. radiodurans *SSB protein forms a homodimer*

Native molecular weight approximation by gel filtration analysis (Fig 4) and sedimentation equilibrium experiments confirmed our prediction that *D. radiodurans* SSB protein exists as a homodimer. In the gel filtration analysis, *D. radiodurans* SSB protein displayed a fractional retention (K_{av}) on a Sephacryl S-200 column of 0.21. *E. coli* SSB protein, ran for comparison, displayed a K_{av} of 0.18. Using a standard molecular weight curve generated for the column, these values correspond to 85 kDa and 101 kDa respectively. The native molecular weight approximation for *D. radiodurans* SSB protein is 2.6 times the molecular weight of a *D. radiodurans* SSB monomer (32.6 kDa), or 1.3 times that predicted for a *D. radiodurans* SSB homodimer (65 kDa). The native molecular weight approximation for *E. coli* SSB protein under these conditions is 5.4 times that of the molecular weight of an *E. coli* SSB monomer (18.8 kDa), or 1.3 times that of an *E. coli* SSB homotetramer (75 kDa). *E. coli* SSB protein is known to form a homotetramer in solution [13,14]. Given the aberrant chromatographic behavior of the native *E. coli* SSB homotetramer, the apparently similar behavior of the native *D. radiodurans* SSB is most consistent with a homodimer. Due to the high sequence identity and similarity of the OB folds between *E. coli* SSB protein and *D. radiodurans* SSB protein, we anticipate that the proteins will have similar properties. In the experiment of Fig. 4, both proteins chromatograph with an apparent mass corresponding to 1.3 times that of their calculated molecular weight, assuming the *D. radiodurans* SSB is a homodimer and the *E. coli* SSB is a homotetramer.

The sedimentation equilibrium data were best described as a single species. There was no evidence for multiple species and modeling attempts using associations or simple mixtures did not yield physically realistic parameters for some fitting variables. The results of global fitting of the all data at absorbance less than 2 gave a molecular weight of 63,300 ± 70, in excellent agreement with the value of 65,183 for a dimer based on the sequence. In fact an increase in the calculated partial specific volume of only 0.008 would yield perfect agreement.

Since sedimentation equilibrium shows the protein to be a homogeneous population of dimers, the results from column chromatography would suggest that the shape of the molecule deviates from that of the globular standards.

The C-terminus of SSB has special functional significance. Although the length and the sequence of the C-terminal regions (the region extending past the OB fold(s)) is variable across bacterial species, the last 10 amino acids at the C-terminus are highly acidic and well conserved across bacterial species [3]. An acidic C-terminus is also seen in the *D. radiodurans* SSB protein C-terminal tail (Fig 3). Studies of *E. coli* SSB C-terminal truncation mutants have indicated that the last 10 amino acids of *E. coli* SSB protein is essential for cell survival and is probably involved in protein-protein interactions [15]. This may be of functional significance to the *D. radiodurans* SSB. As a homodimer, SSB_{Dr} contains 4 OB folds, similar to the *E. coli* SSB homotetramer. However, the same *D. radiodurans*

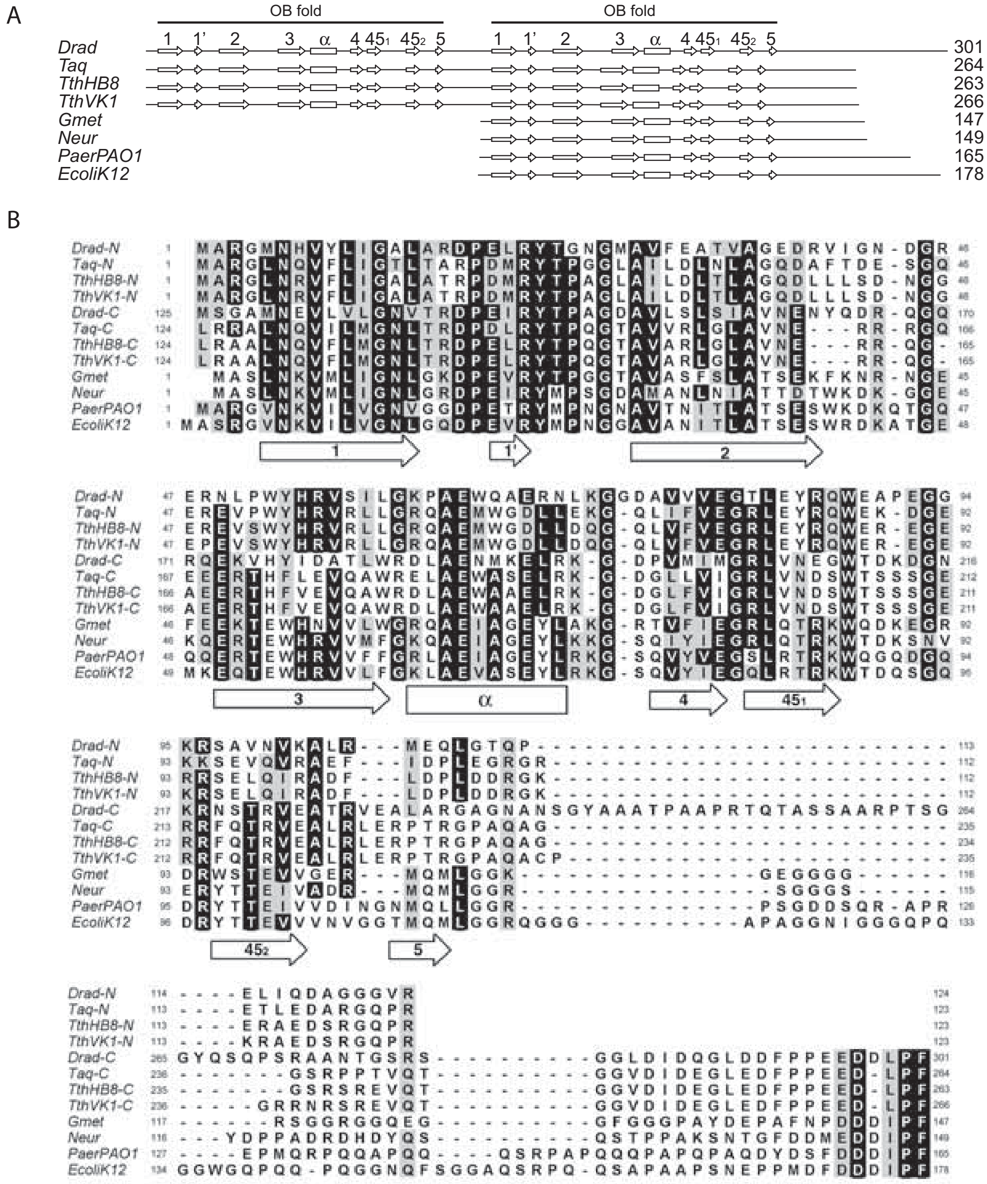

Figure 3

Protein schematics and amino acid sequence alignments of *D. radiodurans* R1 SSB protein and other SSB proteins. A) Schematic representation of the *D. radiodurans* R1 SSB protein and other SSB proteins highlighting the OB fold regions. The *Drad*, *Taq*, *TthHB8* and *TthVK1* SSB proteins contain two OB folds each. The characteristic motifs that make up an OB fold are highlighted with open boxes/arrows and numbered. The *Gmet*, *Neur*, *PaerPAO1* and *EcoliK12* SSB proteins contain one OB fold each and align with both the N-terminal OB fold and the C-terminal OB fold of the proteins that contain two OB folds. Only the C-terminal alignments are shown. The structural assignments are according to the OB fold rule defined by Murzin [11,32]. B) Amino acid sequence alignment of *D. radiodurans* R1 SSB protein and other SSB proteins from closely and distantly related bacteria. The *Drad*, *Taq*, *TthHB8 and TthVK1* SSB protein sequences are divided into N- and C-terminal fragments in order to highlight that each fragment contains an OB fold. The arrows indicate β-sheets and the rectangle indicates the α-helix contained in the secondary structure of the OB fold of SSB protein from *E. coli* strain K12 [11,32]. Sequence identity is shown by white fonts on black boxes, and sequence similarity is shown by black font on gray boxes. Sequence similarity is defined by the following amino acid groupings: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW. Abbreviations: *Drad-N* or *-C*: *D. radiodurans* strain R1, N- or C- terminal fragment; *Taq-N* or *-C*: *T. aquaticus*, N- or C- terminal fragment; *TthHB8-N* or *-C*: *T. thermophilus* strain HB8, N- or C- terminal fragment; *TthVK1-N* or *-C*: *T. thermophilus* strain VK1, N- or C-terminal fragment; *Gmet*: *Geobacter metallireducens*; *Neur*: *Nitrosomonas europaea* ATCC 19718; *PaerPAO1*: *Pseudomonas aeruginosa* PAO1. *EcoliK12*: *E. coli* strain K12.

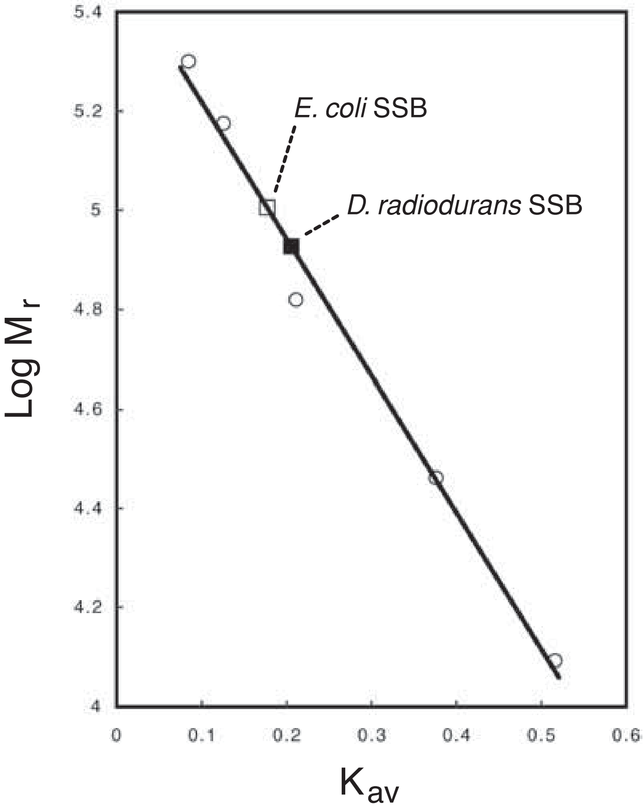

Figure 4

Molecular weight approximation of the *D. radiodurans* SSB protein oligomer. Gel filtration of the standards and SSB proteins were performed as described in the Methods section using a S-200 Sephacryl HR column. The protein standards β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome c (12.4 kDa) were used to calibrate the column as shown by open circles. The elution volume of blue dextran was used to calculate the void volume and the total volume of the column was known. The best-fit line was generated of the log M_r of the protein standards versus K_{av} of the standards. *D. radiodurans* SSB protein and *E. coli* SSB protein samples were injected onto the column in independent experiments and their elution volumes were used to calculate their K_{av}. The position of *D. radiodurans* SSB protein on the standard curve is indicated by a closed square and the position of *E. coli* SSB protein on the standard curve is indicated by an open square.

SSB homodimer contains only two C-terminal regions whereas the *E. coli* SSB homotetramer contains four.

In vitro analyses of the *E. coli* SSB protein showed that the C-terminal third of the protein is not needed for tetramer formation or DNA binding and that the last 10 amino acids, containing the highly conserved acidic region, actually weakens the binding of *E. coli* SSB homotetramer to nucleic acids [15]. The authors of this study conclude that the amino acid sequence contained between the OB fold and the highly acidic 10 amino acid tail acts purely as a spacer to shield the negative charges of the C-terminus from the bound DNA on the SSB protein [15]. In support of this idea, a crystal structure of *E. coli* SSB protein, in the absence of DNA, suggests that the C-terminal region (the region not involved in the OB fold) protrudes from the tetramer and that each of the four C-terminal regions take on different conformations [16].

We speculate that the acidic terminus of the *D. radiodurans* SSB protein is also important for *in vivo* protein-protein interactions, but that there may be some interesting functional differences reflecting the reduced number of C-terminal regions in the *D. radiodurans* SSB oligomer compared to that of *E. coli* and other bacterial SSB oligomers. Further studies involving the *D. radiodurans* SSB C-terminus will help to further understand the role of the spacer and acidic tail regions of bacterial SSB proteins.

A D. radiodurans *SSB protein homodimer facilitates RecA promoted DNA three-strand exchange*

Single-stranded DNA binding proteins, in general, help stimulate recombinase promoted *in vitro* DNA strand exchange reactions. The *E. coli* system has been most intensely studied and *E. coli* SSB protein has been shown to have both pre-synaptic and post-synaptic roles in *E. coli* RecA protein-promoted DNA strand exchange reactions [17,18]. The addition of *E. coli* SSB during the presynaptic formation of RecA filaments helps to remove ssDNA secondary structure, allowing formation of contiguous RecA filaments [17]. Post-synaptically, *E. coli* SSB facilitates the recombination reaction by binding the displaced ssDNA and preventing the reversal of the strand exchange reaction [18].

The SSB proteins from different organisms have frequently been interchanged and found to stimulate the reactions of recombinases from other organisms [19-21]. These observations suggest that there is little species-specific protein-protein interaction. The *E. coli* SSB protein greatly stimulates the *D. radiodurans* RecA protein-promoted DNA strand exchange reaction [22], and in fact an SSB protein is required to observe significant reaction. We have now found that the *D. radiodurans* SSB protein stimulates the DNA strand exchange reactions promoted by both $RecA_{Dr}$ and the *E. coli* $RecA_{Ec}$ (Fig. 5). In the reactions promoted by both *D. radiodurans* RecA protein (Fig 5 Panels A, B and C) and the *E. coli* RecA protein (Figure 5 Panel D), less than half as much *D. radiodurans* SSB protein was needed to generate a yield of nicked circular product equal to that seen in the reactions with *E. coli* SSB under the same conditions. The reduced requirement for the *D.*

radiodurans SSB is again consistent with the formation of a *D. radiodurans* SSB homodimer (Fig 5). Since the *D. radiodurans* SSB forms a homodimer (with 4 total OB folds) and *E. coli* SSB forms a homotetramer (with 4 total OB folds), the concentration of SSB_{Dr} should be multiplied by a factor of 2 to provide an SSB_{Dr}-SSB_{Ec} comparison that reflects equal numbers of OB folds (see dashed line in Fig. 5, panels C and D).

Note that the maximum product generation is slightly different for the *D. radiodurans* and *E. coli* RecA protein-promoted reactions. The optimal conditions are slightly different for the two recombinases, and the conditions used in these reactions were those optimal for the *E. coli* RecA protein.

Conclusions

We conclude that frameshifting does not occur within the *D. radiodurans ssb* gene. The gene consists of a contiguous open reading frame encompassing the regions now labeled DR0099 and DR0100. The predicted amino acid sequence of the *D. radiodurans* SSB protein is closely related to the SSB proteins from *T. aquaticus* and *T. thermophilus*. The purified *D. radiodurans* SSB protein, like the SSB proteins from *T. aquaticus* and *T. thermophilus* [7], forms a homodimer. The *D. radiodurans* SSB homodimer stimulates *D. radiodurans* RecA and also *E. coli* RecA protein-promoted DNA strand exchange reactions with at least the same efficiency as the *E. coli* SSB homotetramer stimulates these reactions. Elucidation of the complete *D. radiodurans ssb* gene and purification of the *D. radiodurans* SSB protein opens the way for further studies of the DNA metabolism of this unusually DNA repair-proficient organism.

Methods

Strain

The *Deinococcus radiodurans* R1 type strain was obtained directly from the American Type Culture Collection (ATCC 13939).

Genomic DNA isolation

Genomic DNA was purified from *Deinococcus radiodurans* R1 type strain cells using the following protocol. TGY media (5% tryptone, 3% yeast extract, 1% glucose) was used in a 5 mL overnight 30°C culture growth. The cells were pelleted and resuspended in 2 mL suspension buffer (10 mM Tris-HCl (pH 8.0), 1 mM Na-EDTA, 0.35 M sucrose). Lysozyme (2 mg) was added and the suspension was incubated for 4 hours at 37°C. Lysis buffer (3 ml; 100 mM Tris-HCl (pH 8.0), 0.3 M NaCl, 20 mM EDTA, and 2% (w/v) SDS, with 2% (v/v) β-mercaptoethanol added just before use) was added to lyse the cells. The suspension was extracted with equal volumes of TE saturated phenol once, then once with 25:24:1 phenol/chloroform/isoamyl alcohol and then once with 24:1 chloroform/isoamyl alcohol. The non-aqueous layer was removed, and the DNA was ethanol precipitated. The dry DNA pellet was dissolved in 200 μl TE, and 1 μl was used in subsequent PCR amplifications.

Sequencing approaches

The sequence of the *ssb* gene was checked in three independent analyses using different PCR primers and sequencing primers. In a direct approach, primers starting in the flanking DR0098 ribosomal protein S6 gene (5' CATCAAGGCTTCGGGCAAC 3') and the DR0101 ribosomal protein S18 gene (5' TTGCGTTCGCCGCTGTTTCC 3') were used to PCR amplify the *ssb* gene and flanking regions. Sequencing primers were used to directly sequence this PCR fragment with a 2–4-fold coverage. In addition to this direct approach, the *ssb* gene was independently PCR amplified and inserted into two separate plasmids, pRAD1 (obtained as a gift from Mary E. Lidstrom) [9] and pET21A (Novagen) and sequenced from these plasmids. The first clone, with the *ssb* gene inserted between the *Xho*I and *Hin*dIII cloning sites of vector pRAD1, used the PCR forward primer containing the *Xho*I site (underlined) (5' CTACCGCTCGAGAGTGGAAGACCAAGAAGGCCTGAGC 3') and the reverse primer containing the *Hin*dIII site (5' TATGAAGCTTTTAGTGGTGGTGGTGATGATGGCCACCAAAGGGCAGGTCG 3'). The forward primer for the pRAD1 insertion PCR was designed to anneal in the upstream flanking ribosomal protein S6 gene while the reverse primer was designed to anneal at the end of the *ssb* gene and included a C-terminal histidine tag sequence. The *ssb* gene was sequenced in this clone with a two-fold coverage. This clone was not used for any further experimentation contained within this report. A second, independently created clone called pEAW328, with the *ssb* gene inserted between the *Eco*RI and the *Nde*I cloning sites of pET21A, used the PCR forward primer containing the *Nde*I site (5' CGTATTCCATATGGCC*CGT*GGCATGAACCAC 3') and the reverse primer containing the *Eco*RI site (5' CGGAATTCTTAAAAGGGCAGGTCGTC 3'). The forward primer for the pET21A insertion PCR was designed to anneal at the initiating codon of the correct *ssb* gene and contained a CGT codon (italicized) rather than the natural CGA codon for Arg in order to replace a low use Arg codon (the third amino acid in the primary sequence) with a higher use Arg codon in preparation for overexpression. The reverse primer for the pET21 insertion PCR was designed to anneal at the natural termination codon of the correct *ssb* gene. No tag sequence was used in the pET21A vector construction. Three separate pET21A insertion clones were sequenced twice. All restriction enzymes were purchased from New England Biolabs.

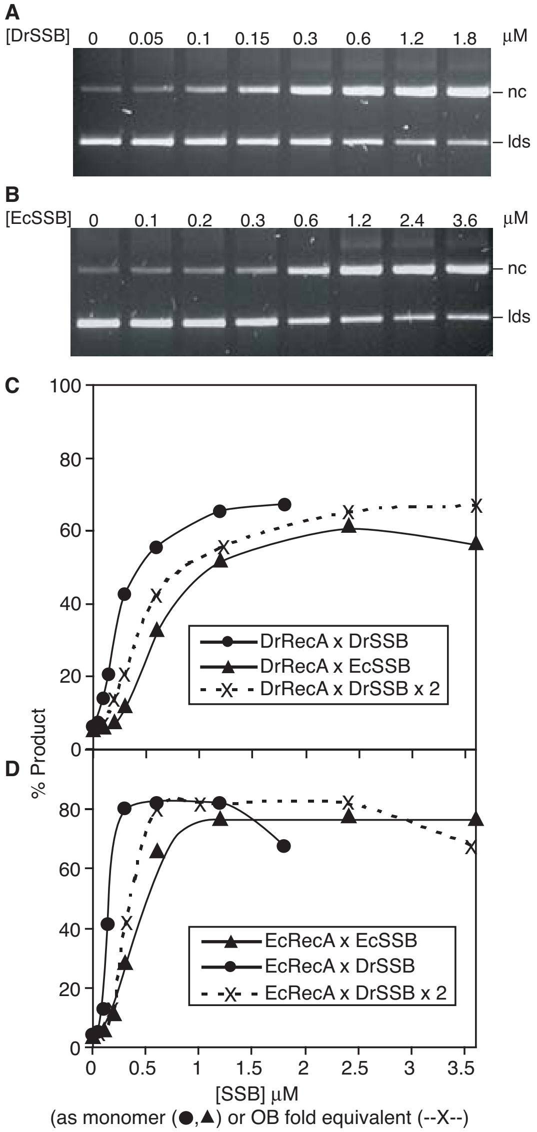

Figure 5

DNA strand exchange reactions promoted by *D. radiodurans* RecA and *E. coli* RecA with SSB titrations. Reactions were carried out as described in the Methods section. Circular single-stranded DNA (css) was preincubated with either *D. radiodurans* or *E. coli* RecA. ATP and SSB protein (either *D. radiodurans* or *E. coli* SSB protein as indicated) were then added and incubated, followed by the addition of homologous linear double-stranded DNA (lds) which initiated the DNA three-strand exchange reaction. The nicked circular double-stranded DNA product (nc) is distinguishable by agarose gel electrophoresis and quantifiable. Panels A and B show the agarose gel electrophoresis results of reactions promoted by *D. radiodurans* RecA with various monomer concentrations of *D. radiodurans* SSB protein and *E. coli* SSB protein, respectively. These results are quantitated in panel C, with the data from reactions with *D. radiodurans* SSB (closed circles) and *E. coli* SSB (closed triangles) coming from Panels A and B respectively. Panel D shows the quantitated results of similar reactions promoted by *E. coli* RecA with *D. radiodurans* SSB (closed circle) and *E. coli* SSB (closed triangle) (agarose gel not shown). Since an SSB_{Dr} monomer has two OB folds and an SSBEc monomer has only one, the improved reactions seen in the reactions containing the former protein could simply reflect the higher effective concentration of OB folds when monomeric SSB concentrations are compared. In panels C and D, the dashed lines represent a plot in which the percentage reaction product generated in the reactions using SSB_{Dr} are plotted against the actual concentration of OB folds in these reactions (twice the actual concentration of *D. radiodurans* SSB monomers). The dashed lines (--X--) thus allow a direct comparison of the reactions observed with the *D. radiodurans* SSB protein with the reactions observed with the *E. coli* SSB protein (closed triangles). The production of nicked circular double-stranded DNA product is calculated as a percentage of total duplex DNA (the sum of linear double-stranded DNA substrate, nicked circular double-stranded DNA product and any network products near the well).

Protein sequence analysis

Standard BLAST pair wise analysis (Blosum62 matrix at default settings) of the protein sequences was used to calculate percent identity and similarity values between proteins. ClustalX was used to generate the multiple sequence alignments.

Expression and purification of D. radiodurans *SSB Protein*

As described above, the *D. radiodurans ssb* gene PCR product was inserted into the *EcoRI* and the *NdeI* cloning sites of pET21A (Novagen) to yield construct pEAW328. The *ssb* gene in this construct did not contain a histidine tag or any modification that would lead to a translation product that would differ from that encoded by the chromosomal *D. radiodurans ssb* gene. Construct pEAW328 was transformed into BL21 Codon Plus (DE3) (Stratagene) *E. coli* cells. These cells were grown in LB broth in the presence of 100 μg/ml of ampicillin and 25 μg/ml of chloramphenicol at 35°C to an optical density at 600 nm of 0.8. The cells were then induced with 0.4 mM IPTG and grown at 35°C for three more hours before harvest. Harvested cells were resuspended in Buffer A (25 mM Tris-HCl, pH 8.3, 12% w/v glycerol, 0.5 mM EDTA) with the buffer addition corresponding to five times the cell volume. Lysozyme was added to a final concentration of 0.2 mg/ml. Cells were stirred at 4°C for 1 hour and then sonicated on ice. All subsequent purification steps were performed at 4°C. Cell debris and insoluble material were removed by 3 successive 20 min centrifugations at 38,000 g with insoluble material being removed between centrifugations. Solid NaCl was then dissolved in the supernatant to a concentration of 0.18 M NaCl. DNA and proteins were precipitated by drop-by-drop addition of 10% (w/v) polyethylenimine, pH 7.5, with constant stirring to a final concentration of 0.4% (w/v) polyethylenimine. The solution was stirred for 15–60 minutes and then centrifuged for 15 min at 10,000 g. The SSB protein remained in the pellet at this point and was eluted from the pellet with Buffer B (25 mM Tris-HCl, pH 8.3, 0.4 M NaCl, 12% w/v glycerol, 0.5 mM EDTA, 1 mM β-mercaptoethanol) in a volume equal to the volume initially used to resuspend the cells. The polyethylenimine pellet was broken up and resuspended using a plastic spatula and then a glass homogenizer or small mortar and pestle. This suspension was stirred for 30 min and then centrifuged for 15 min at 10,000 g. The SSB was found in the supernatant. Solid Ammonium Sulfate was slowly added to the supernatant while stirring over the course of 30 minutes to a 30% saturation of the solution. This solution was allowed to continue stirring for approximately 3 hours and the precipitated proteins, including much of the SSB, were collected by centrifugation at 10,000 g for 20 min. The Ammonium Sulfate pellet was dissolved in TGE Buffer (50 mM Tris-HCl, pH 8.3, 20% w/v glycerol, 1 mM EDTA, 1 mM β-mercaptoethanol) with gentle stirring at a volume ~65% of the initial volume used to resuspend the cells. The resuspended solution was centrifuged for 20 min at 30,000 g to remove non-dissolved proteins. The supernatant was dialyzed against TGE Buffer. Following dialysis, the solution was cleared of precipitate by a 5 min centrifugation at 3,800 g and the supernatant was dialyzed against P (20 mM) buffer (20 mM potassium phosphate, pH 7.5, 10% (w/v) glycerol, 0.1 mM EDTA, 1 mM β-mercaptoethanol). Again the solution was cleared by centrifugation following dialysis. A hydroxyapatite column (BioRad) was pre-equilibrated with P (20 mM). EDTA was not used in buffers applied to the column other than the EDTA contained in the load volume. The hydroxyapatite bed volume used and found adequate was 110% of the initial volume used to resuspend the cells. The dialyzed protein was loaded on the column and eluted over a 10 column volume linear gradient going from Buffer P (20 mM) to Buffer P (152 mM, i.e. 152 mM potassium phosphate, pH 7.5, 10% (w/v) glycerol, 1 mM β-mercaptoethanol). The SSB completed elution approximately half way through the linear gradient. Fractions containing mostly SSB and some very minor degradation bands by SDS-PAGE were pooled and dialyzed against TGE Buffer. The dialyzed protein was then loaded onto a DEAE Sepharose (Amersham Pharmacia Biotech) column (bed volume of 110% of the initial volume used to resuspend the cells was found adequate) and eluted in a 10 column volume linear gradient of TGE Buffer going from 0.0 M NaCl to 0.3 M NaCl. Degraded protein eluted before the pure SSB protein eluted. Elution was complete at about half way through the gradient. Fractions containing pure SSB protein were pooled and dialyzed into storage buffer (20 mM Tris-HCl, pH 8.3, 0.5 M NaCl, 50% (w/v) glycerol, 1 mM EDTA, 1 mM β-mercaptoethanol), aliquotted, snap frozen in liquid nitrogen, and stored at -80°C.

The extinction coefficient for the *D. radiodurans* SSB protein was determined as described previously [23,24]. Eleven determinations at three different concentrations of *D. radiodurans* SSB protein gave an average extinction coefficient of $\varepsilon_{280} = (4.1 \pm 0.2) \times 10^4\ M^{-1}cm^{-1}$. The N-terminal sequence analysis was performed by the Protein and Nucleic Acid Chemistry Laboratories, Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, Mo., without the laboratory personnel knowing the identity of the protein.

Native Molecular Weight Determination by Gel Filtration

The native molecular weight of *D. radiodurans* SSB was approximated by gel filtration FPLC. A Sephacryl S-200 HR column (31.4 cm × 1.0 cm) was used at a flow rate of 0.05 ml/min. The buffer in all experiments was 50 mM Tris-HCl, pH 7.5, 100 mM KCl, as recommended by Sigma for the protein standards. Chromatography was performed at 4°C while A_{280} was measured. The column was calibrated using Sigma Gel Filtration Molecular Weight Markers: blue dextran (2,000 kDa), beta-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome c (12.4 kDa). The standards were loaded independently at the concentrations recommended by Sigma in 100 μl sample volumes. Approximately 70 μg of *E. coli* SSB protein (18.8 kDa per monomer) and 200 μg of *D. radiodurans* SSB protein (32.6 kDa per monomer) were independently loaded on the column in 100 μl sample volumes. Standards and SSB loads were dissolved or dialyzed respectively into the recommended buffer plus 5% (w/v) glycerol. The elution volume of blue dextran determined the void volume (V_o), and the total volume (V_t) was determined by volume measurements of the column before packing and measurement of the column tubing used. The peak elution volumes (V_e) were calculated from the chromatogram and fractional retentions, K_{av}, were calculated using the equation: $K_{av} = (V_e - V_o)/(V_t - V_o)$. A standard curve was determined by plotting the K_{av} of the protein standards against the $\log_{10}M_r$ of the standards. The native molecular weight of *D. radiodurans* SSB protein was approximated by comparing its K_{av} value to the standard curve. The native molecular weight of *E. coli* SSB protein was determined for comparison in like manner. *E. coli* SSB protein was purified as previously described [25], and an extinction coefficient of 2.38×10^4 $M^{-1}cm^{-1}$ was used to calculate the concentration of *E. coli* SSB.

Sedimentation Equilibrium Measurements

To prepare samples for sedimentation equilibrium a 1 mL *D. radiodurans* SSB 157 μM protein sample (in 35 mM Tris-HCl (pH 7.7), 25% w/v glycerol, 300 mM NaCl, 1 mM EDTA, 0.5 mM β-mercaptoethanol, 1 mM EDTA buffer) was dialyzed for 1 hour at 4°C against 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA dialysis buffer, and then again against 2 L of fresh dialysis buffer overnight at 4°C. The protein concentration of the resulting sample was 80.4 μM. The final dialysate was used to dilute aliquots of this stock to concentrations of 4.67 μM, 12.4 μM and 20.4 μM.

100 μL of each sample was placed in 12 mm double-sector charcoal-filled Epon centerpieces with about 105 μL of the final dialysate as reference. Centrifugation was performed at 4°C in a Beckman model XL-A analytical ultracentrifuge. The protein gradients were recorded at 280 nm every 2–3 hours until they became superimposable. Data were collected at six speeds (5200, 8200, 11000, 14000, 16500, 18500 rpm). With the various speeds and starting concentrations, absorbance in the gradients ranged from ~0 to >2.8, which corresponds to ~0 to >60 μM protein. Baseline absorbances were measured for each sample after high speed depletion and was less than 0.03 in all cases and was subtracted prior to curvefittings. The partial specific volume was calculated from the composition as 0.722 mL/gm. The dialysate density was measured as 1.007518 gm/mL at 4°C using an Anton Paar DMA5000 density meter.

The data from the three samples at six speeds were globally tested against models of a single species, two noninteracting species and two species in equilibrium. Programs for analysis were written in Igor Pro (Wavemetrics Inc., Lake Oswego, OR) by Darrell R. McCaslin.

DNA Three-Strand Exchange Reactions

Enzymes and Reagents

The *E. coli* RecA and *D. radiodurans* RecA proteins were purified by polyethylenimine precipitation followed by a DEAE-Sepharose column and a hydroxyapatite column as described [26]. Protein concentrations were determined by absorbance at 280 nm using the extinction coefficients $\varepsilon_{280} = 2.23 \times 10^4$ M^{-1} cm^{-1} for *E. coli* RecA [27], 1.41×10^4 M^{-1} cm^{-1} for *D. radiodurans* RecA [22] and 2.38×10^4 $M^{-1}cm^{-1}$ for *E. coli* SSB [28]. *E. coli* SSB protein was purified as previously described [25].

Preparation of DNA substrates

Duplex supercoiled DNA and circular ssDNA (css DNA) substrates from bacteriophage M13mp7 (7238 bp) were purified as described [29-31]. The linear dsDNA substrate was prepared from M13mp7 supercoiled dsDNA, which was cut with *BsmBI* and purified by electrophoresis on 1.0% agarose gel. The concentrations of ssDNA and dsDNA solutions were determined by absorbance at 260 nm, using 36 and 50 μg ml^{-1} A_{260}^{-1}, respectively, as conversion factors. The concentrations of DNA and proteins reported below are the final concentrations after addition of all components and DNA concentrations are in terms of total nucleotides.

DNA three-strand exchange reaction

The DNA three-strand exchange reactions were carried out at 37°C in solutions containing 25 mM Tris-OAc (80% H^+, pH 7.5), 10 mM $Mg(OAc)_2$ (Fisher Scientific), 1 mM DTT (Research Organics), 3 mM Potassium Glutamate, 5% (w/v) glycerol and an ATP regeneration system (12 mM phosphocreatine and 10 units/ml phosphocreatine kinase (Boehringer Mannheim)). Two μM RecA protein (from *E. coli* or *D. radiodurans*) was pre-incubated with 6 μM css DNA in the reaction buffer and regeneration system for 10 min. ATP (3 mM) and SSB (indicated concentrations) were then added, followed by another 20 min incubation. The reactions were initiated by addition of 10 μM lds DNA. The reactions were incubated for 2 hr and stopped by the addition of 1.2 μl 10% SDS, 0.3 μl 0.5 M EDTA and 0.6 μl 20 mg/ml Proteinase K followed by 30 min incubation. Aliquots mixed with 2.5 μl 6 × loading buffer (15% Ficoll, 0.25% bromphenol blue, 0.25%

xylene cyanole FF) were loaded on an 1% agarose gel and electrophoresed at 25–35 V for 16 hr at room temperature. To visualize the DNA bands, the gels were stained with ethidium bromide, and exposed to UV light. Gel images were captured with a digital CCD camera utilizing GelExpert software (Nucleotech). The intensity of DNA bands was quantitated with the software package TotalLab v1.10 from Phoretix.

List of abbreviations

ssb, single-strand DNA-binding protein gene; SSB, single-stranded DNA-binding protein; ORF, open reading frame; OB, oligonucleotide/oligosaccharide-binding; ATCC, American Type Culture Collection; PCR, polymerase chain reaction; NCBI, National Center for Biotechnology Information; RBS, ribosomal binding site; SSB_{Dr}, *D. radiodurans* single-stranded DNA-binding protein; ssDNA, single-stranded DNA; css, circular single-stranded DNA; lds, linear double-stranded DNA; nc, nicked circular double-stranded DNA ; *Drad, D. radiodurans* strain R1; *Taq, Thermus aquaticus*; *TthHB8, Thermus thermophilus* strain HB8; *TthVK1, Thermus thermophilus* strain VK1; *Gmet, Geobacter metallireducens*; *Neur, Nitrosomonas europaea* ATCC 19718; *PaerPAO1, Pseudomonas aeruginosa* PAO1; *EcoliK12, Escherichia coli* strain K12.

Authors' contributions

J.M.E. predicted and confirmed the full length *D. radiodurans ssb* gene, identified one of two original sequencing errors in the *ssb* gene, performed the sequence alignments, purified the *D. radiodurans* SSB protein, determined the oligomeric state of the protein by analytical gel filtration, assisted in the sedimentation equilibrium analysis and drafted the manuscript. N.H. performed and quantitated the DNA three-strand exchange experiments. E.A.W. identified one of two original sequencing errors in the *D. radiodurans ssb* gene, participated in the confirmation of the full length *ssb* gene and cloned the *D. radiodurans ssb* gene in the over-expression vector. M.M.C. conceived of the study, participated in its design and coordination and edited the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors thank Dr. Darrell McCaslin, Director of the Biophysics Instrumentation Facility (BIF) at UW-Madison, for advice and assistance with the sedimentation equilibrium study. Funding for the establishment of the BIF was provided by the University of Wisconsin-Madison, grant BIR-9512577 from the National Science Foundation, and grant S10 RR13790 from the National Institutes of Health. We also thank John Battista for reading and commenting on early versions of this manuscript. The work reported herein was supported by grant GM52725 from the National Institutes of Health.

References

1. Battista JR: **Against all odds – the survival strategies of *deinococcus radiodurans*.** *Annu Rev Microbiol* 1997, **51:**203-224.
2. Makarova KS, Aravind L, Wolf YI, Tatusov RL, Minton KW, Koonin EV, Daly MJ: **Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics.** *Microbiol Mol Biol Rev* 2001, **65:**44-79.
3. Lohman TM, Ferrari ME: **Escherichia coli single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities.** *Annu Rev Biochemistry* 1994, **63:**527-570.
4. White O, Eisen JA, Heidelberg JF, Hickey EK, Peterson JD, Dodson RJ, Haft DH, Gwinn ML, Nelson WC, Richardson DL, Moffat KS, Qin HY, Jiang LX, Pamphile W, Crosby M, Shen M, Vamathevan JJ, Lam P, McDonald L, Utterback T, Zalewski C, Makarova KS, Aravind L, Daly MJ, Minton KW, Fraser CM *et al.*: **Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.** *Science* 1999, **286:**1571-1577.
5. Lipton MS, Pasa-Tolic L, Anderson GA, Anderson DJ, Auberry DL, Battista JR, Daly MJ, Fredrickson J, Hixson KK, Kostandarithes H, Masselon C, Markillie LM, Moore RJ, Romine MF, Shen Y, Stritmatter E, Tolic N, Udseth HR, Venkateswaran A, Wong KK, Zhao R, Smith RD: **Global analysis of the *Deinococcus radiodurans* proteome by using accurate mass tags.** *Proc Natl Acad Sci U S A* 2002, **99:**11049-11054.
6. Liu Y, Zhou J, Omelchenko MV, Beliaev AS, Venkateswaran A, Stair J, Wu L, Thompson DK, Xu D, Rogozin IB, Gaidamakova EK, Zhai M, Makarova KS, Koonin EV, Daly MJ: **Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation.** *Proc Natl Acad Sci U S A* 2003, **100:**4191-4196.
7. Dabrowski S, Olszewski M, Piatek R, Brillowska-Dabrowska A, Konopa G, Kur J: **Identification and characterization of single-stranded-DNA-binding proteins from Thermus thermophilus and *Thermus aquaticus* – new arrangement of binding domains.** *Microbiology* 2002, **148:**3307-3315.
8. Ma J, Campbell A, Karlin S: **Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures.** *J Bacteriol* 2002, **184:**5733-5745.
9. Meima R, Lidstrom ME: **Characterization of the minimal replicon of a cryptic *Deinococcus radiodurans* SARK plasmid and development of versatile *Escherichia coli-D.radiodurans* shuttle vectors.** *Appl Environ Microbiol* 2000, **66:**3856-3867.
10. Meima R, Rothfuss HM, Gewin L, Lidstrom ME: **Promoter cloning in the radioresistant bacterium *Deinococcus radiodurans*.** *J Bacteriol* 2001, **183:**3169-3175.
11. Murzin AG: **OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences.** *EMBO J* 1993, **12:**861-867.
12. Chedin F, Seitz EM, Kowalczykowski SC: **Novel homologs of replication protein A in archaea: implications for the evolution of ssDNA-binding proteins.** *Trends Biochem Sci* 1998, **23:**273-277.
13. Weiner JH, Bertsch LL, Kornberg A: **The deoxyribonucleic acid unwinding protein of *Escherichia coli*. Properties and functions in replication.** *J Biol Chem* 1975, **250:**1972-1980.
14. Raghunathan S, Kozlov AG, Lohman TM, Waksman G: **Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA.** *Nat Struct Biol* 2000, **7:**648-652.
15. Curth U, Genschel J, Urbanke C, Greipel J: **In vitro and in vivo function of the C-terminus of *Escherichia coli* single-stranded DNA binding protein.** *Nucleic Acids Res* 1996, **24:**2706-2711.
16. Matsumoto T, Morimoto Y, Shibata N, Kinebuchi T, Shimamoto N, Tsukihara T, Yasuoka N: **Roles of functional loops and the C-terminal segment of a single-stranded DNA binding protein elucidated by x-ray structure analysis.** *J Biochem* 2000, **127:**329-335.
17. Kowalczykowski SC, Krupp RA: **Effects of *Escherichia coli* SSB protein on the single-stranded DNA-dependent ATPase activity of *Escherichia coli* RecA protein. Evidence that SSB protein facilitates the binding of RecA protein to regions of secondary structure within single-stranded DNA.** *J Mol Biol* 1987, **193:**97-113.
18. Lavery PE, Kowalczykowski SC: **A postsynaptic role for single-stranded DNA-binding protein in recA protein-promoted DNA strand exchange.** *J Biol Chem* 1992, **267:**9315-9320.
19. Ganesh N, Muniyappa K: **Characterization of DNA strand transfer promoted by *Mycobacterium smegmatis* RecA reveals functional diversity with *Mycobacterium tuberculosis* RecA.** *Biochemistry* 2003, **42:**7216-7225.

20. Alani E, Thresher R, Griffith JD, Kolodner RD: **Characterization of DNA-binding and strand-exchange stimulation properties of y-RPA, a yeast single-strand-DNA-binding protein.** *J Mol Biol* 1992, **227:**54-71.
21. Namsaraev EA, Berg P: **Rad51 uses one mechanism to drive DNA strand exchange in both directions.** *J Biol Chem* 2000, **275:**3970-3976.
22. Kim JI, Sharma AK, Abbott SN, Wood EA, Dwyer DW, Jambura A, Minton KW, Inman RB, Daly MJ, Cox MM: **RecA protein from the extremely radioresistant bacterium *Deinococcus radiodurans:* Expression, purification, and characterization.** *J Bacteriol* 2002, **184:**1649-1660.
23. Lohman TM, Chao K, Green JM, Sage S, Runyon GT: **Large-scale purification and characterization of the *Escherichia coli* rep gene product.** *J Biol Chem* 1989, **264:**10139-10147.
24. Marrione PE, Cox MM: **RuvB protein-mediated ATP hydrolysis: functional asymmetry in the RuvB hexamer.** *Biochemistry* 1995, **34:**9809-9818.
25. Shan Q, Cox MM, Inman RB: **DNA strand exchange promoted by RecA K72R. Two reaction phases with different Mg2+ requirements.** *J Biol Chem* 1996, **271:**5712-5724.
26. Lusetti SL, Wood EA, Fleming CD, Modica MJ, Korth J, Abbott L, Dwyer DW, Roca AI, Inman RB, Cox MM: **C-terminal deletions of the *Escherichia coli* RecA protein – Characterization of in vivo and in vitro effects.** *J Biol Chem* 2003, **278:**16372-16380.
27. Craig NL, Roberts JW: **Function of nucleoside triphosphate and polynucleotide in *Escherichia coli* recA protein-directed cleavage of phage lambda repressor.** *J Biol Chem* 1981, **256:**8039-8044.
28. Lohman TM, Overman LB: **Two binding modes in *Escherichia coli* single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration.** *J Biol Chem* 1985, **260:**3594-3603.
29. Messing J: **New M13 vectors for cloning.** *Methods Enzymol* 1983, **101:**20-78.
30. Neuendorf SK, Cox MM: **Exchange of RecA protein between adjacent RecA protein-single-stranded DNA complexes.** *J Biol Chem* 1986, **261:**8276-8282.
31. Haruta N, Yu X, Yang S, Egelman EH, Cox MM: **A DNA pairing-enhanced conformation of bacterial RecA proteins.** *J Biol Chem* 2003, **278:**52710-52723. (JBC papers in press – M308563200)
32. Raghunathan S, Ricard CS, Lohman TM, Waksman G: **Crystal structure of the homo-tetrameric DNA binding domain of *Escherichia coli* single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-A resolution.** *Proc Natl Acad Sci U S A* 1997, **94:**6652-6657.